CERN

This Helix Nebula Science Cloud (HNSciCloud) Pilot Phase Open Session

Geneva, Switzerland

Frédéric Hemmer
CERN - IT Department

Innovation

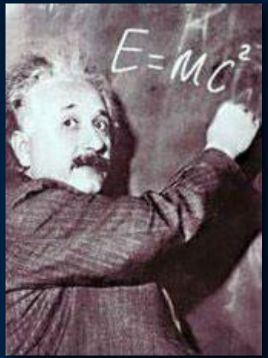
Education

CERN

uniting people

Research

The Mission of CERN



- **Push back** the frontiers of knowledge

  E.g. the secrets of the Big Bang …what was the matter like within the first moments of the Universe's existence?

- **Develop** new technologies for accelerators and detectors

  Information technology - the Web and the GRID

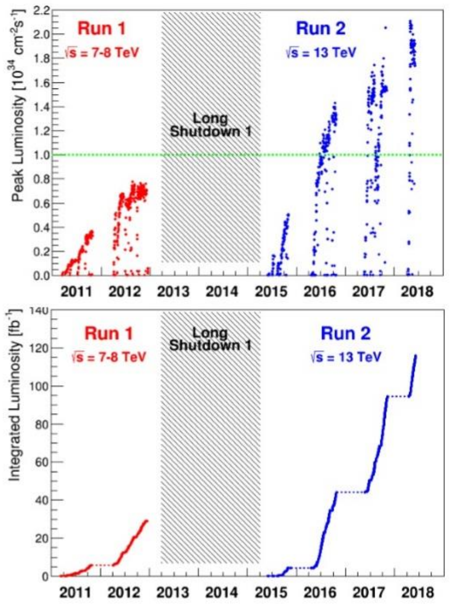
  Medicine - diagnosis and therapy

- **Train** scientists and engineers of tomorrow

- **Unite** people from different countries and cultures

LHC Progress - 2018

Run1 + Run 2: Luminosity Production



Peak Luminosity

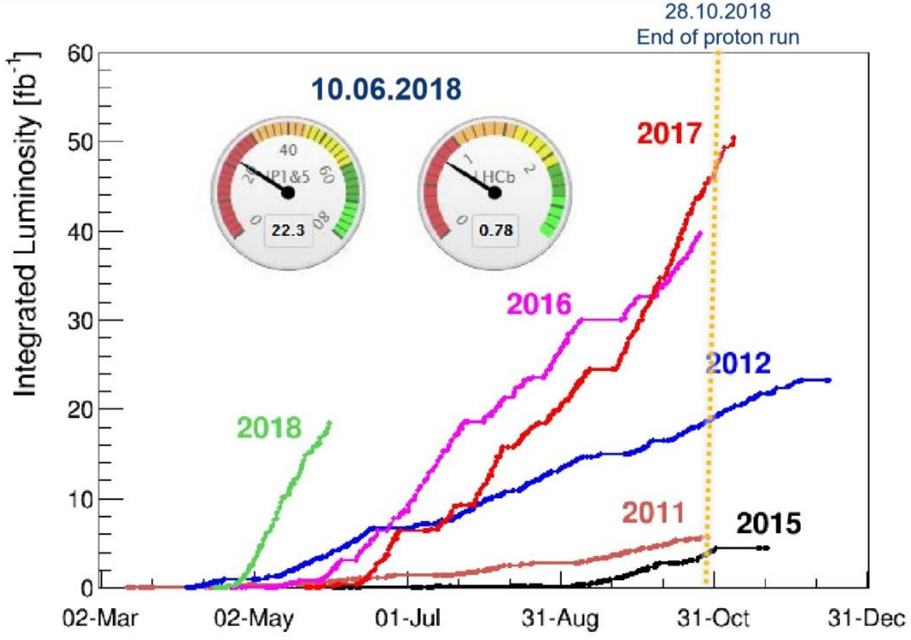
2018 shows steepest increase in peak luminosity of all years

Period	Int. Luminosity [fb^{-1}]
Run 1	29.2
Run 2: 2015	4.2
Run 2: 2016	39.7
Run 2: 2017	50.2
Run 2: 2018	22.3
Total Run 1+ 2	145.6

