Ian Bird

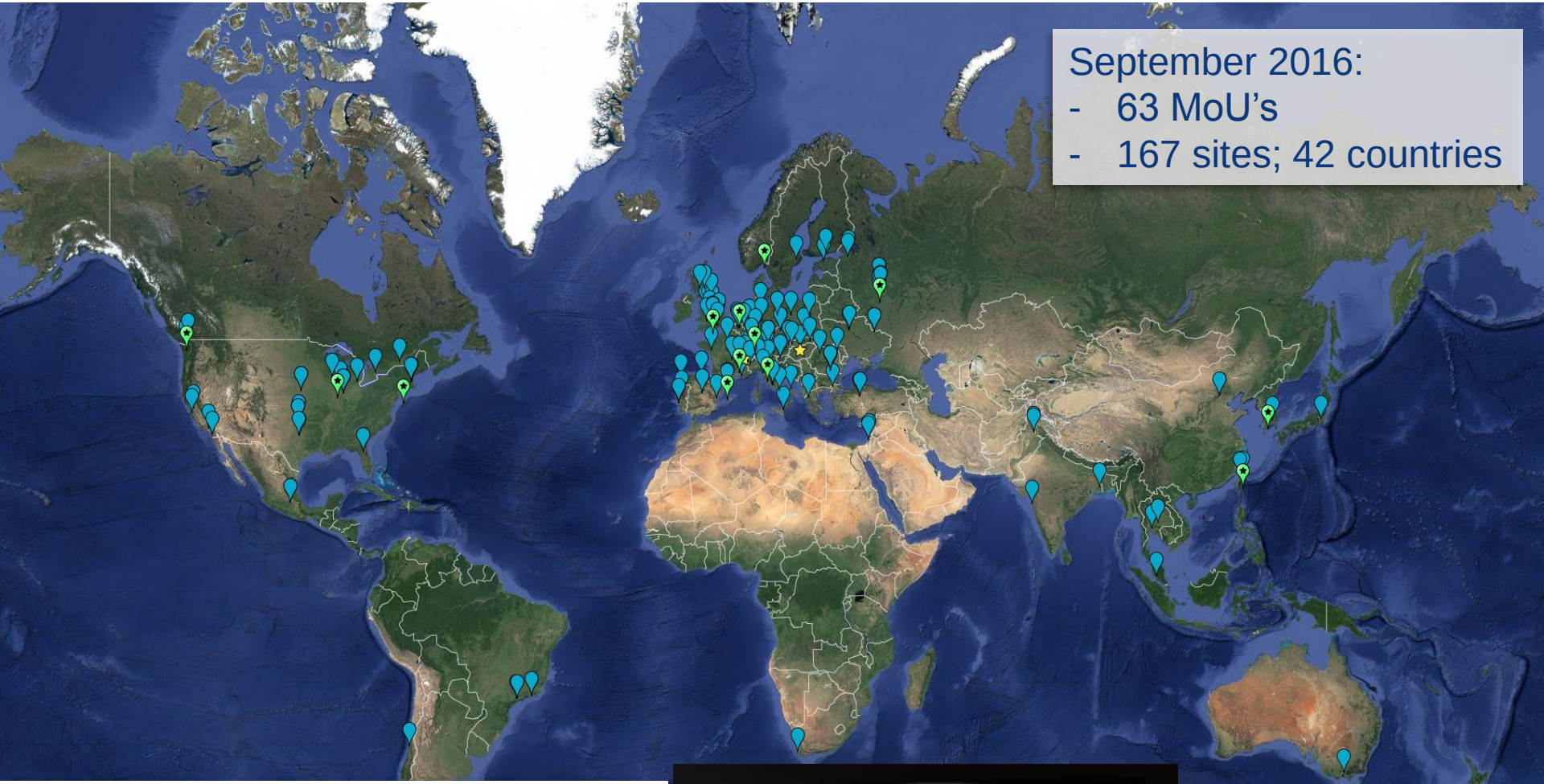
WLCG Workshop

San Francisco, 8th October 2016

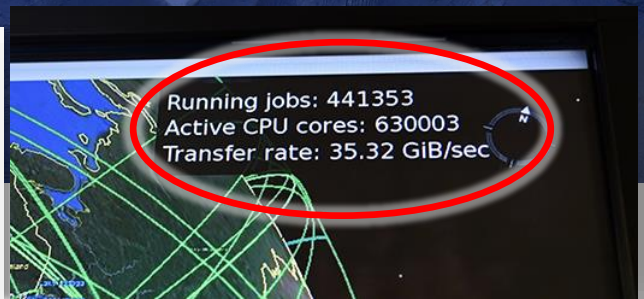
Workshop Introduction

Context of the workshop: Half-way through Run 2; Preparing for Run 3, Run 4

WLCG Collaboration



