
TAPE resources at the ATLAS

— T. Javurek on behalf of ADC —

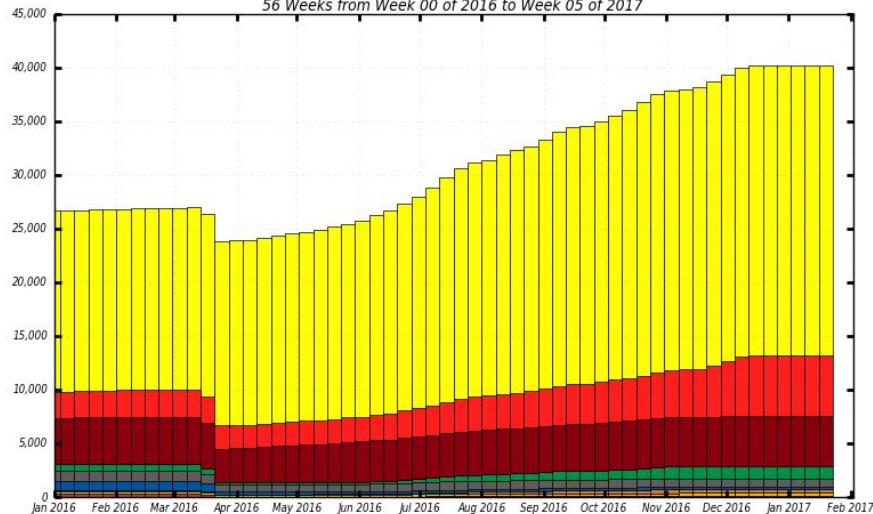
How ADC uses tapes

- Writing
 - RAW export from T0 (according to MoU share)
 - Archiving AOD, ESD, DESD, DRAW, RDO, HITS (according to free space)
 - No direct writing from jobs
- Reading
 - Organised reprocessing campaigns
 - Chaotic analysis (in case that datasets are not available on DISK)
 - Data is always copied to DATADISK, never read from buffer
- Deleting
 - Obsolete data is never removed automatically from tapes (safety feature)
 - Cleaning campaigns are run manually periodically
 - Recently we have done it once per year in spring before the LHC starts

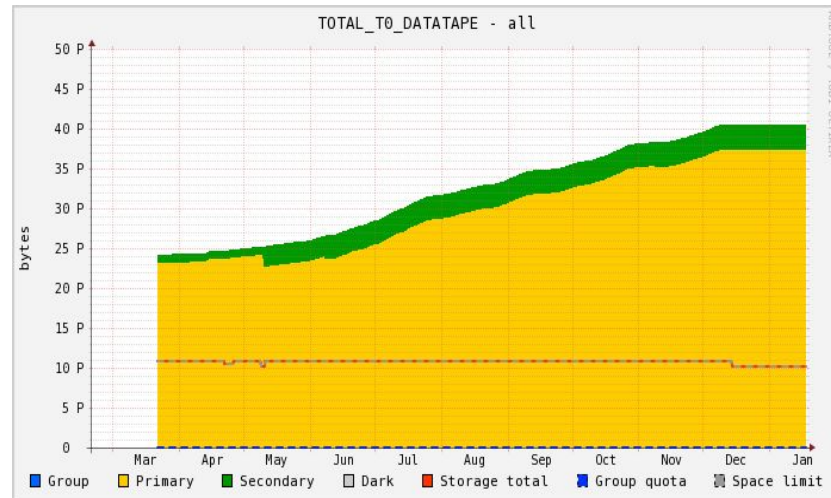
Overview - CERN tape



Number of Physical Bytes (in TBs)
56 Weeks from Week 00 of 2016 to Week 05 of 2017