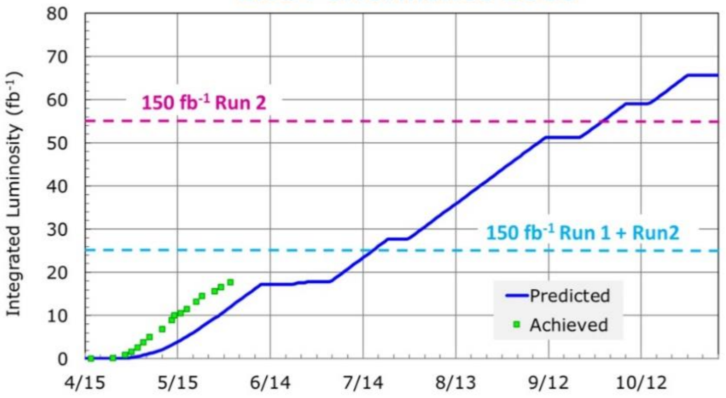
116.4 (Run 2: 2015–2018)

10.06.2018

Multi-annual Integrated Performance



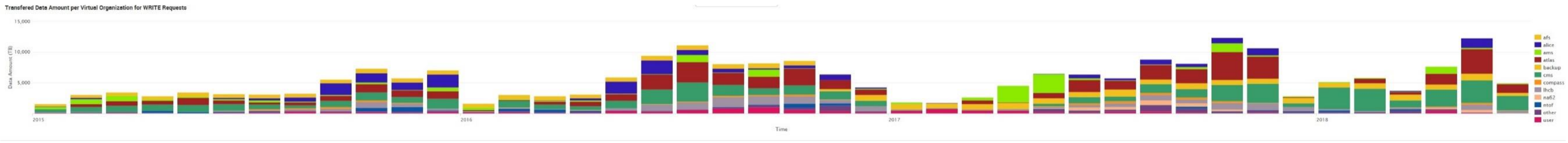
LHC Performance 2018

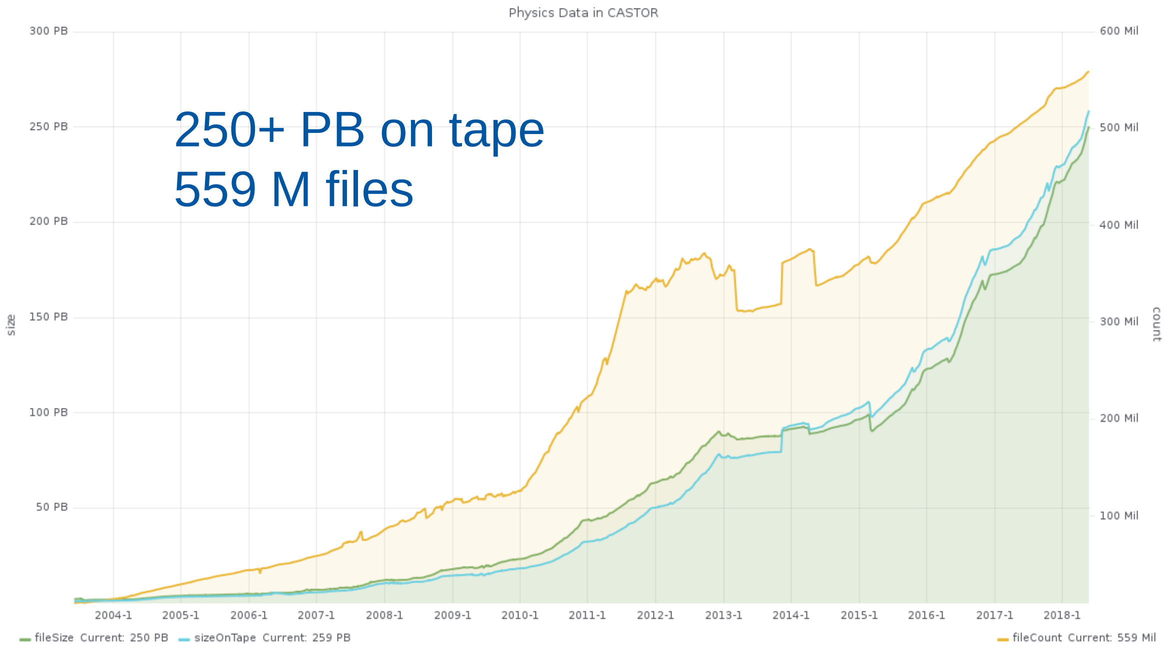


2018:
A Production Year to complete Run2

Document Classification: **Public**

