

Setup and numbers

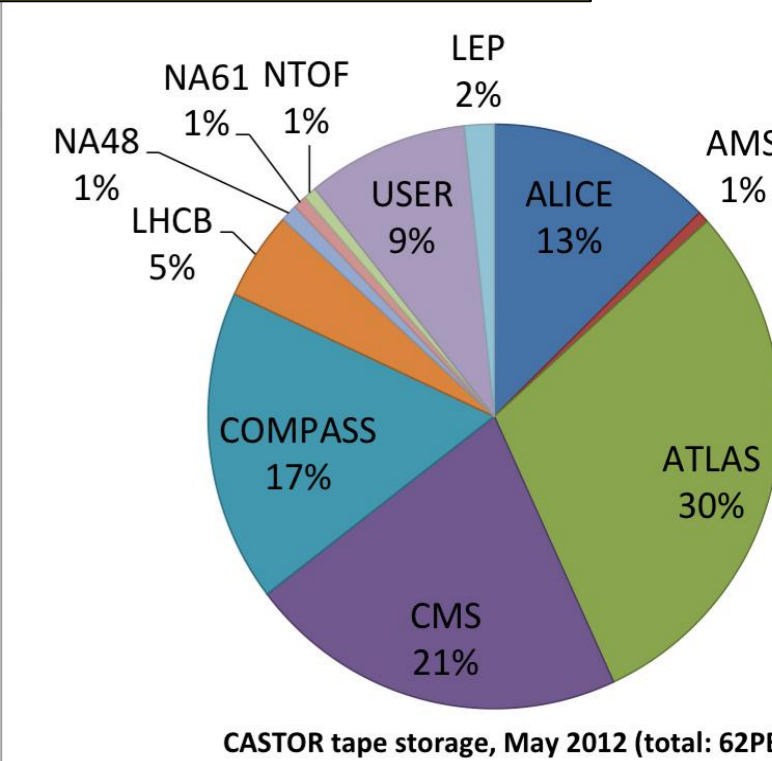
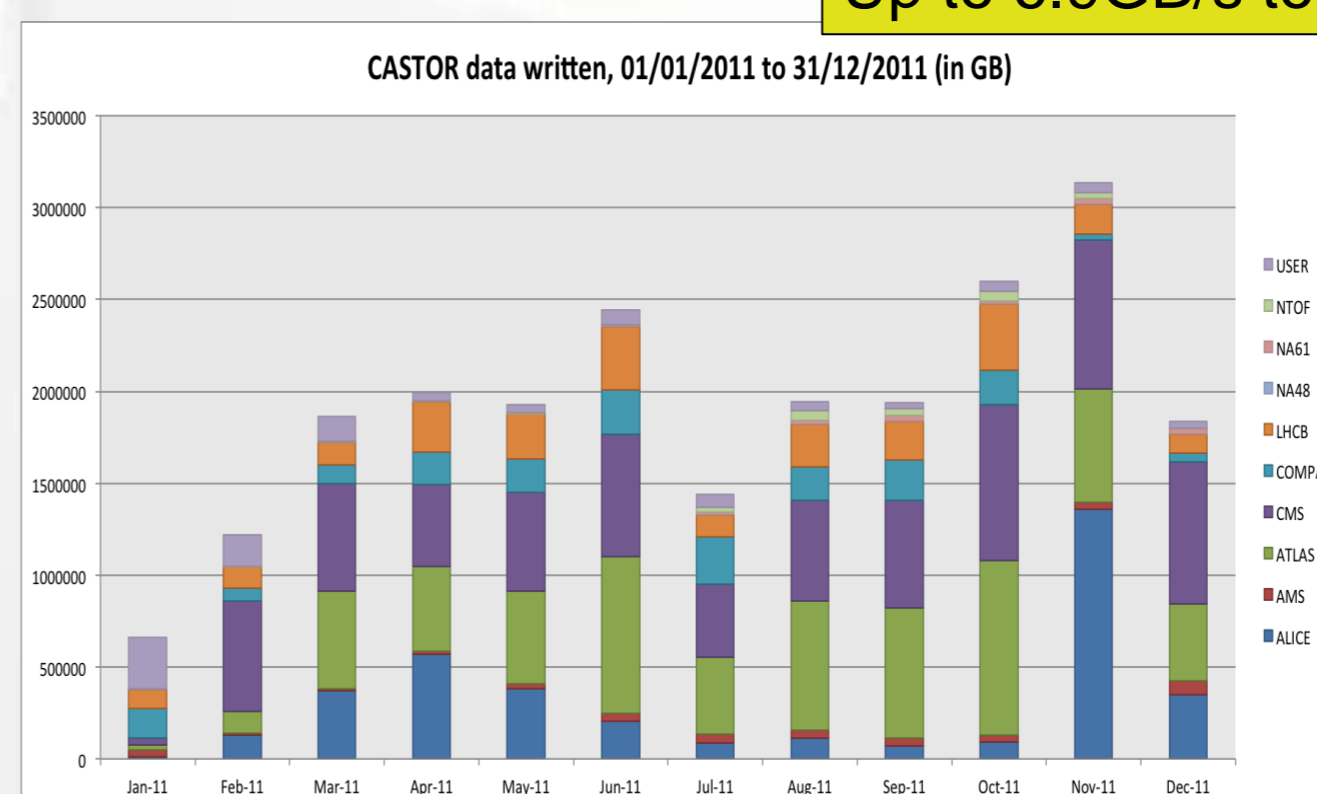
62PB data in 208M files and 52K tapes
1.5 .. 3PB new data per month
Up to 6.9GB/s to tape during HI period



