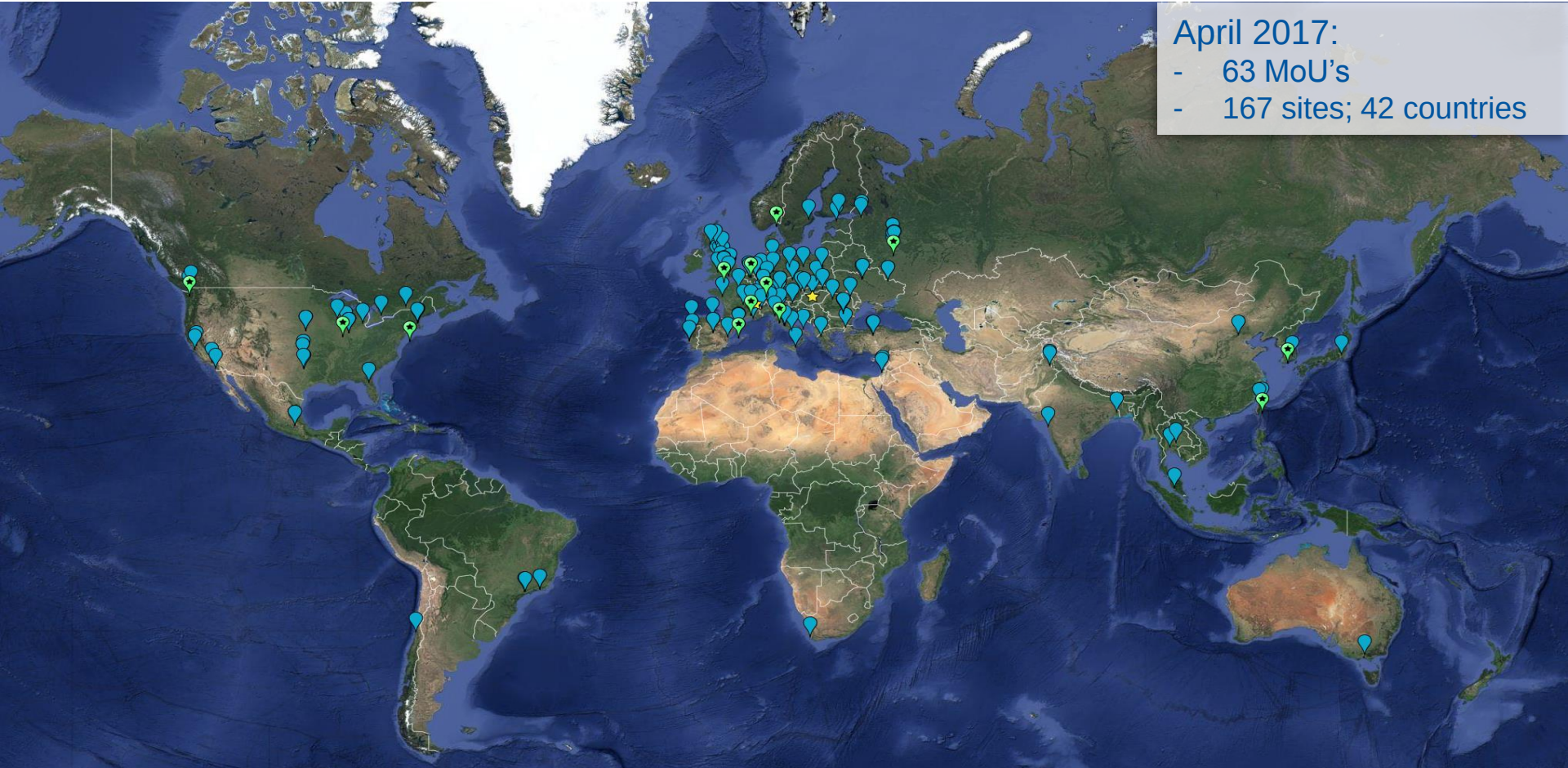
WLCG Status Report

Ian Bird

Computing RRB

CERN, 25th April 2017

WLCG Collaboration



WLCG Funding & Expenditure

LHC Future Computing Funding and Expenditure Estimates

(all figures in MCHF), data extracted 12th April 2017

	2016	2017	2018	2019	2020
Funding					
From CERN Budget *					
- Personnel	17.9	19.2	19.7	19.0	19.5
- Materials **	21.6	19.5	21.0	19.6	18.6
Total Funding	**39.5**	**38.6**	**40.6**	**38.6**	**38.1**
Expenditure					
- Personnel ***	**19.0**	**18.9**	**19.3**	**18.8**	**18.7**
- Materials	**22.8**	**21.8**	**17.3**	**19.1**	**15.9**
Total Planned Expenditure	**41.8**	**40.7**	**36.7**	**37.8**	**34.6**
Balance Personnel	**-1.2**	**0.2**	**0.3**	**0.3**	**0.8**
Balance Materials	**-1.1**	**-2.3**	**3.6**	**0.5**	**2.7**

* Internal budget 2017

** Includes carry-forward/carry-back, EUR/CHF exchange rate penalty applied since 2015 and negative CVI for 2017 and beyond

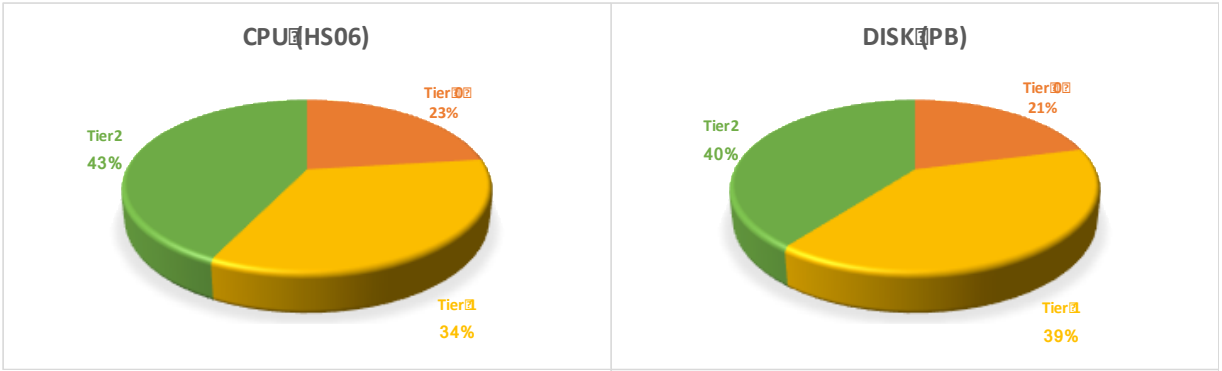
*** Excluding Data Centre Operators

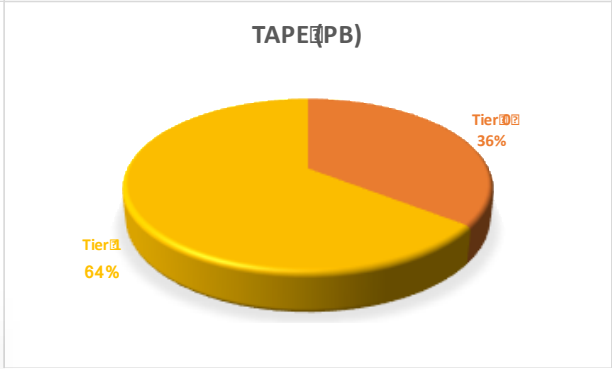
Personnel: balanced situation

Materials planning based on currently understood parameters:

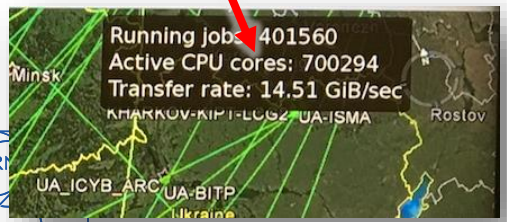
- Assumes 2018 requirements as discussed today;
  - LS2: 2019 - 2020
- Cost extrapolations based on recent experience
- Includes cost of additional purchases for 2017 at end of 2016
- Foresee 2nd network hub costs; internal network upgrades; tape infrastructure updates
- Positive balance in 2020 is preparation for ramp up to Run 3

Pledged resources 2017



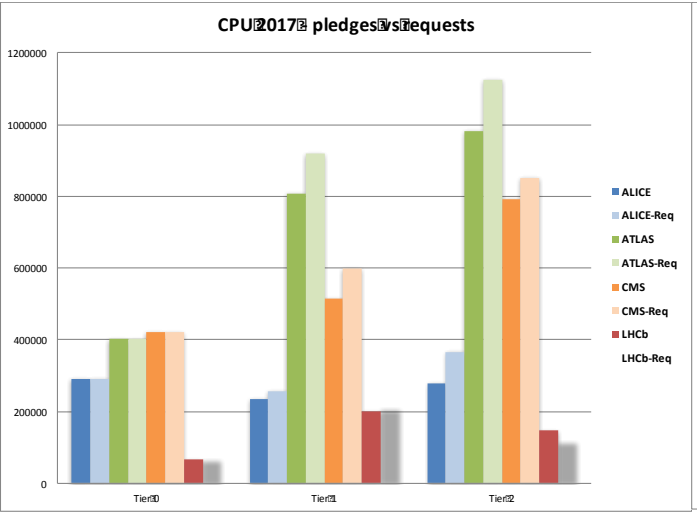


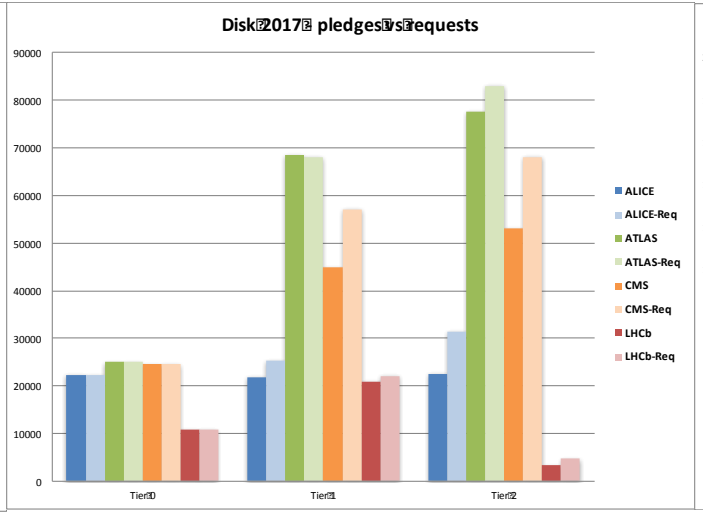
- 5.2 M HS06
  - >500 K cores (if bought today)
  - Actually many more

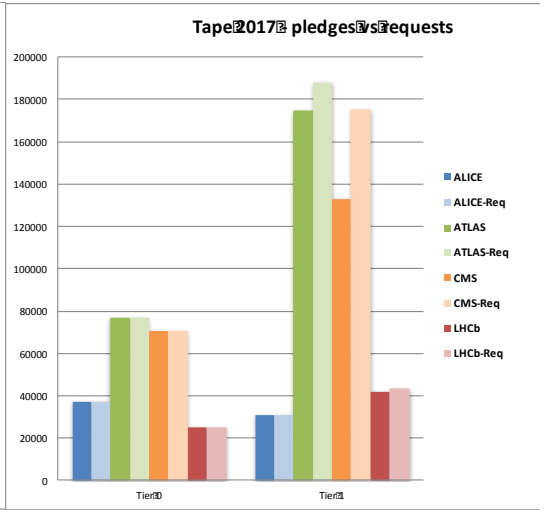


- 985 PB Storage
  - 395 PB disk
  - 590 PB tape

2017 Pledge situation





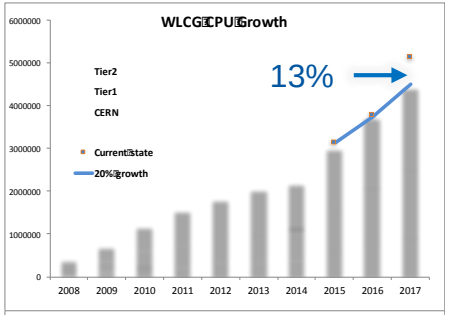


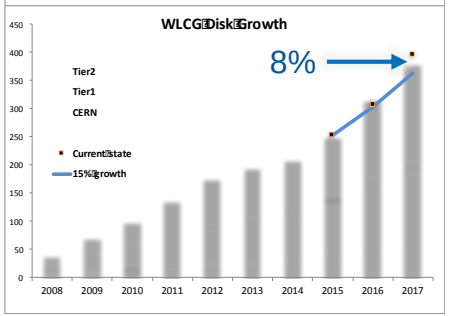
Not all is deployed yet for 2017 – a few delays
Full resources expected by June

