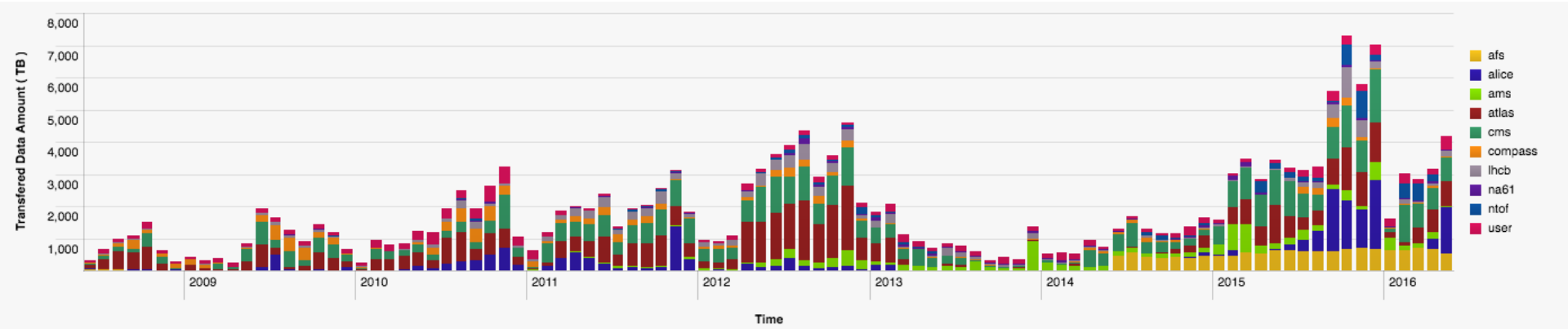
Ian Bird, CERN

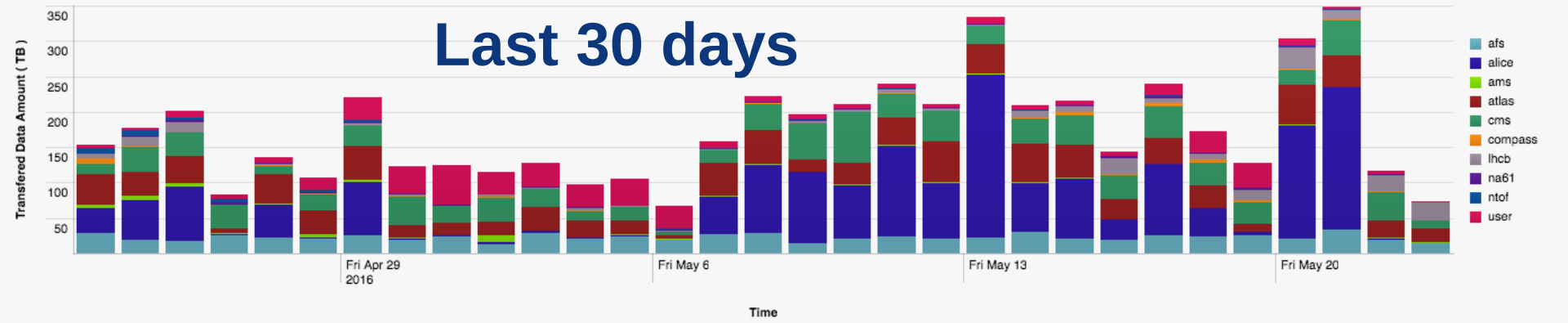
WLCG LHCC Referee Meeting

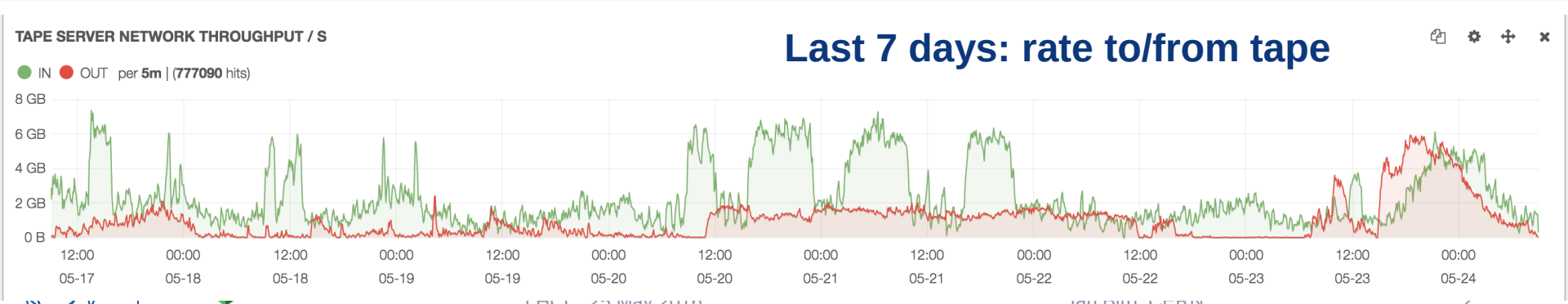
25th May 2016

WLCG Topics

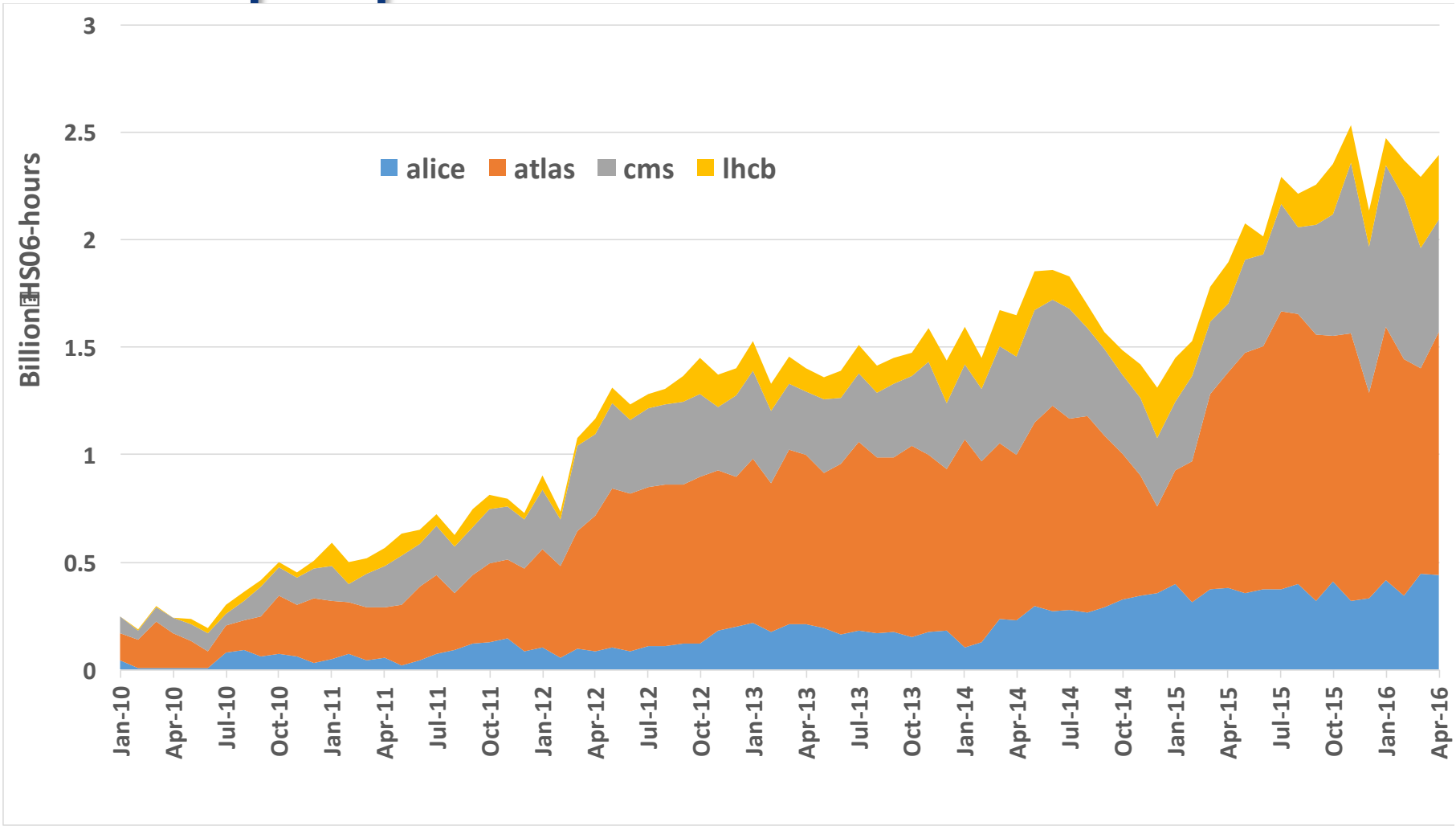
2016 data in Tier 0







Ramp-up of WLCG CPU



Tier 0

MEYRIN DATA CENTRE

	last_value
Number of Cores in Meyrin	151,107
Number of Drives in Meyrin	83,702
Number of 10G NIC in Meyrin	9,305