4x Oracle SL8500
40x T10KC drives
+ legacy T10KBs

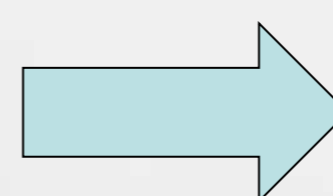


5x IBM TS3500
40x TS1140 drives
+ legacy TS1130s



From nearline to archiving...

Tape technology is **evolving**...
Increased per-tape capacity and transfer speed
Large-scale archive market building up (legal, media, cloud providers)
... but library and drive mount/positioning times **not** improving!

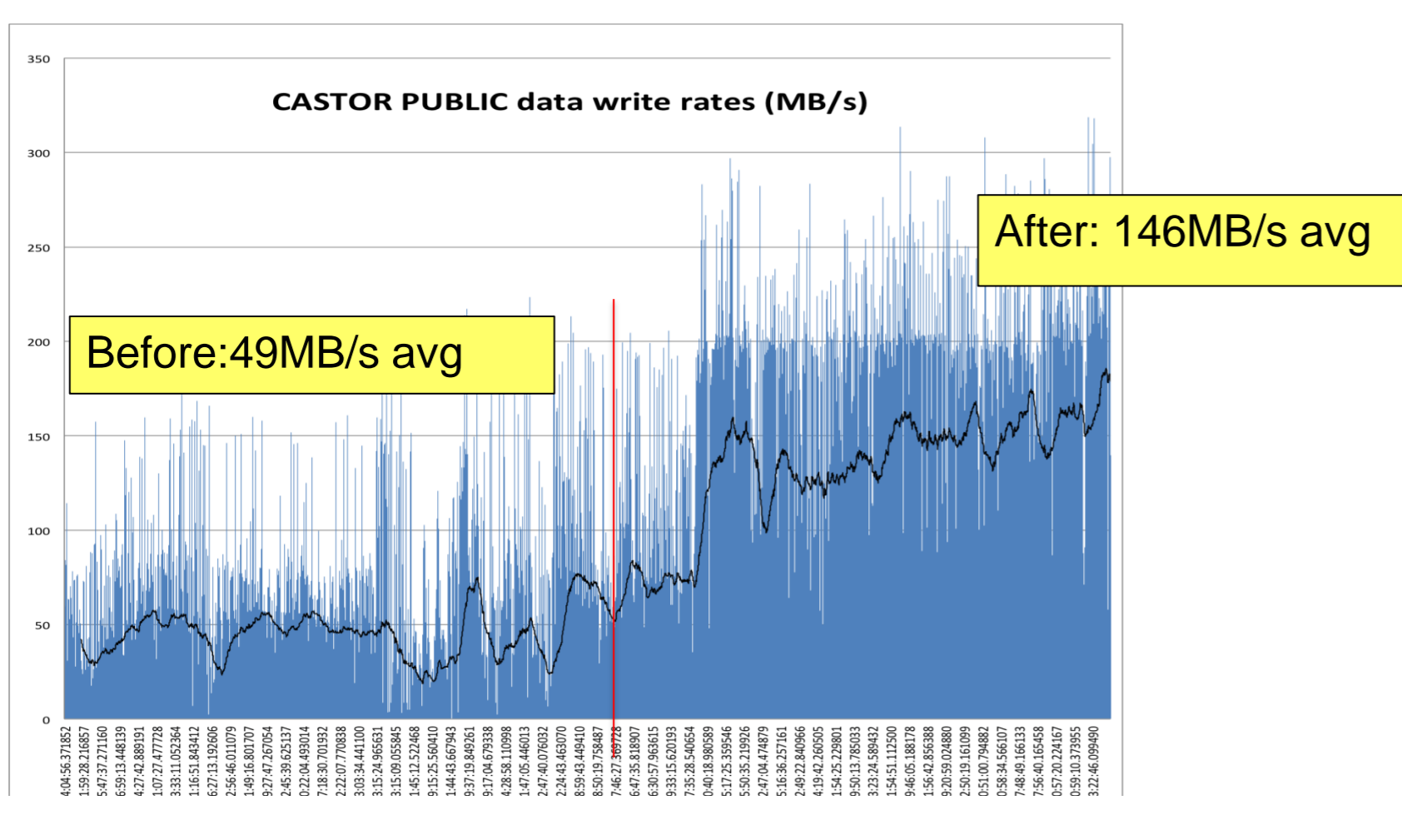


Speed up streaming access and