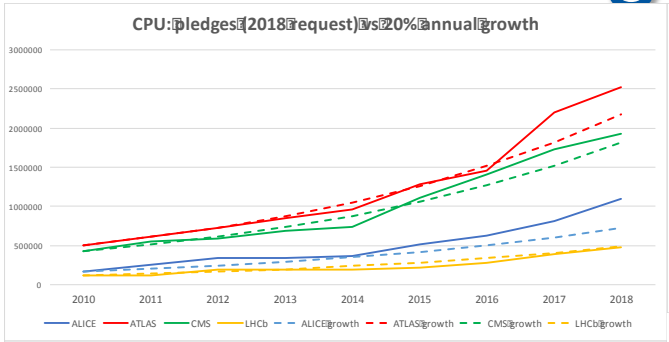
Mitigation measures reviewed by LHCC

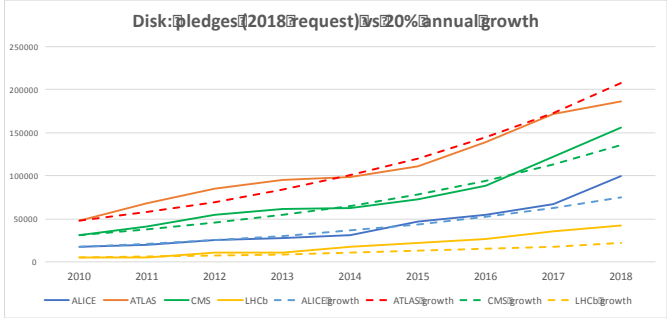
- In February the LHCC reviewed the measures taken by the experiments to mitigate the shortfall in resources relative to the exceptional LHC performance
- Concluded that: (CERN-LHCC-2017-004)
  - “The **LHCC congratulates** the LCG and experiments on the successful implementation of mitigation measures to cope with the increased data load. “
  - “The **LHCC notes** that the margins to reduce the resource usage in the short term without impact on physics have been exhausted. “

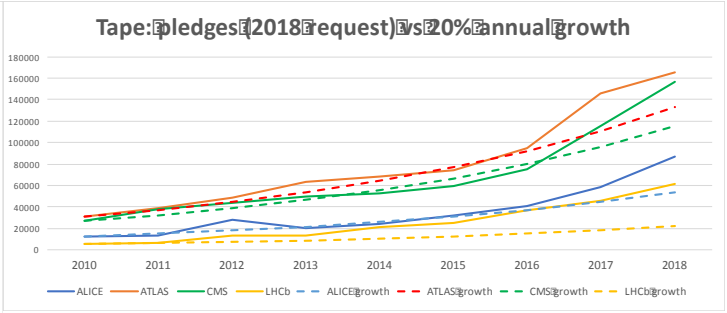
Comments on flat budgets









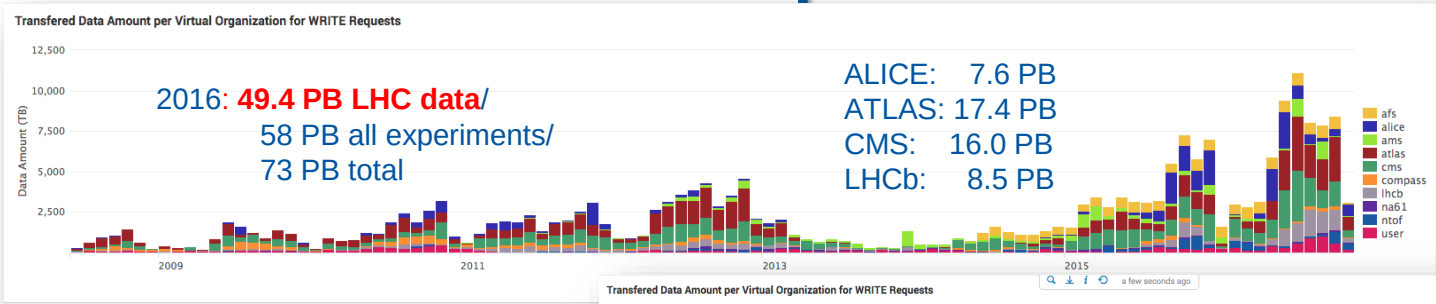


Extrapolations from 2010:

- Ignore no investment in 2013,14
- Deviations from “flat budget” are generally not enormous, and are corrected
- Jump in 2017 – LHC performance
- Tape needs still increase

- We need to clarify what is meant by flat budgets:
  - We assume: constant budget/investment even in long shutdown years
  - This did not happen in LS1