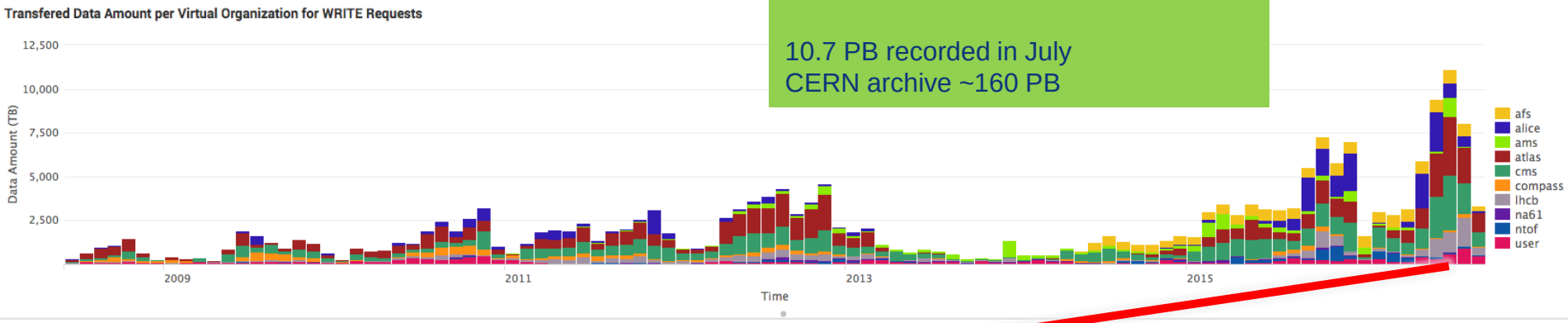
- CPU: 3.8 M HepSpec06
  - If today's fastest cores: ~ 350,000 cores
  - Actually many more (up to 5 yr old cores)
- Disk 310 PB
- Tape 390 PB

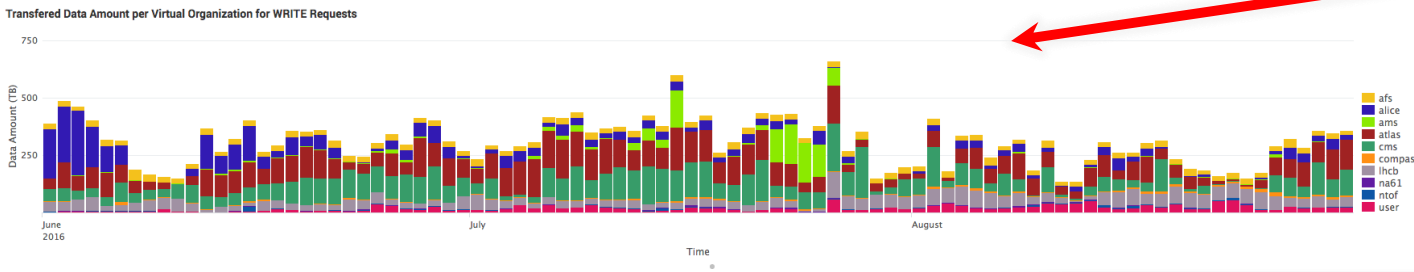


2016 data

LHC data – Continue to break records:

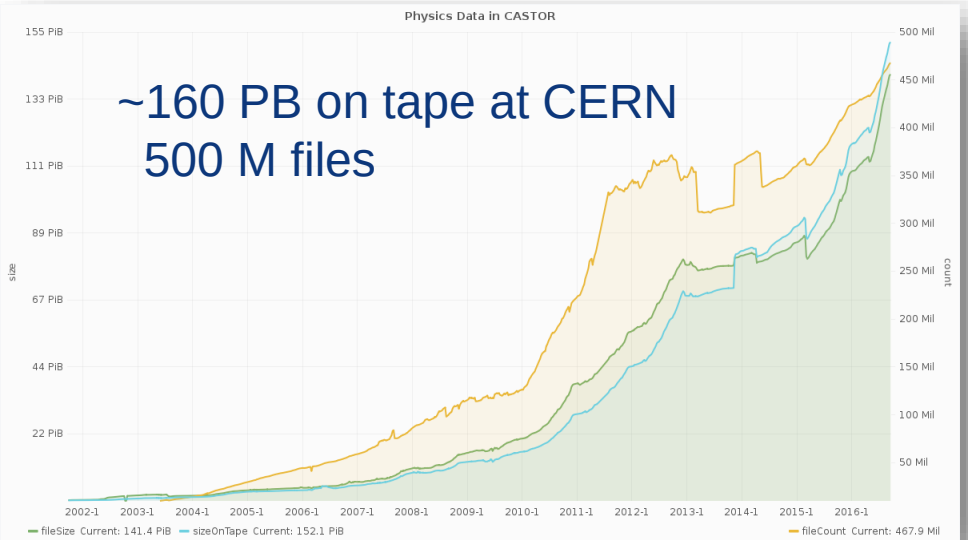
10.7 PB recorded in July
CERN archive ~160 PB





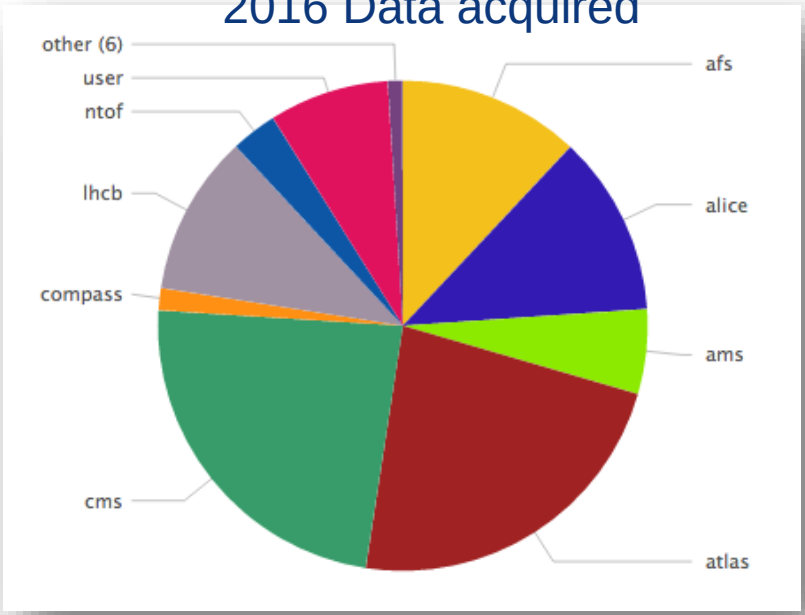
June-Aug 2016
>500 TB / day
(Run 1 peak for HI was 220 TB)

2016 to date: 35 PB LHC data:
ALICE 6, ATLAS 11.6, CMS 11.9, LHCb 5.4)



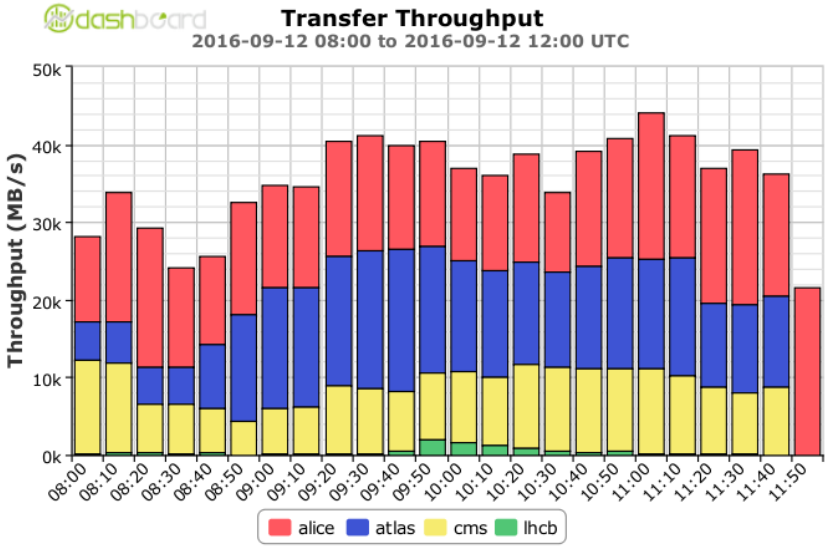
~160 PB on tape at CERN
500 M files

2016 Data acquired



Data distribution

- Global transfer rates increased to > 40 GB/s (=2 x Run1)