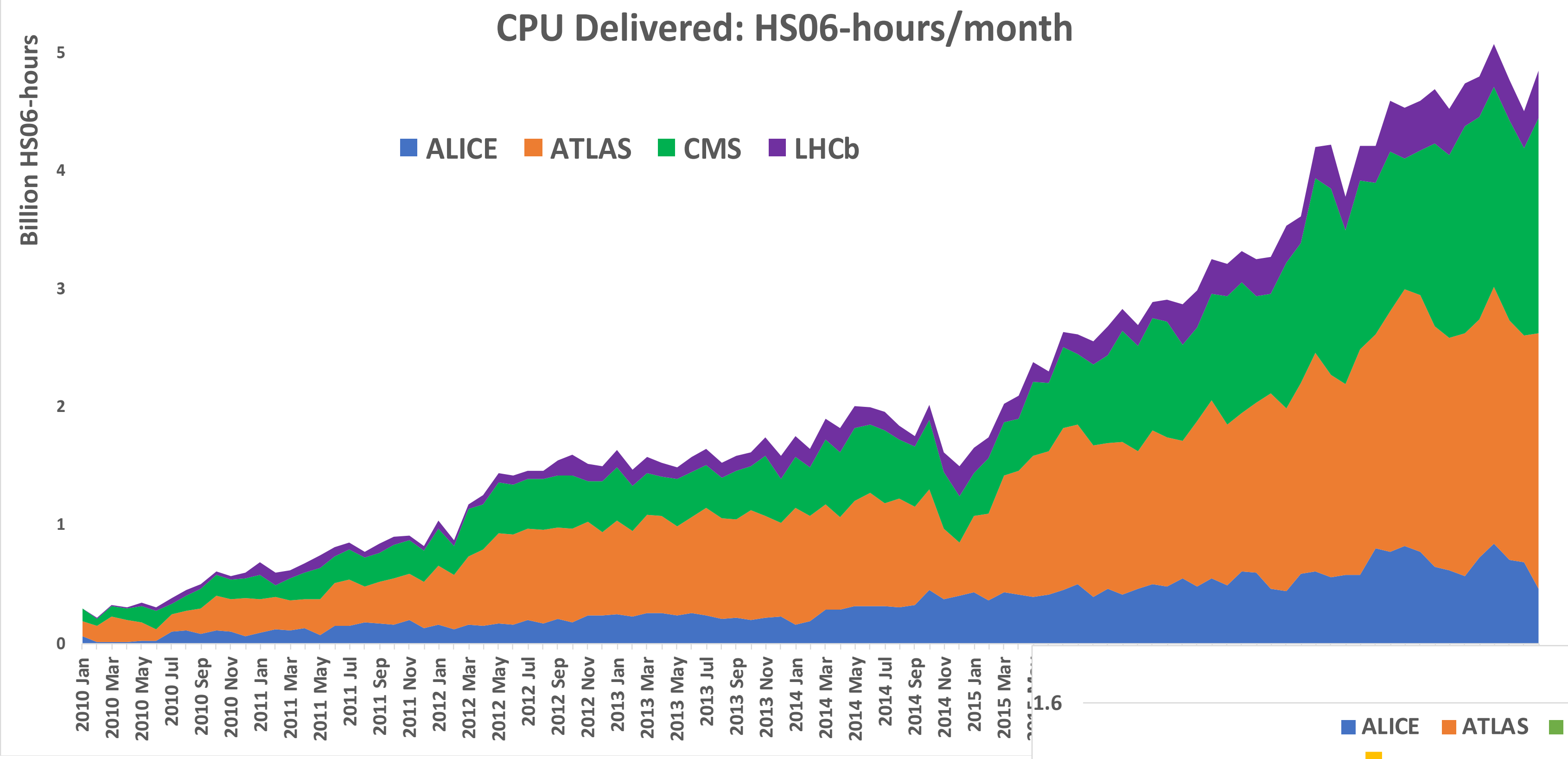
Data in 2018

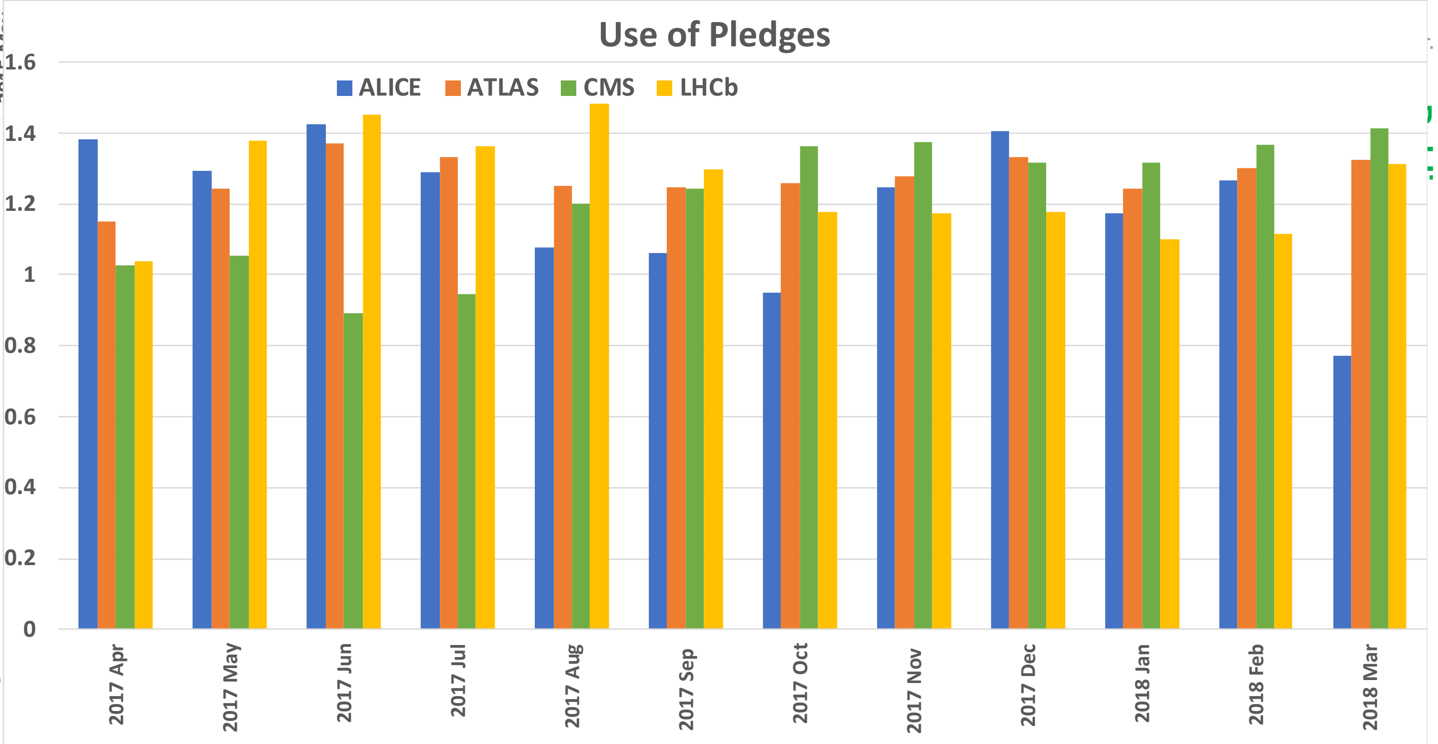




Year	LHC (PB)	Total (PB)
2016	48.3	69.8
2017	38.8	64.3
2018	***28.4!***	38.8

CPU Delivered





New peak: ~210 M HS06-days/month

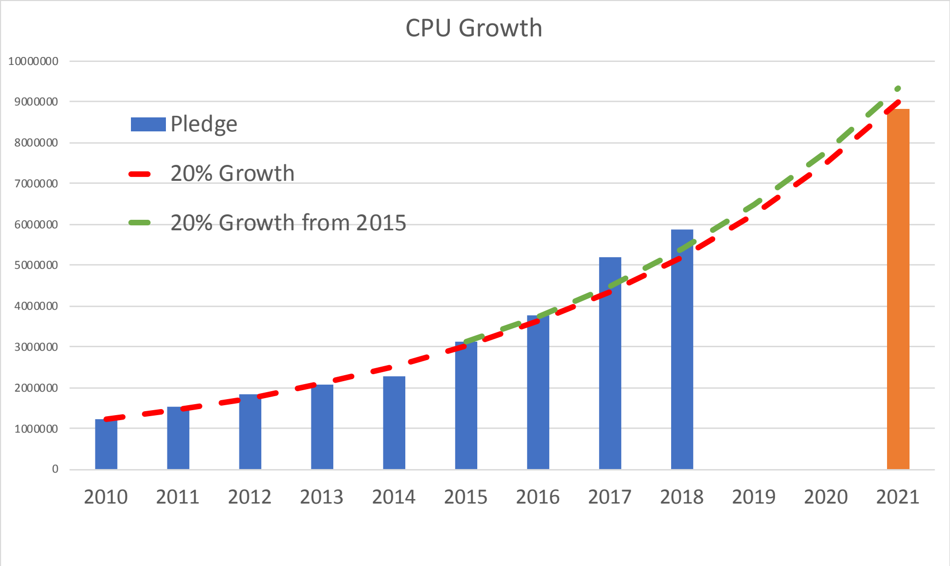
~ 685 k cores continuous

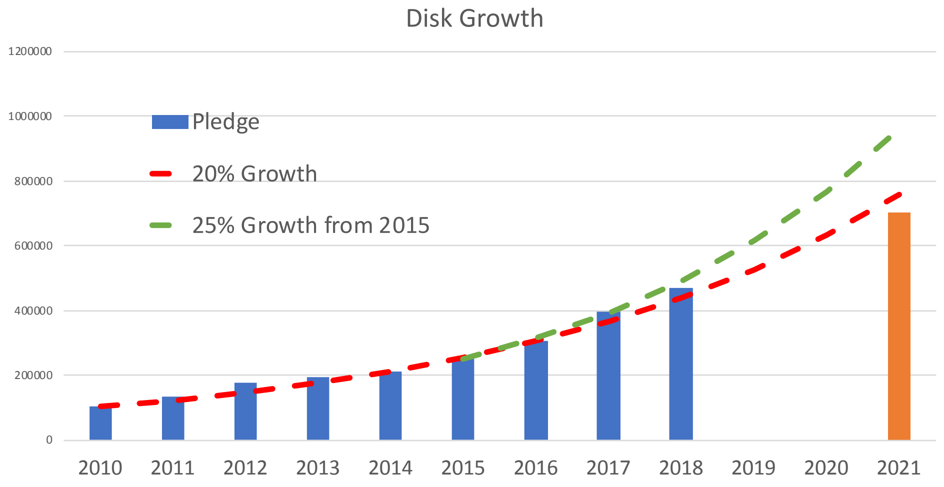