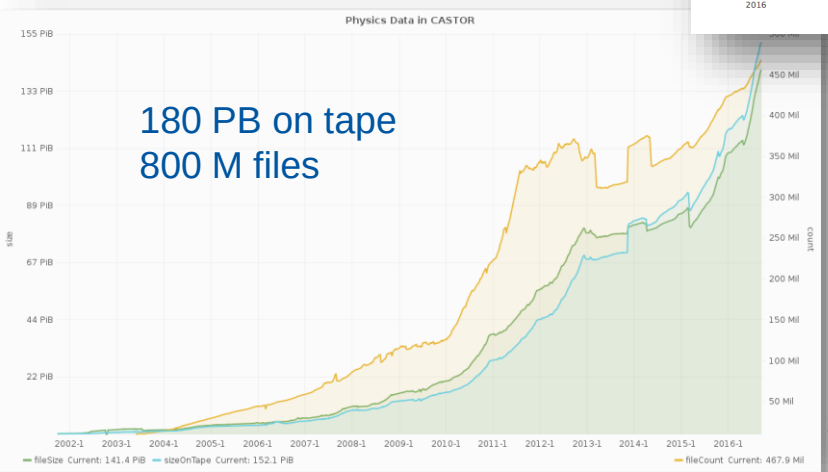
Data in 2016 - updated



2016: **49.4 PB LHC data**/
58 PB all experiments/
73 PB total

ALICE: 7.6 PB
ATLAS: 17.4 PB
CMS: 16.0 PB
LHCb: 8.5 PB

11 PB in July

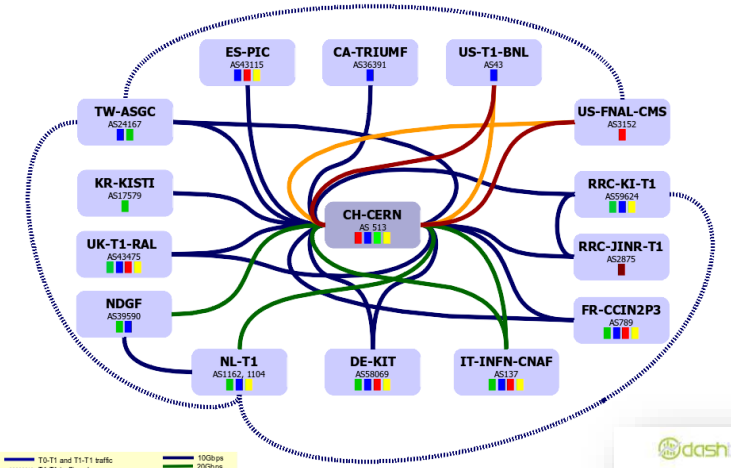


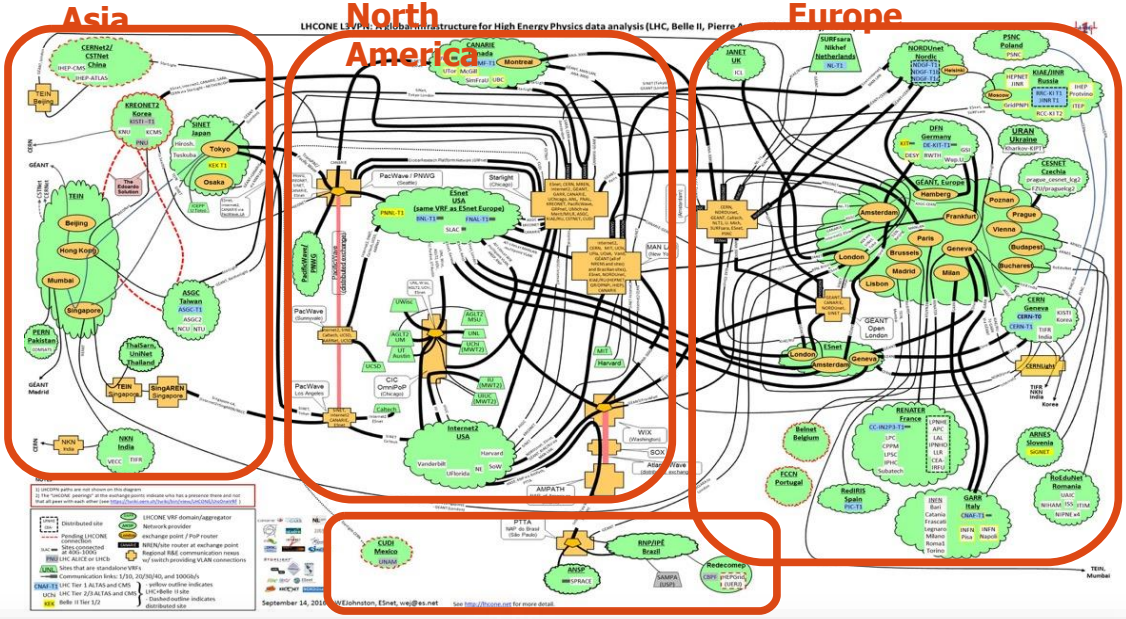
180 PB on tape
800 M files

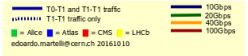
2017

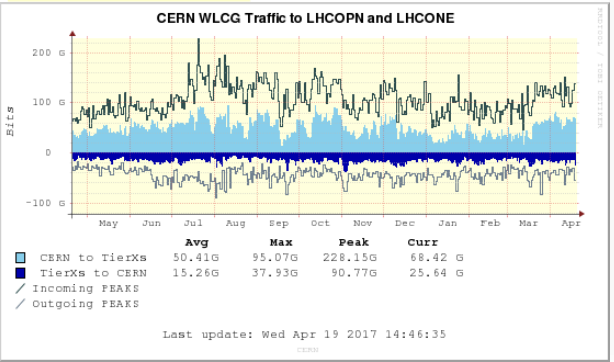
Data transfers

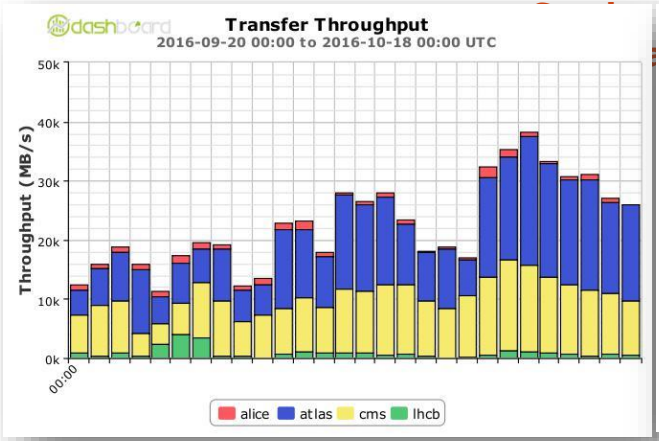
LHCOPN

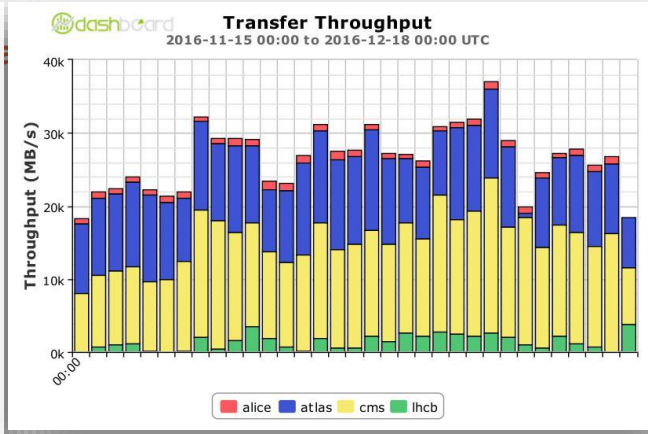




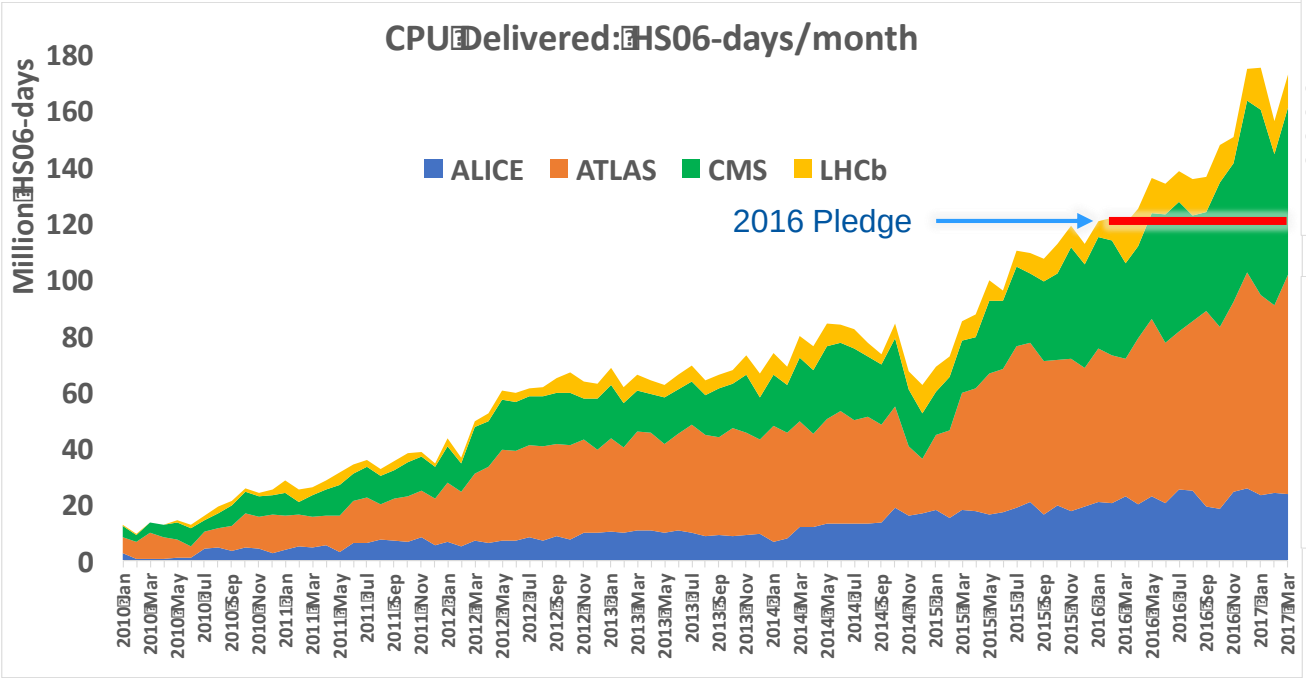


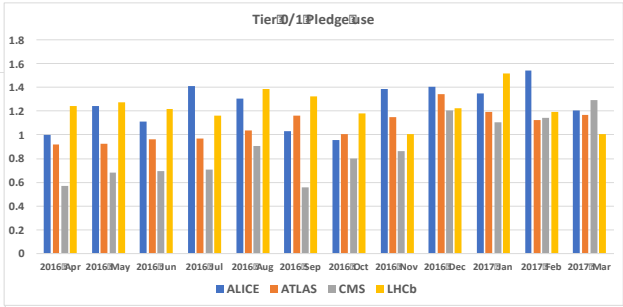