Monthly traffic growth on LHCONE

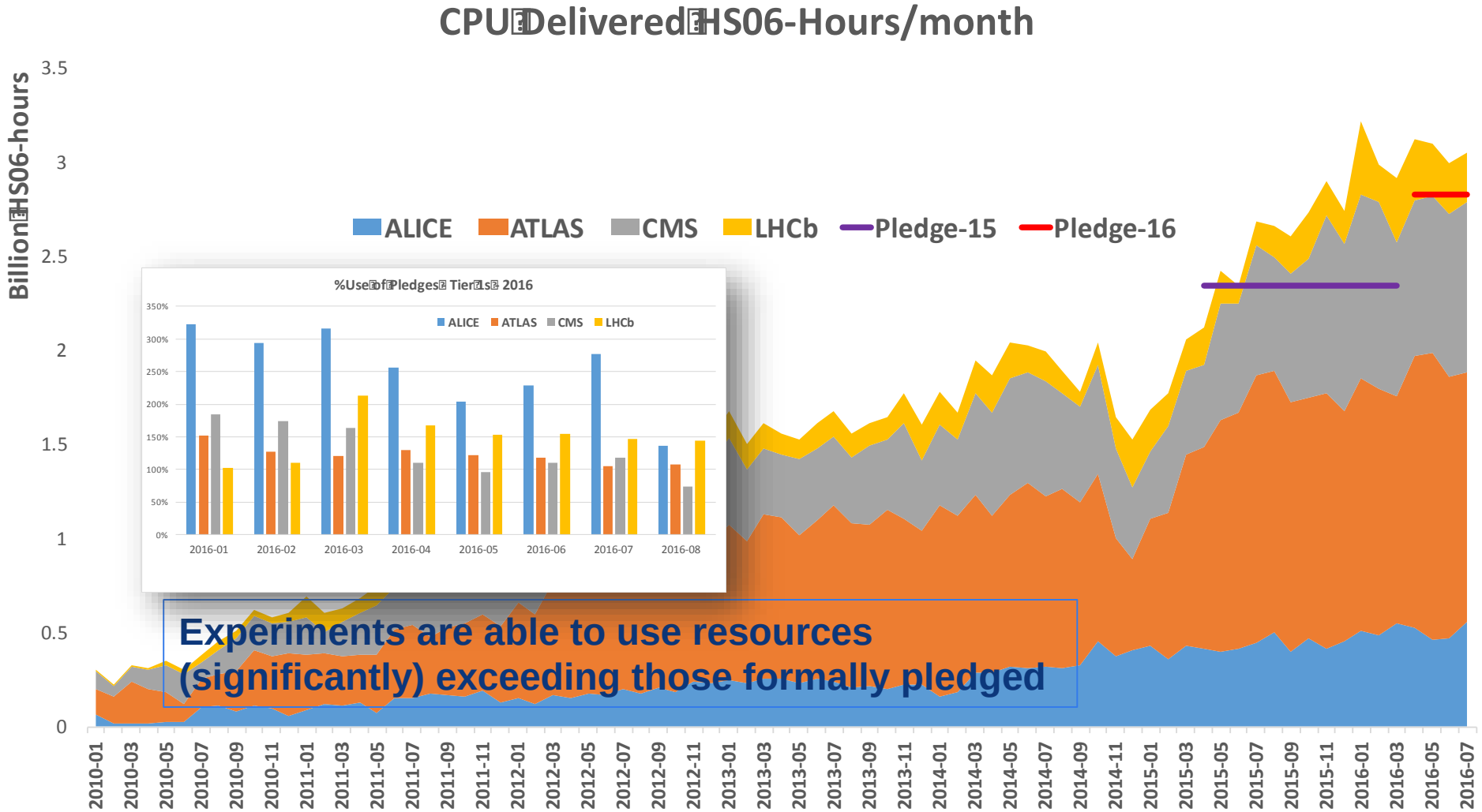
Increased performance everywhere:

- Data acquisition >10PB / month
- Data transfer rates > 40 GB/s globally

Several Tier 1s have increased network bandwidth to CERN to manage new data rates;
GEANT has deployed additional capacity for LHC

Regular transfers of 80 PB/month with 100 PB/month during July-Aug (many billions of files)

CPU delivered

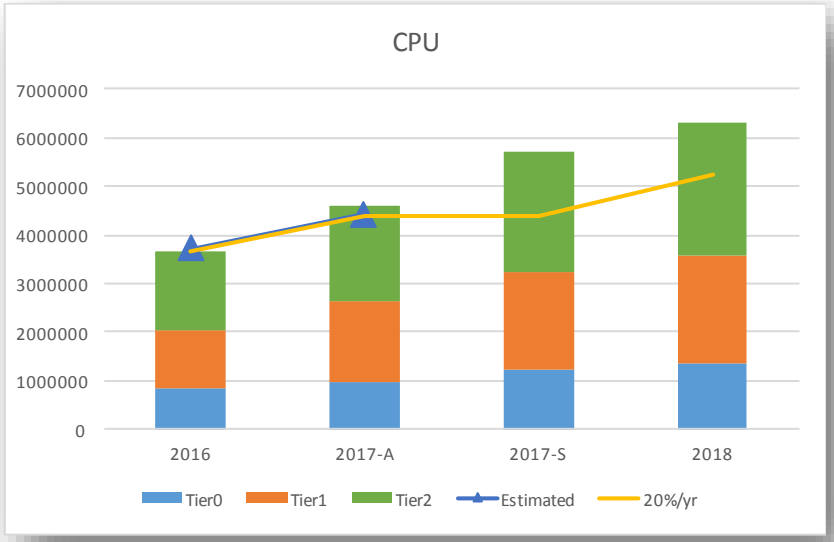


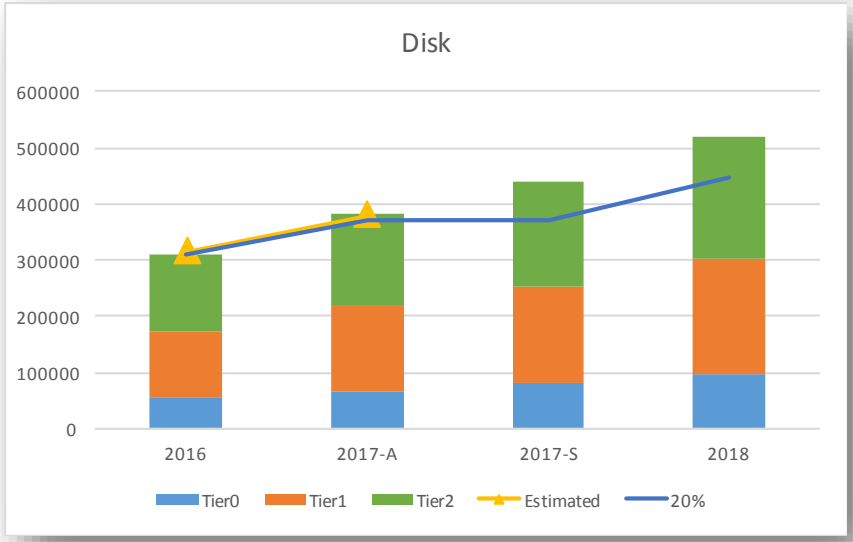
Resource requirement evolution

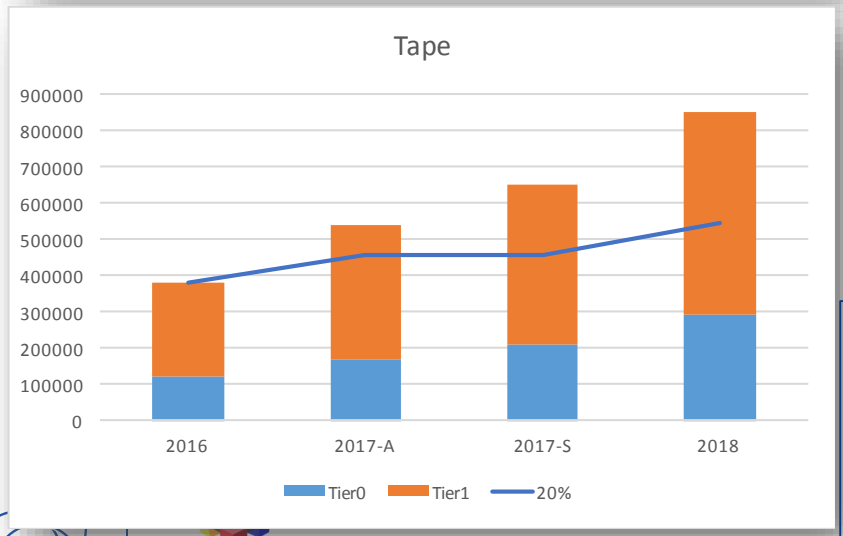
Run2: Increased computing needs

- LHC performance is above expectations:
- Computing needs driven by (mainly):
  - LHC live time (37% → > 60%)
  - Luminosity ($1.0x10^{34}$ → $1.2x10^{34}$ or better)
  - Pile-up (CMS, ATLAS) (21 → 33 on average)
- For 2016, the available resources will be sufficient
  - More tapes at CERN have been bought
- Re-analysis for 2017,18
  - Just done in time for RRB
  - Not yet scrutinized by RSG