Run 3 Planning (2021-2023):

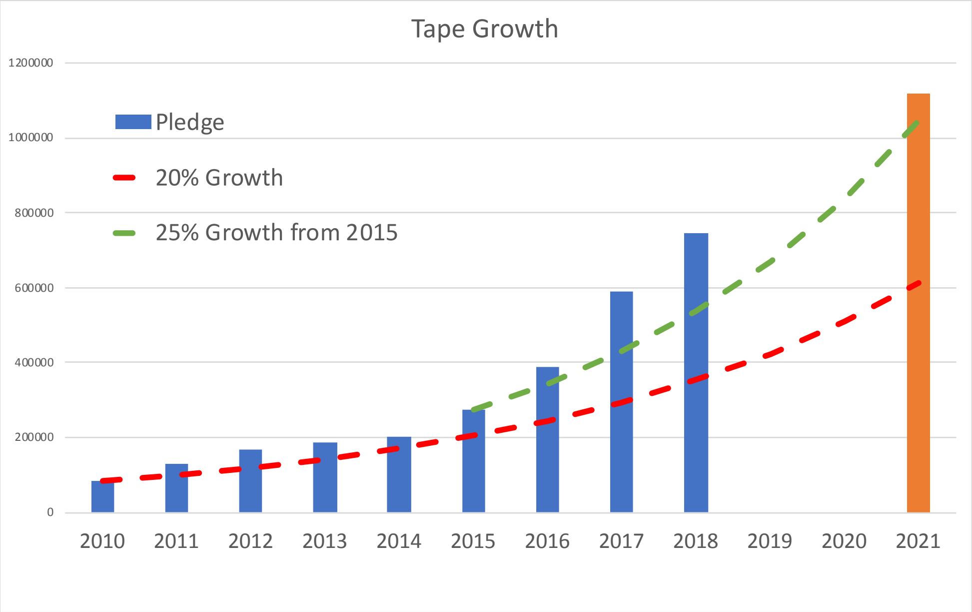
Similar to 2018

- If the experiments luminosity level at a higher pile-up and for longer ➔
  - Potentially higher average pileup
  - Non-linear increase in CPU time
- Possibly less time between fills – more live time
- Overall the best estimate is 30% (50% conservatively) more resources needed than in 2018
  - But we have not seen 2018 yet
- For 2021: 1st year after LS2, could be only half-year live time but ramp up to optimal conditions rapidly
- Unknown:
  - Still need plans for experiment trigger rates
  - And plans for luminosity levelling