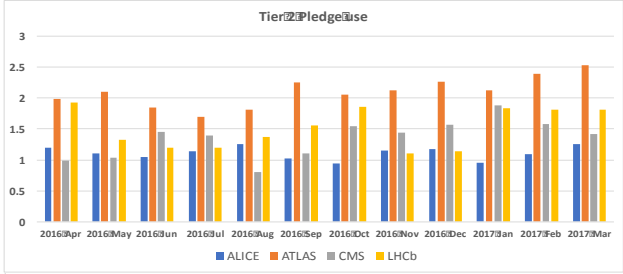




CPU Delivered



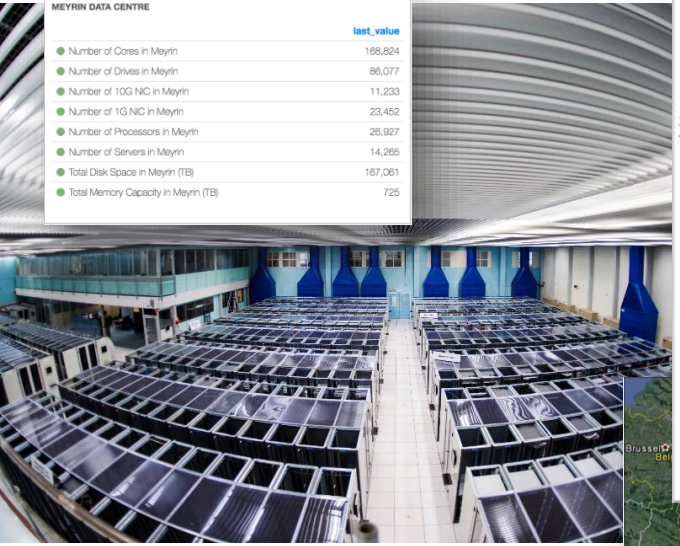




New peak: ~180 M HS06-days/month

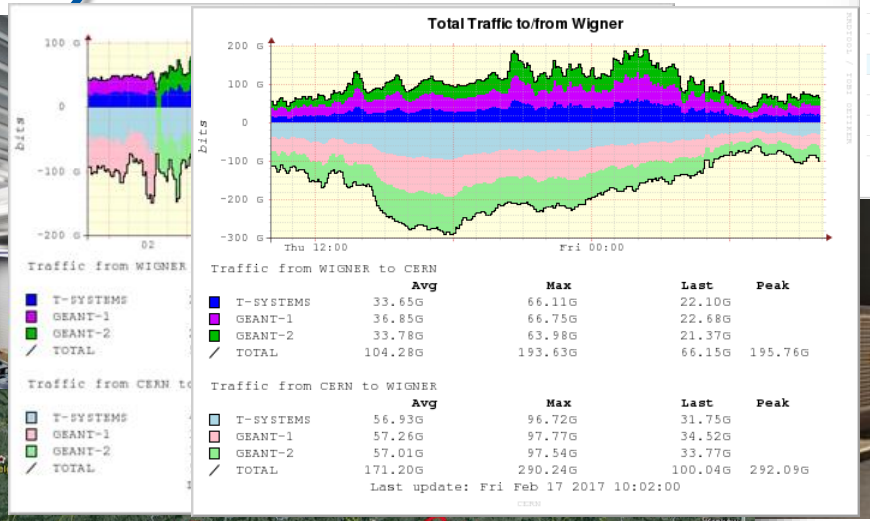
~ 600 k cores continuous

CERN (Tier 0) Facilities



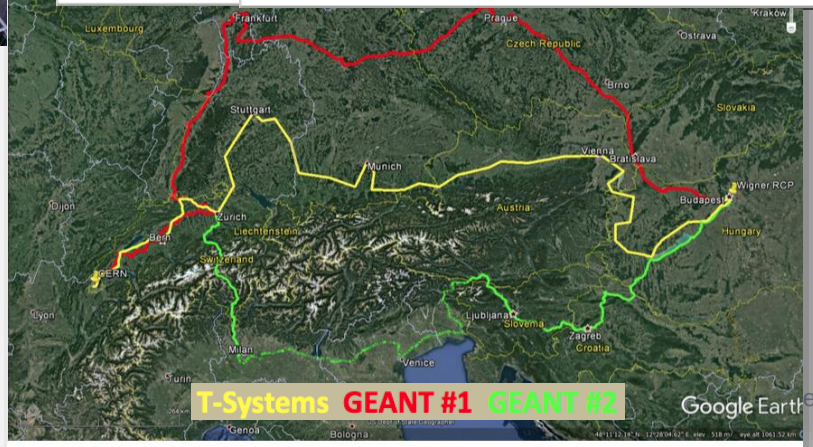
MEYRIN DATA CENTRE

	last_value
Number of Cores in Meyrin	168,824
Number of Drives in Meyrin	86,077
Number of 10G NIC in Meyrin	11,233
Number of 1G NIC in Meyrin	23,452
Number of Processors in Meyrin	26,927
Number of Servers in Meyrin	14,265
Total Disk Space in Meyrin (TB)	167,061
Total Memory Capacity in Meyrin (TB)	725