  - But: expectations are increased requirement above previous estimates of 15-30%

Re-assessment of needs







Estimated: Estimates made in 2014 for Run 2 up to 2017

20%: Growth of 20%/yr starting in 2016 ("flat budget")

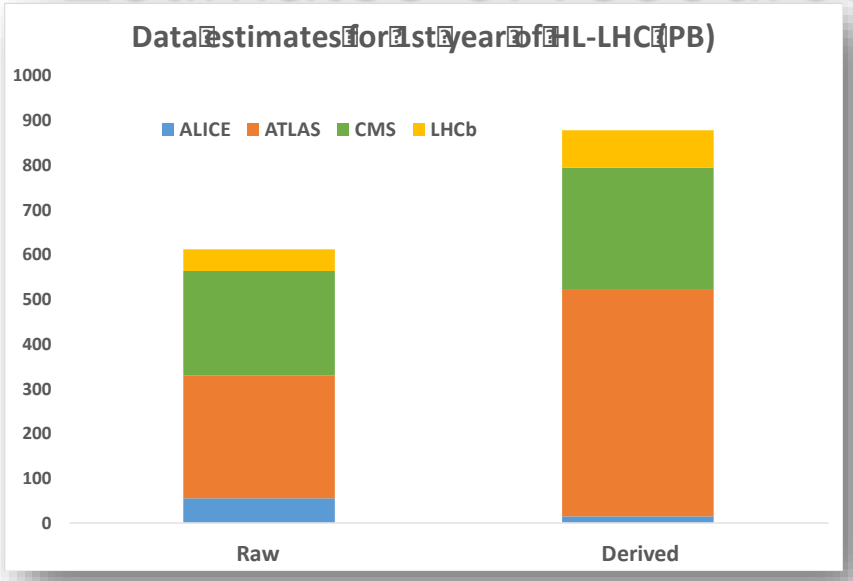
The reliability of resource predictions is continually improving, the largest uncertainties being the LHC running conditions.

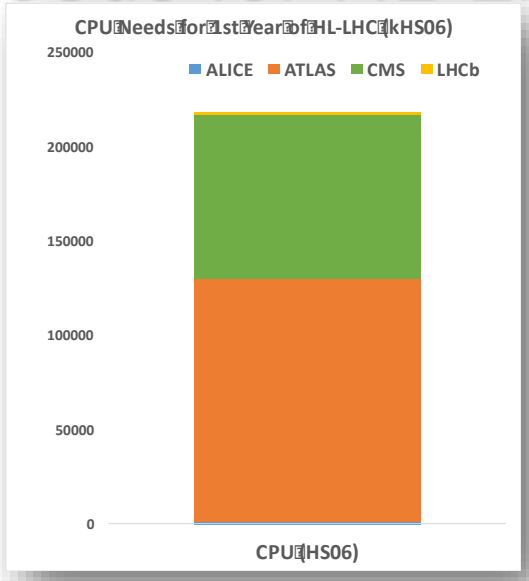
Funding guidance: flat budgets for computing from 2014

Outlook

- Ongoing and continual evolution
  - Computing models & software performance in the experiments
  - Infrastructure – use of clouds, HPC, volunteer computing etc., etc.
- Anticipate:
  - Run 2 and Run 3 will be manageable with an ~evolutionary approach
    - But making use of technology advances where useful
    - ALICE Upgrade TDR done, LHCb this year
  - HL-LHC will require more revolutionary thinking