minimise file/user based HSM

a) boost write streaming

Full deployment of 'buffered' tape marks (cf. poster 415, S. Murray)

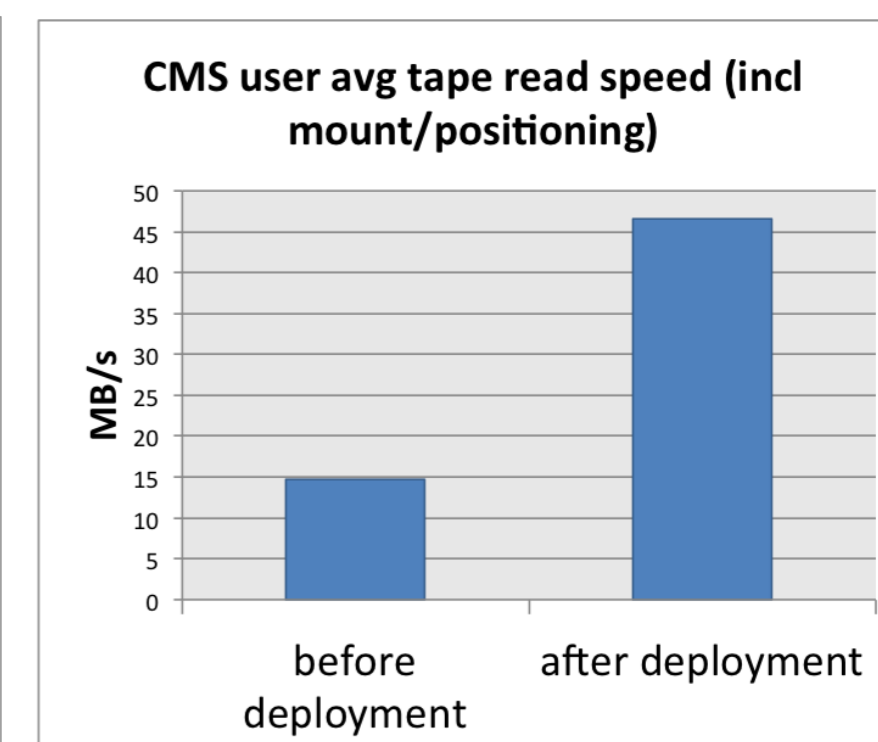
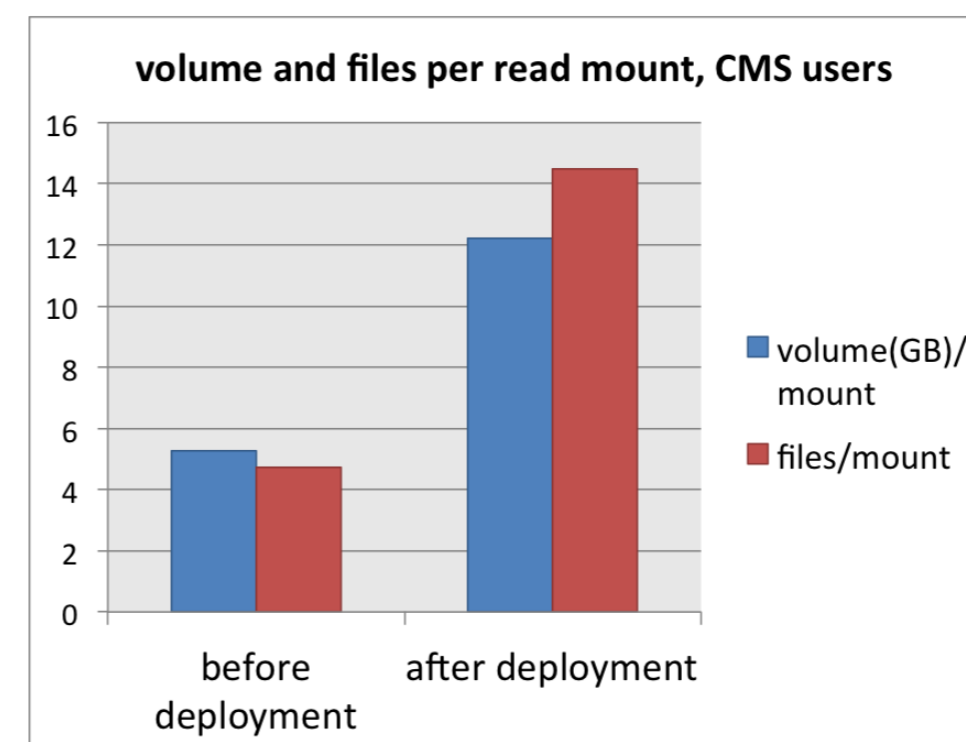
Early 2011, reduced sync tape marks from 3 to 1 per CASTOR file
speed increase from ~15MB/s to ~40MB/s
Achieve near-native drive speed by writing data in effective chunks of ~32GB
eliminates "backhitch" (tape repositioning) inefficiencies