WIGNER DATA CENTRE

	last_value
Number of Cores in Wigner	56,000
Number of Drives in Wigner	29,694
Number of 10G NIC in Wigner	2,981
Numer of 1G NIC in Wigner	6,579
Number of Processors in Wigner	7,002
Number of Servers in Wigner	3,504
Total Disk Space in Wigner (TB)	97,315
Total Memory Capacity in Wigner (TB)	221



2017:

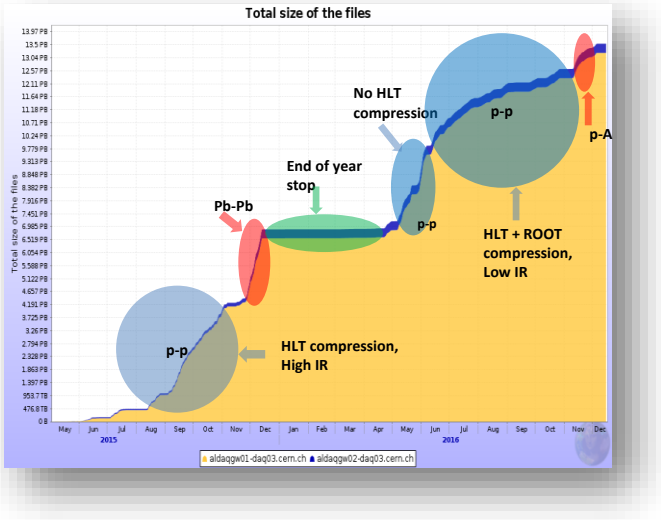
- 225k cores → 325k
- 150 PB raw → 250 PB

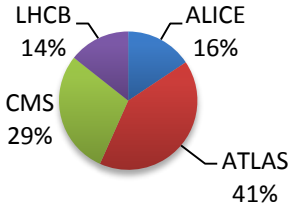
2017-18/19

- Upgrade internal networking capacity
- Refresh tape infrastructure

Experiment updates

ALICE: Run 2 progress and processing status





Data on tape at CERN in 2016

Data taking plan follows strictly the approved physics programme

- Better data quality by limiting the Interaction Rate
- Reducing RAW data volume by HLT data compression
  - RAW data size reduction by factor ~4.3 to ~5.5
  - New Huffman algorithm - 10-20% reduction of TPC event size

- Data processing – new high-precision calibration schema developed and certified
  - Up to 130K concurrent jobs, 80% overall CPU/wall efficiency
  - HLT providing 5% of the CPU resources
  - 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
- Analysis
  - Unprecedented level of analysis activity, up to 3PB data analyzed daily
- Simulation
  - Validation and performance improvements of Geant4 and its use for specific simulation campaigns

ALICE: Concerns

- Requirements
  - Growth within acceptable rate since 2014, all requests approved
- Pledges and delivered
  - Systematic under-pledges and under-delivery at T2s over several years
  - This artificially skews our 2018 requirements/pledges ratio
- Radical measures already taken
  - Reducing replicas, now at absolute minimum (no replicas for any file but AODs)

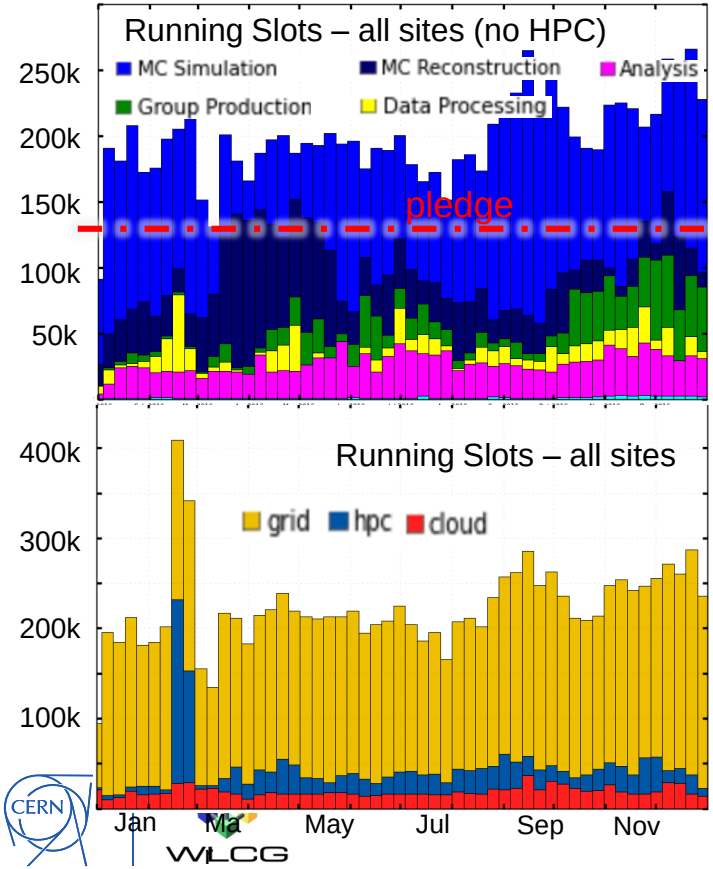
- Recommendations of CRSG ***not followed***, resulting in deficit of disk storage
  - A negative message will further erode the resources delivery
- 2018 will be particularly data intensive - pp and PbPb runs with high multiplicity central triggers
- **To successfully process and analyze the data we need our requirements to be fulfilled**

Tier	Pledge Type	ALICE	Required	Balance
Tier 0	CPU (HEP-SPEC06)	292,000	292,000	0%
Tier 0	Disk (Tbytes)	22,400	22,400	0%
Tier 0	Tape (Tbytes)	36,900	36,900	0%
Tier 1	CPU (HEP-SPEC06)	235,481	256,000	-8%
Tier 1	Disk (Tbytes)	21,808	25,400	-14%
Tier 1	Tape (Tbytes)	30,611	30,900	-1%
Tier 2	CPU (HEP-SPEC06)	277,660	366,000	-24%
Tier 2	Disk (Tbytes)	22,537	31,400	-28%