Estimates of resource needs for HL-LHC





Data:

- Raw 2016: 50 PB → 2027: 600 PB
- Derived (1 copy): 2016: 80 PB → 2027: 900 PB

CPU:

- x60 from 2016

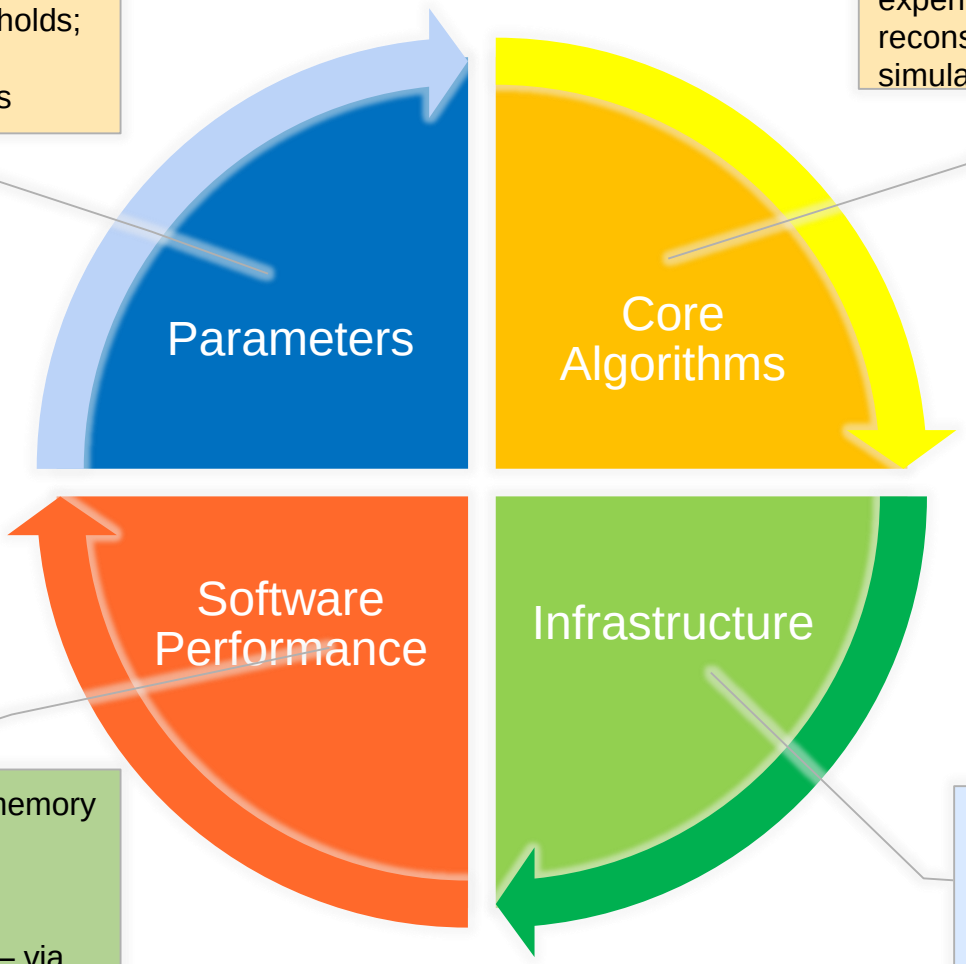
Technology at ~20%/year will bring x6-10 in 10-11 years

- Simple model based on today's computing models, but with expected HL-LHC operating parameters (pile-up, trigger rates, etc.)
- At least x10 above what is realistic to expect from technology with reasonably constant cost

HL-LHC computing cost parameters

Business of the experiments: amount of Raw data, thresholds; Detector design long term computing cost implications

Business of the experiments: reconstruction, and simulation algorithms



Performance/architectures/memory etc.;
Tools to support: automated build/validation
Collaboration with externals – via HSF

New grid/cloud models; optimize CPU/disk/network; economies of scale via clouds, joint procurements etc.

Future work

- Understanding how to make best use of available resources
  - Not just for HL-LHC, but already now
- Have to be efficient in all aspects: infrastructure, applications, people
- The easy gains have been made – we need a sustained effort to optimise
- May require some radical changes