b) optimise data retrieval

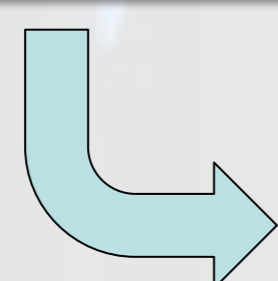
Deployed "traffic lights" and "bus lanes" for prioritising bulk read requests

Access to robotics and drives is the bottleneck
Define thresholds for tape mounts:
minimum volume, max wait time, concurrent drive usage; group related tape requests
Different thresholds + priorities for production coordinators and end users

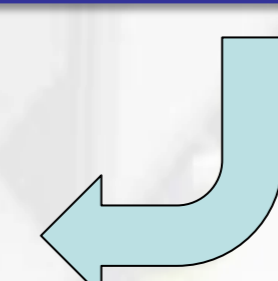


3x files / volume per mount -> 3x increase in effective tape access speed
~50% less tape mounts (~7K to 3.5K mounts per day)

Additional ~30% less tape mounts expected when moving CMS end users to EOS



higher throughput and efficiency allowing substantial
decrease of production tape drives (from 120 to 80)



... now evaluating commodity tape

Tape usage at CERN was heavy-duty requiring enterprise-class tape equipment from IBM and Oracle
With far less demand in terms of small file writes and read mounts, "commodity" tape (LTO)
becomes a serious option, i.e. for "dusty" archived data which is infrequently accessed
Market share: LTO (~90%) vs. enterprise media (~2%)

Since Nov 2011, testing a SpectraLogic T-Finity library (max 120 drives, 30K slots)

Test configuration: 5-10 LTO-5 drives, 1000 cartridges (1.5PB)

Necessary CASTOR adaptations coded and released

Field test drives, library, and vendor support – storing 2nd copies of experiment data