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	

ATLAS

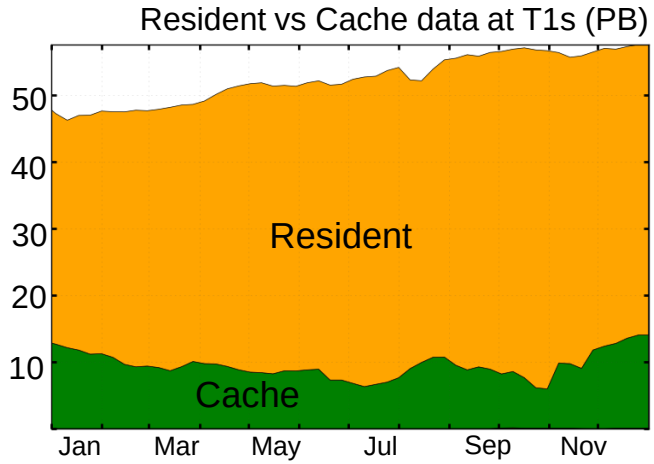
Considerable effort invested to mitigate the resource needs: 2018 requests below “flat budget” growth w.r.t 2017 requests



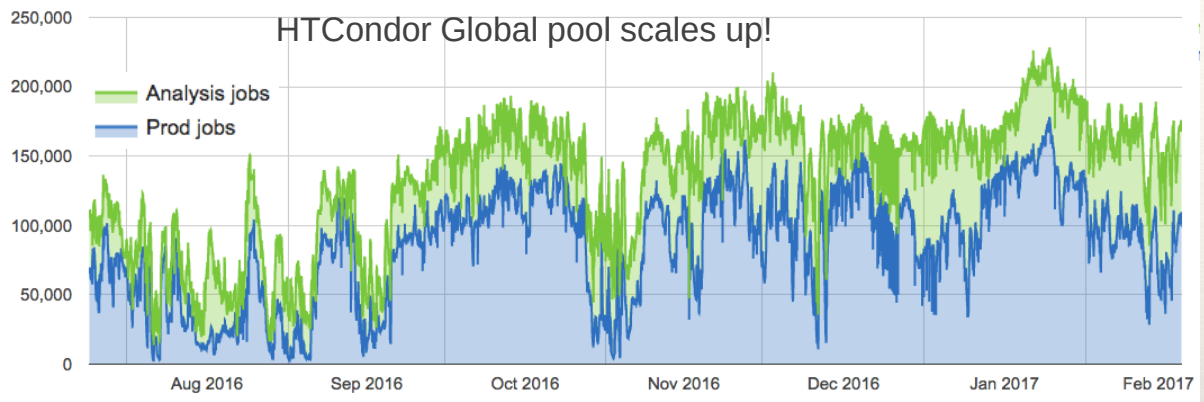
Full and constant use of WLCG pledged resources and benefit of extra 50% opportunistic CPU capacity

Investment in software and computing tools to integrate non Grid resources is showing considerable profit: 15% of CPU capacity from Cloud and HPC

Lifetime model and dynamic data management facilitate space optimization



CMS - resources management *[1/2]*



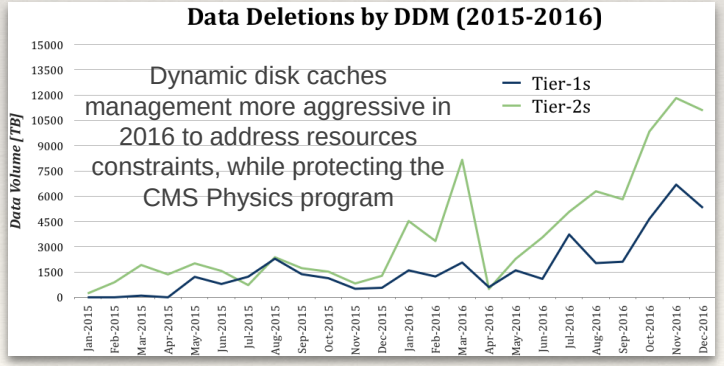
Very high usage of resources in 2016.

On average CMS used:

103% of the T1 CPU pledges

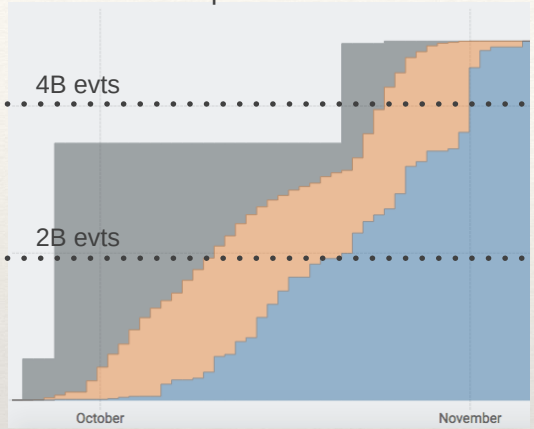
129% of the T2 CPU pledges

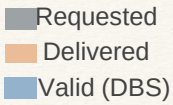
- 96% and 134% as compared to the CMS requirements, respectively



CMS - activities *[2/2]*

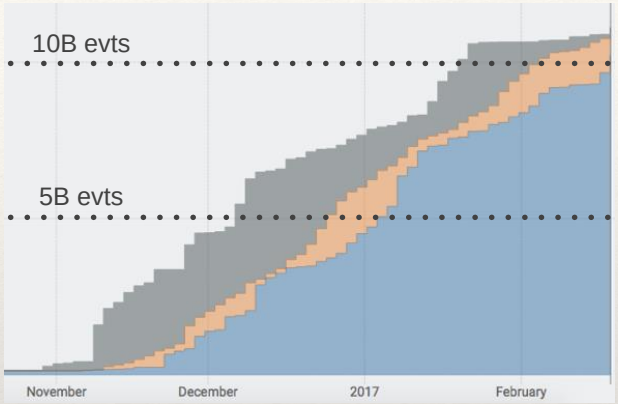
Full 2016 data rereco:
~5B evts completed in ~5 weeks