Number of 1G NIC in Meyrin	23,641
Number of Processors in Meyrin	25,207
Number of Servers in Meyrin	13,373
Total Disk Space in Meyrin (TB)	175,893
Total Memory Capacity in Meyrin (TB)	613

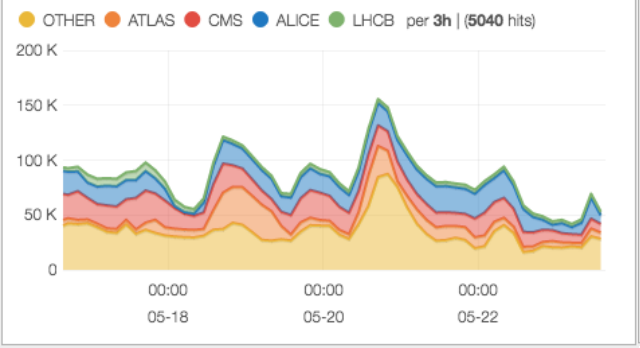
WIGNER DATA CENTRE

	last_value
Number of Cores in Wigner	43,328
Number of Drives in Wigner	23,180
Number of 10G NIC in Wigner	1,399
Numer of 1G NIC in Wigner	5,067
Number of Processors in Wigner	5,418
Number of Servers in Wigner	2,712
Total Disk Space in Wigner (TB)	71,738
Total Memory Capacity in Wigner (TB)	172

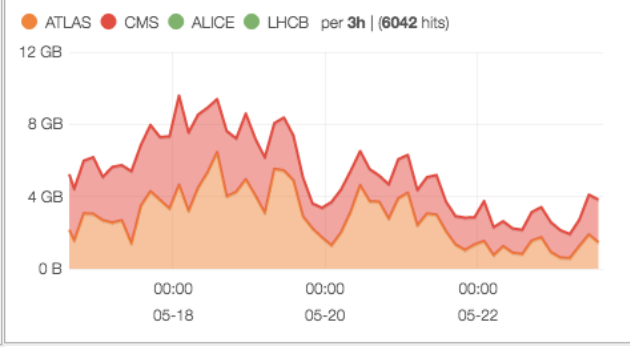
NETWORK AND STORAGE

	last_value
Tape Drives	104
Tape Cartridges	20,517
Data Volume on Tape (TB)	140,606
Free Space on Tape (TB)	44,024
Routers (GPN)	140
Routers (TN)	30
Routers (Others)	107
Switches	3,708

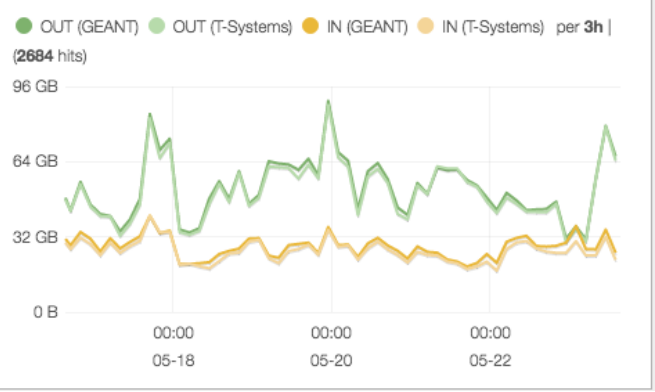
BATCH JOBS (#)