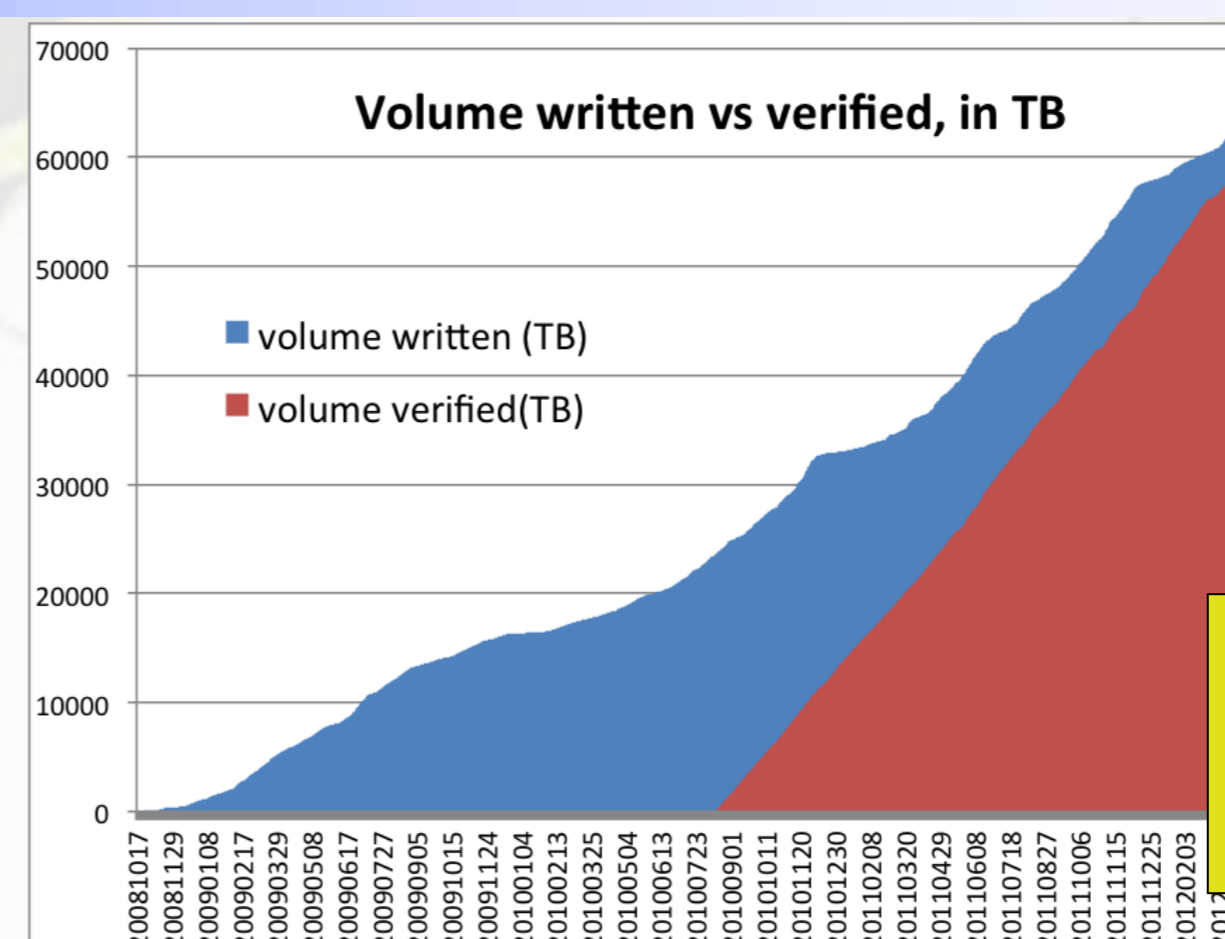
So far satisfactory results in general



Active monitoring and verification

- Automated handling of tape, drive and library malfunction
- User activity is monitored and unusual/inefficient usage automatically reported
- "bus lane" concept encourages users to pre-stage data

STAGER USER:GROUP	#MOUNTS	#TAPES	RATIO	Avg#FILES	AvgFILESIZE(MB)
PUBLIC vkxxxx:vy	2167	66	32.8	1.3	1047
ALICE a1xxxx:z2	834	85	9.8	1.7	4305
LHCb saxxxxx:z5	285	5	57	1.4	160
CMS mpxxxx:zh	119	2	59.5	1	202



Archive verification:
re-read
• all newly filled tapes
• tapes not seen recently
turnaround time ~2 years
speed ~ 3PB / month

Reliability (approximate!):
~0.0175% unavailable (under repair)
~0.000062% data loss (37.2GB)