Maximum: 40,207 , Minimum: 0.00 , Average: 30,526 , Current: 40,207

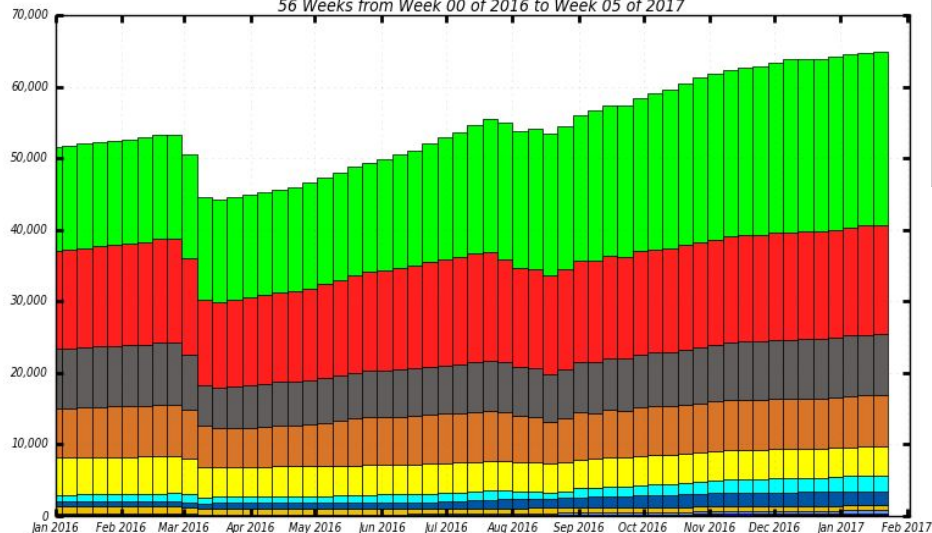


- Fastest growth: 2.4 PB/month ⇒ max. 28 PB in 2017
- 2 PB of obsolete data

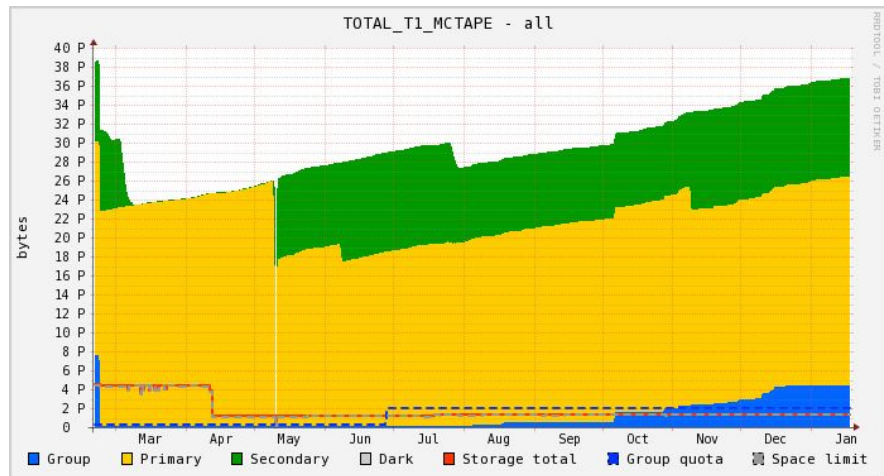
Overview - t1-tapes



Number of Physical Bytes (in TBs)
56 Weeks from Week 00 of 2016 to Week 05 of 2017