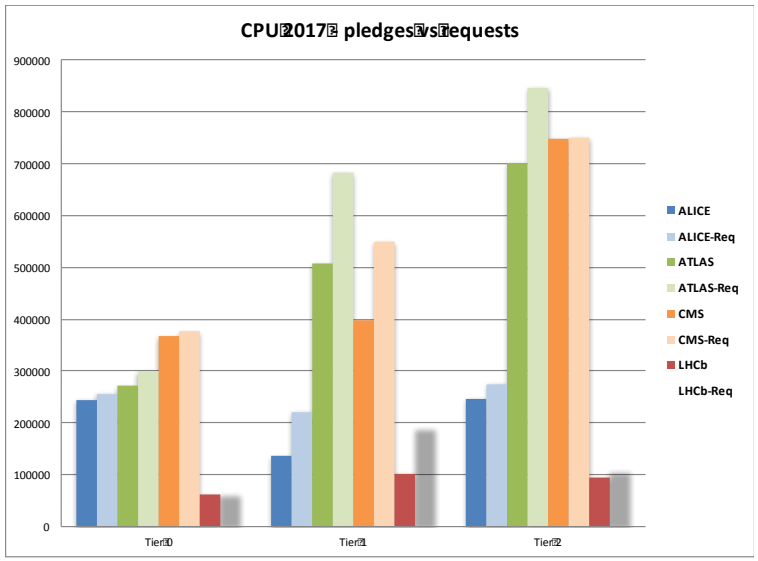
FILE TRANSFER THROUGHPUT (GB/S)

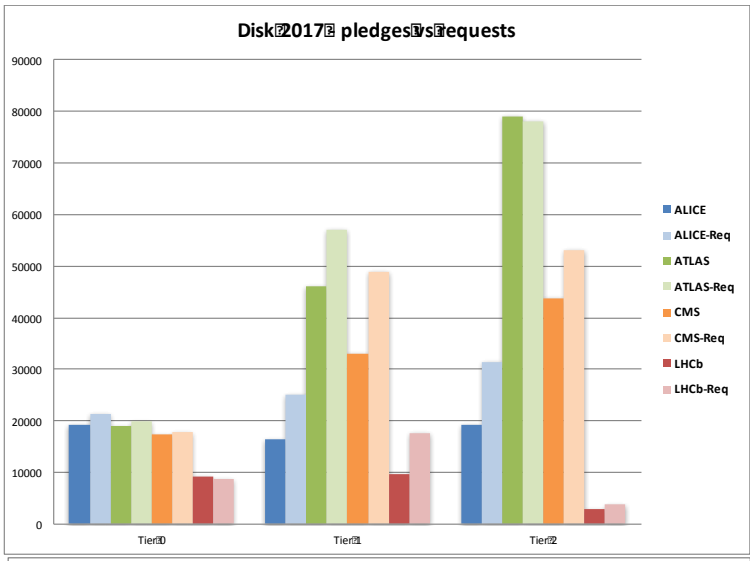


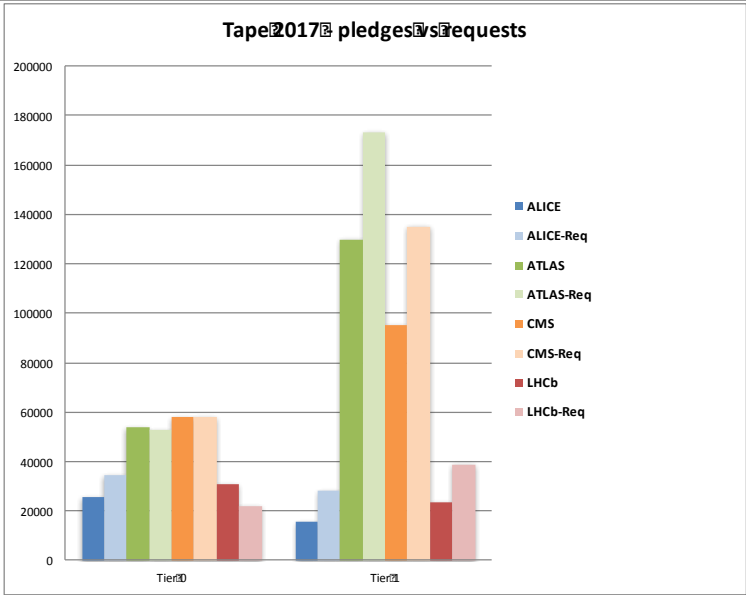
WIGNER NETWORK LINKS (GBIT/S)