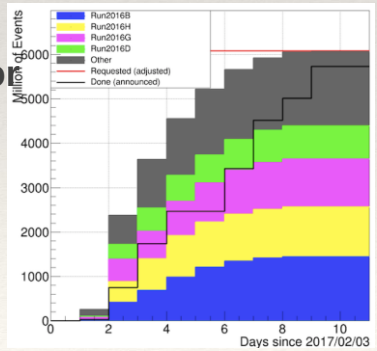


Remarkable slope of delivered events compared to requests injection rate

Major MC re-DigiReco campaign:
>10B evts completed in ~3 months



Re-MiniAOD needed for Moriond'17:
completed in ~1 week

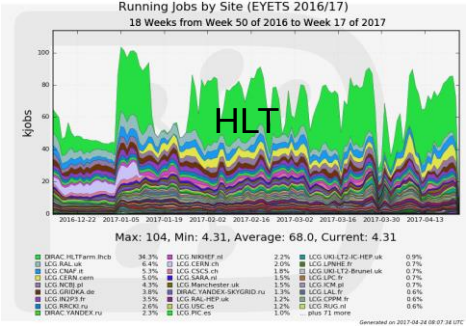


Now and next:

- Phase-I and Phase-II preparations flowing in
- legacy re-reco planned to start around end of March

LHCb Resource Optimizations

2016 mitigation actions to cope with increased resource needs included reduction of disk replicas (3 to 2), parking of ~ 1/3 of TURBO data on tape and purging of ~1 PB of disk resident data via popularity

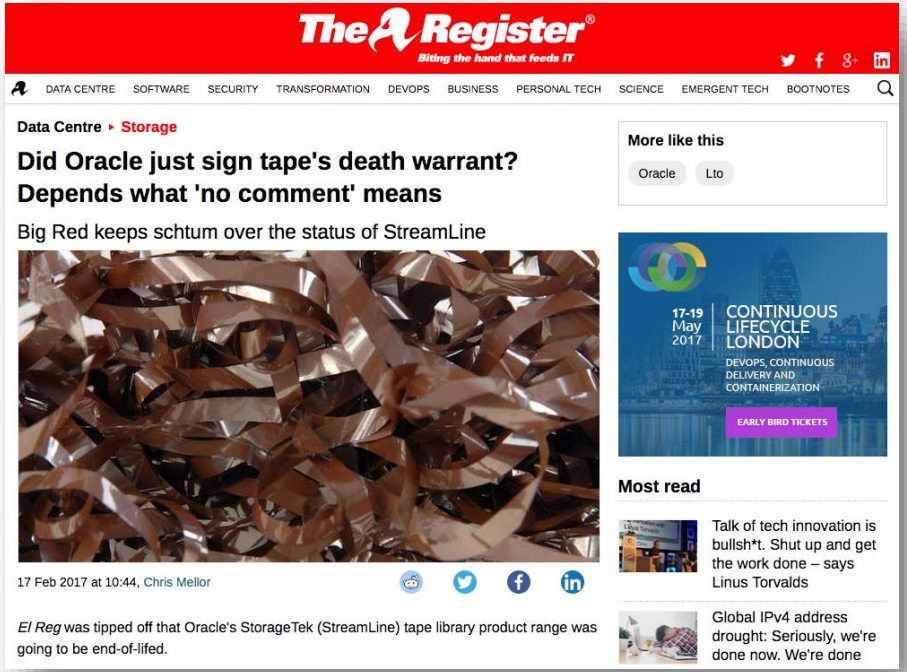


During EYETS very good usage of HLT farm for offline computing. At peak times ~ doubling computing resources

Further application optimizations for 2017 data taking in place via code vectorization in close connection to the experiment's Run3 upgrade activities

Evolution & planning

Concerns over tape …



- Oracle will no longer produce “enterprise” class drives/media
  - Focus on LTO
- Not a huge impact on most Tier 1s
  - Use LTO, IBM
  - Some plan LTO migration
- Has cost implication for CERN (~40 PB cut from costed plan);
  - mitigate with IBM, and introduction of LTO (investment)
- However, long term concern is that IBM now dominates the tape market
- …

CERN | WLCG

Community White Paper

- Mentioned at previous RRB
- Goal to have a Community White Paper (CWP) on overall strategy & roadmap for software/computing for HL-LHC
  - Deliverable of an NSF-funded pre-project
  - Also takes account of Belle-II, ILC, neutrinos, etc.
- To be delivered by summer 2017
- Kick-off workshop held in San Diego 23-26 Jan
- Final workshop planned for end of June
- Will be used as input for the LHCC report later this year, developing roadmap towards TDR for HL-LHC computing in 2020

CWP kick-off workshop

- UC San Diego (Jan 23-26)
- ~110 people, 80% US, 20% other
  - All LHC experiments represented
  - Participation from industry and CS
  - Wider HEP audience – included experts in trigger, reconstruction, ML, etc
- One day of plenary & panel discussions; ~2 days of parallel working group meetings

Plenary discussions

- First day of presentations giving the context for working group discussions:
  - Physics and Computing Challenges of the HL-LHC
  - Situation for Linear Collider, LIGO, FNAL program, Chinese efforts
  - Status of Cloud Technology
  - Situation in (US) funding agencies
  - Licenses and Citation for Software
- Multiple discussion panels on
  - Compute architectures, platforms and software performance
  - Data centres and facilities technologies (clouds, networks, storage)
  - Machine Learning in HEP

Working groups active

- Working groups defining scope, challenges, questions, workplan:
  - Computing Models, Facilities, Distributed Computing
  - Detector Simulation
  - Event Reconstruction and Triggering
  - Data Access and Management
  - Data Analytics and Machine Learning
  - Event Processing Frameworks
  - Workflow and Resource Management
  - Data Analysis and Interpretation
  - Data and Software Preservation
  - Software Development, Deployment and Validation/Verification
  - Role and Future of ROOT
  - Visualization
- All documents are visible at: http://hepsoftwarefoundation.org/cwp/cwp-working-groups.html

Follow up

- Very productive workshop – large engagement of the community
- Working groups have plans for completing the documents
  - Meetings co-located at various community events
- Final CWP workshop to be held in June

HL-LHC Computing TDR

- Agreed with LHCC to produce TDR for HL-LHC computing in 2020
- In 2017 we will provide a document describing the roadmap to the TDR
  - Using the CWP as input
  - Describing potential new computing models
  - Defining prototyping and R&D work that will be needed
- The TDR will not be the end – technology evolution in 6-7 years will be significant, cannot afford not to follow it
- NB. Very different situation from the original TDR –
  - we have a working and well-understood system that must continue to operate and evolve into the HL-LHC computing programme

Conclusions

- Run 2 in 2016 delivered 50 PB of new data, following exceptional performance of the LHC
  - Continued to set new performance records in all areas
- WLCG infrastructure continued to be even more active in the EYETS
- 2017/18 look to be challenging in terms of resource availability, esp if LHC meets expected luminosities, availability
- Activity (& engagement) is ramping up to look at evolution of the computing models for the future