Maximum: 64,815 , Minimum: 0.00 , Average: 53,640 , Current: 64,815

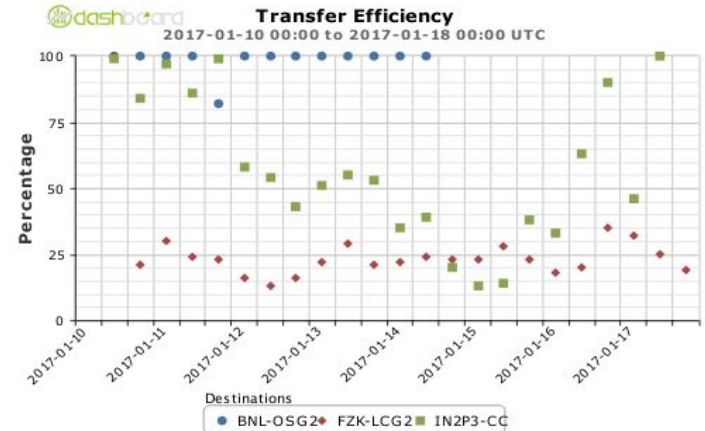
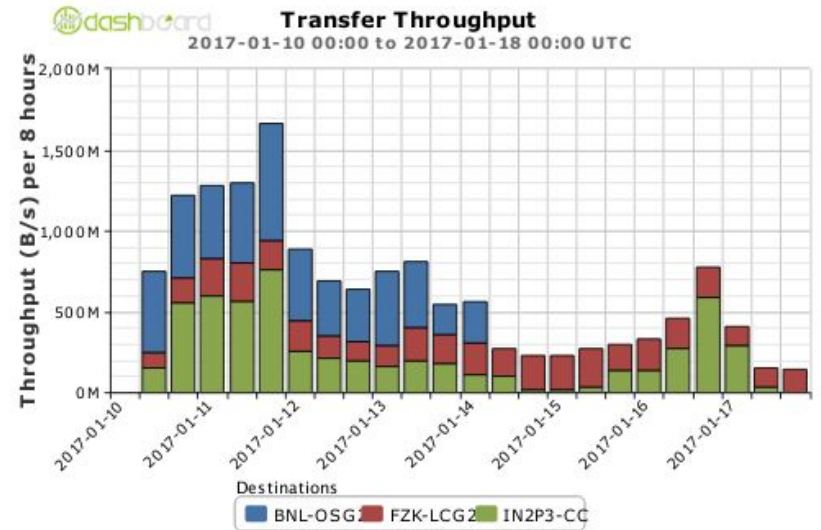


- Fastest growth in 2016: 2.6 PB/month \Rightarrow 32 PB in 2017 to all t1 tapes
- Deletion foreseen: 13 PB
- 65 PB = 60 mil. files \Rightarrow avg. = \sim 1GB/file (also reported by BNL) discussed later \downarrow

Staging tests

BNL, FZK, IN2P3-CC

- TAPE → DISK at given site
- 150 TBs of AODs for each of the sites
- 310k files ~ 1.5 GB/file
- Results:
 - BNL: 4 days ⇒ 430 MB/sec.
 - IN2P3-CC: 7 days ⇒ 250 MB/sec.
 - FZK: 100MB/7 days (ongoing) ⇒ 165 MB/sec.
- FZK+BNL+IN2P3-CC ~ 50% MoU ⇒
 - (430+250+165) × 2 ~ 1.7 GB/sec from all tapes ⇒
 - 7 days to stage 1 PB (optimistic estimate)



Feedback from sites - ATLAS 2017 pledges

Do you see any problem with fulfilling ATLAS 2017 pledges?

IN2P3-CC (17.8 PB ✓), NDGF-T1 (9 PB ✓), RAL-LCG2 (22.75 PB ✓), TRIUMF (18.8 PB - ✓ from 1st April), INFN-T1 (16.9 PB ✓), FZK-LCG2 (22.1 PB ✓), PIC (8.8 PB ✓ from mid-2017), BNL (44.4 PB ✓), CERN (77 PB ✓)