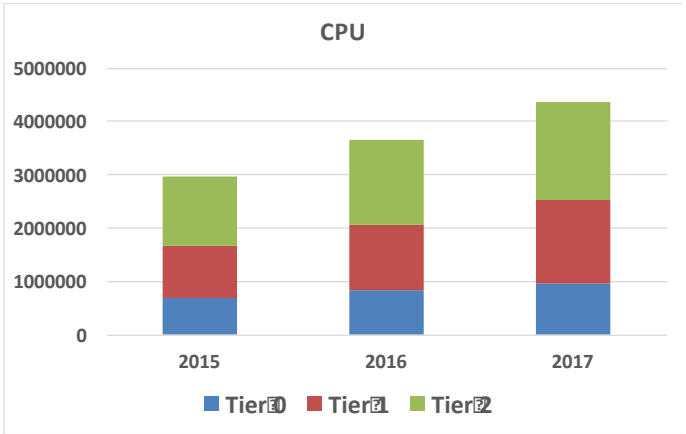
Resources - 2017

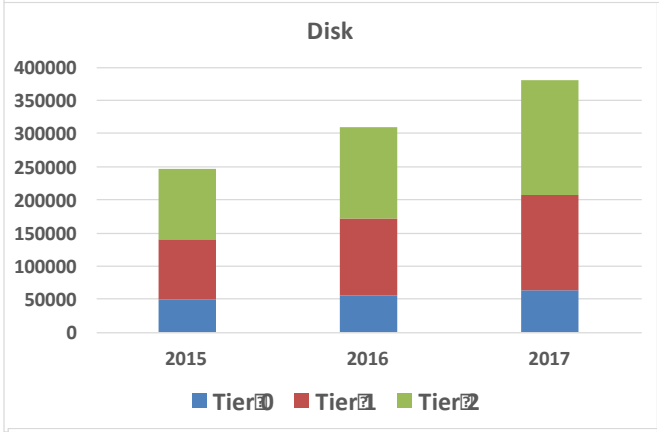


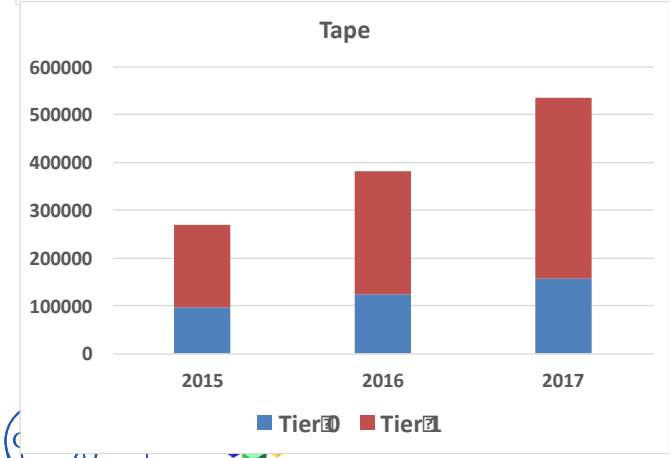


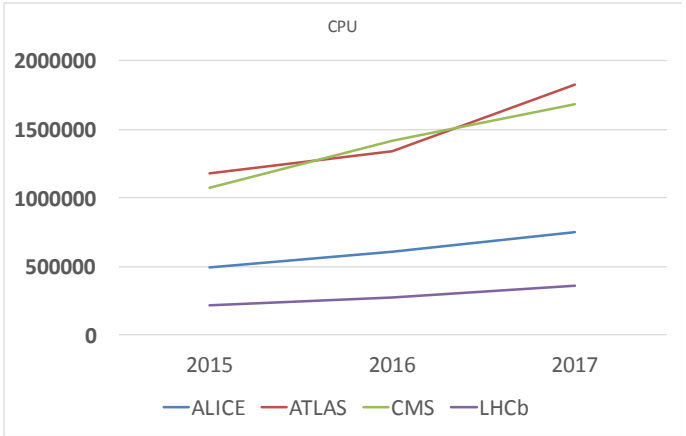


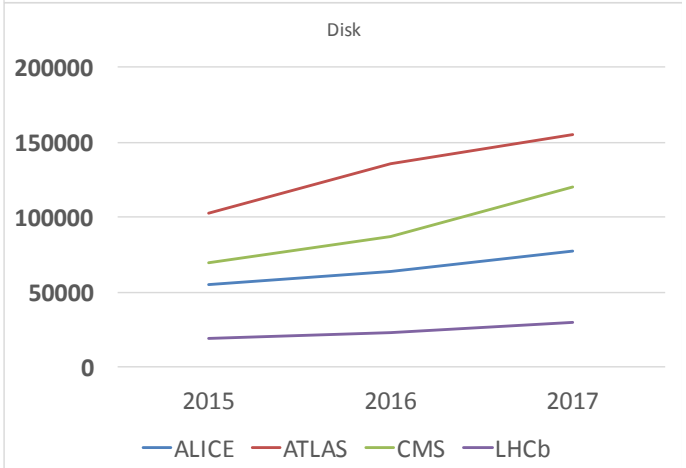
Growth

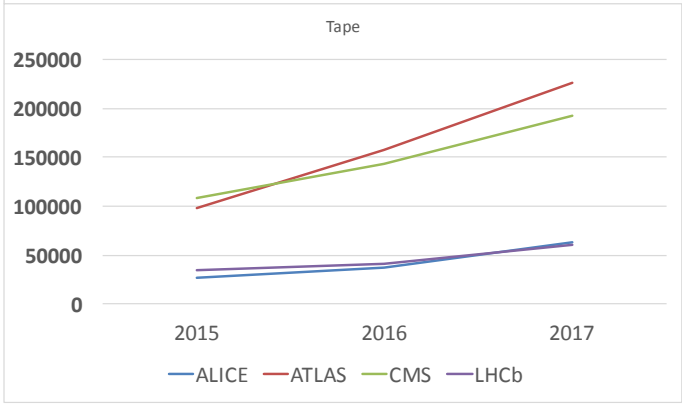

CPU
5000000
4000000
3000000
2000000
1000000
0
2015
2016
2017
Tier0
Tier1
Tier2


Disk
400000
350000
300000
250000
200000
150000
100000
50000
0
2015
2016
2017
Tier0
Tier1
Tier2


Tape
600000
500000
400000
300000
200000
100000
0
2015
2016
2017
Tier0
Tier1


CPU
2000000
1500000
1000000
500000
0
2015
2016
2017
ALICE
ATLAS
CMS
LHCb


Disk
200000
150000
100000
50000
0
2015
2016
2017
ALICE
ATLAS
CMS
LHCb


Tape
250000
200000
150000
100000
50000
0
2015
2016
2017
ALICE
ATLAS
CMS
LHCb

Planning - progress

- Planning assumptions
  - Work has started to document baseline assumptions for Run 4 - ongoing
  - Following format of resource outlook from Run 2 document
- Software topics
  - See Benedikt's talk today, following HSF workshop; including first performance workshop