SARA-MATRIX (11.7 PB), RRC-KI-T1 (3 PB) ... no response

Total t1s pledge = 175 PB, used ~ 70 PB, + 32 PB in 2017

Total t0 pledge = 77 PB, used ~ 40 PB, + 28 PB in 2017

Feedback from sites - repacking

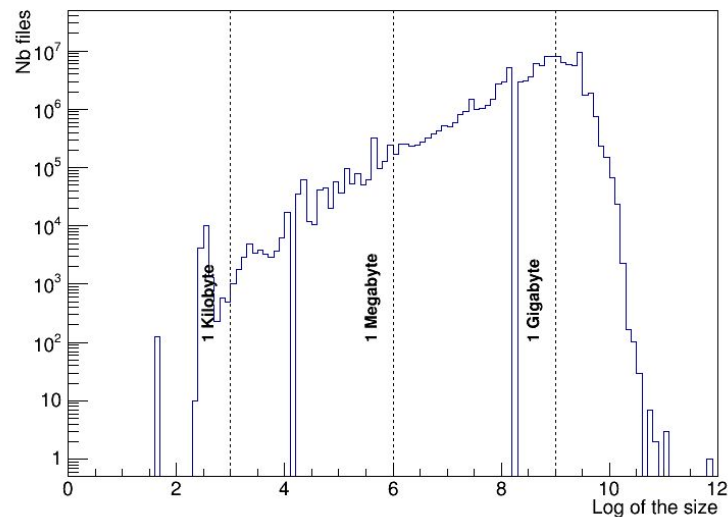
How often do you plan to proceed repacking campaign?

- Continuous repacking/When needed: NDGF-T1, PIC, BNL, CERN
- Major repacks:
 - soon: IN2P3-CC
 - mid-2017: CERN, INFN-T1 (adding new drives)
 - later in 2017: RAL-LCG2, FZK-LCG2, INFN-T1
 - will be announced: TRIUMF, FZK-LCG2

Feedback from sites - discussion

Is there anything that you wish that ATLAS will improve in the workflow involving TAPES?

- **File size**
- **Requests in bulks - organized r/w**
- **Massive write lead to errors**
- **Deletion soon (IN2P3C) and more frequently (repacks)**
- **Announce large requests**
- **publish foreseen tape usage for 2017**



Thanks for attention!
Thanks for feedbacks!

Feedback from sites - detail

Is there anything that you wish that ATLAS will improve in the workflow involving TAPes?

IN2P3-CC: massive staging requests (~PBs in few days) lead to timeouts.

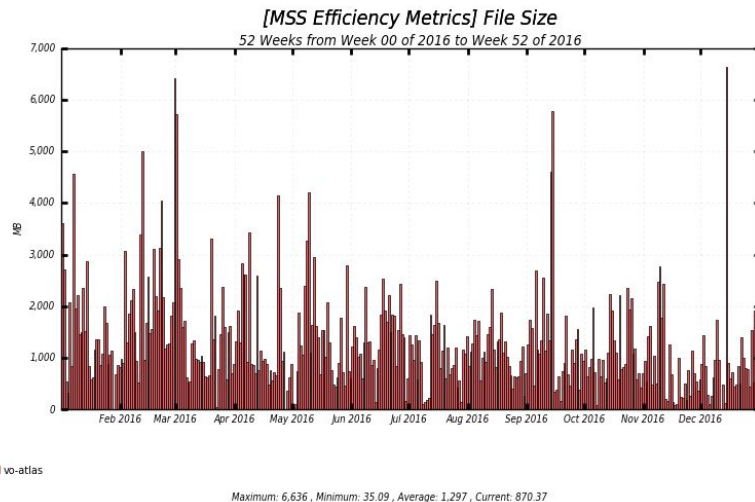
NDGF-T1: more concurrent requests (?), restrict cp to disk when file is already about to be copied there (?)

RAL: tape families (?), publish foreseen tape usage for 2017, What do we write there?

TRIUMF: read in bulks e.g. in one go if large derivation campaign foreseen (organized staging); file size; massive write (~PBs) lead to errors,

Feedback from sites - detail

PIC: file size - fig. from Pepe.



Feedback from sites - detail

INFN: prediction of write and read; not to write short-lifetime data (we don't do)

FZK: publish expected throughputs; larger file size

BNL: larger file size (5 GB avg. (1 GB now));

CERN: larger file size (fig.); organized writing