Planning: Demonstrators

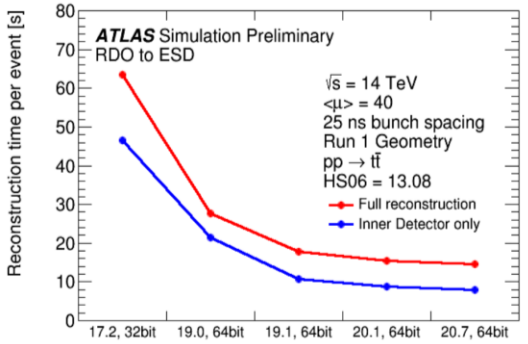
- Work has started on prototypes for:
  - Storage federations:
    - Reliable storage at large sites (clustering of storage between sites) – single view for experiments;
      - concrete proposal for this based on EOS, to be prototyped in Russia initially
    - Opportunistic (caches) at small sites – essentially to support active work, but not long term
      - No concrete work yet, but could be based on DPM
  - Machine Learning:
    - Interest from all experiments; investigate anomaly detection in production workflows, and in data management systems; in particular combining information from >1 experiment
      - Some funding in Russia, efforts in ATLAS and LHCb, CMS also interested
- Under discussion:
  - How to instantiate a “Tier 2 in a box” → simplified deployment of a site with minimal effort needed
    - Several ideas being investigated – to be followed up

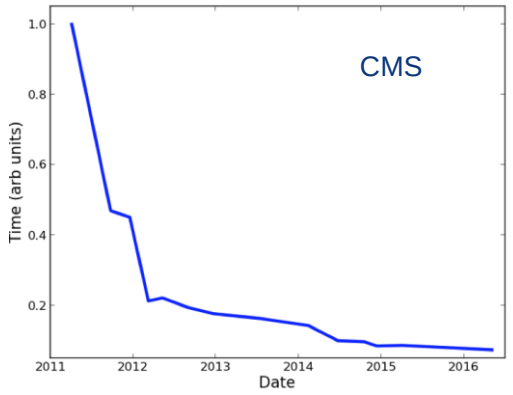
Performance session at HSF

- https://indico.cern.ch/event/496146/timetable/
  - Alice
  - ATLAS
  - CMS
  - Astro Physics
  - Root
  - Geant(V)
  - Art/LArSoft
- Followed by a panel discussion: (summary) https://indico.cern.ch/event/496146/contributions/2147420/attachments/1265860/1895094/SummaryPanel.pdf

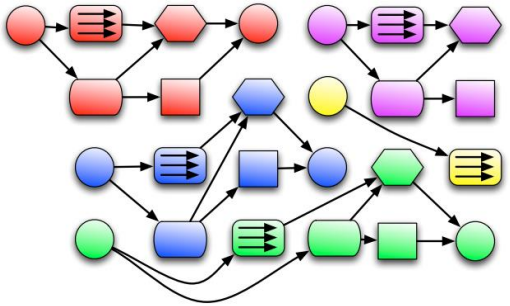
Observations:

- Massive work has been invested in improving the code used in Run1
  - Selection of efficient libs (CLHEP→ Eigen)
  - Tracking in magnetic fields …….....
  - Static code analysis
  - Performance monitoring
    - Many different tools
  - Reorganisation of data structures
  - Experiments know where the time is spent
    - Which doesn't mean that it is clear how to speed it up…
  - Huge improvements, but:
    - gains are getting smaller → region of diminishing return?
- ALICE invested heavily in understanding the use of GPUs
  - New tracking code developed with GPUs and parallelism in mind
    - Improved performance also in CPUs
    - Many systematic studies (see slides)
- Cooperation between experiments exists, especially in tracking
- Everyone states that what matters is throughput, but most work is focussed on acceleration of code
  - Event/sec is used as unit, not event/sec per unit of hardware investment
- Experience from other communities: Astro
  - Classical HPC applications, "standard" codes, limited number of algorithms
  - HEP use cases much more divers, many algorithms in the same job, data intensive ….





Future directions



- Exploiting parallelism wherever possible
  - Starting from the frameworks (Gaudi, AthenaMT, ROOT,…)
  - GPUs, accelerators, multi core
- Part of the community is convinced that HEP computing will depend in the future on efficient usage of HPC machines
  - Part of the community disagrees strongly
- Evolution not revolution
  - Effort levels exclude a revolution
  - Not clear how much can be gained from evolution
- Agreement that common approach is needed for HL-LHC
  - Community Whitepaper
  - Working groups