Resource evolution







- 2010-2018 – pledges
- 2021 assume 1.5 x 2018

Document Classification: **Public**

However …

- ALICE and LHCb are upgrading during LS2, so the expectations of their needs do not follow the assumptions in the previous slides:

  - LHCb:
    - luminosity and pileup increase by factor 5.
    - Major changes in computing model result in higher trigger rate and HLT output bandwidth.
    - LHCC milestone for computing model in Q3/2018, together with engineering TDR – currently under review

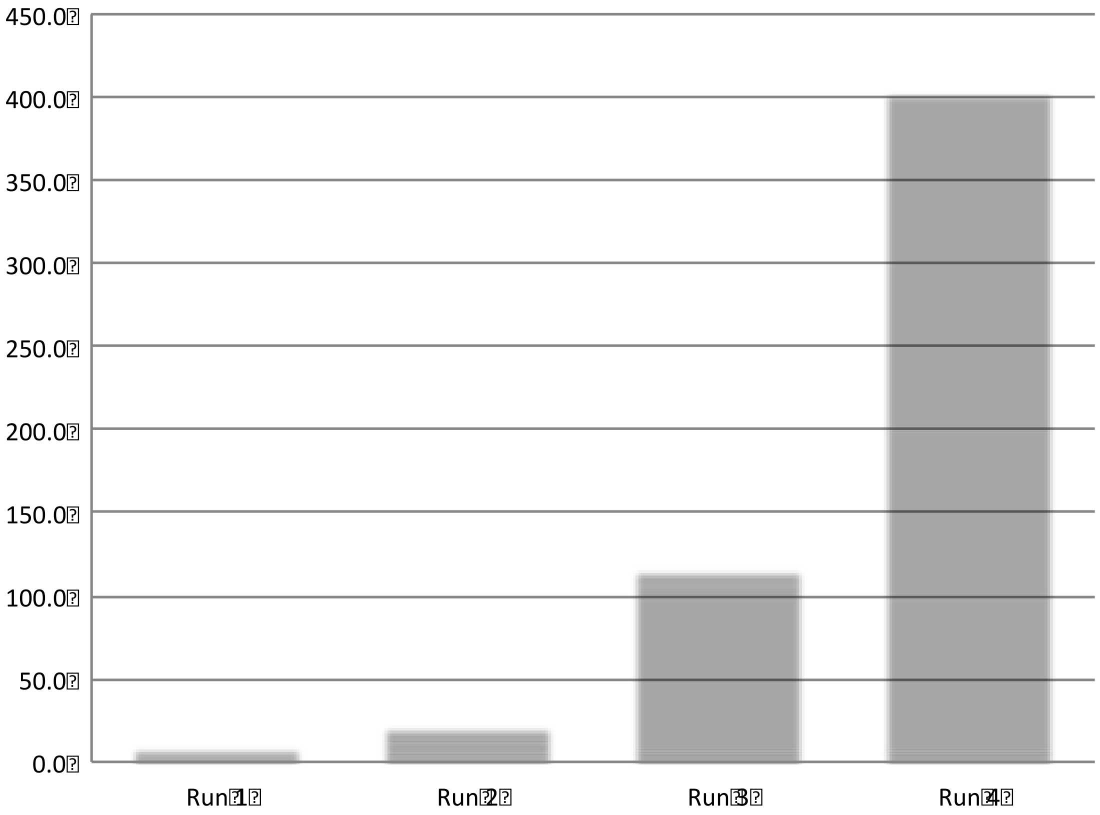
  - ALICE:
    - Factor 100 increase in readout rate (50 kHz)
    - Data volume increase mitigated by online reconstruction and raw data compression in new O2 facility
    - O2 TDR is approved; summary needs are:
    - Increases in 2021 wrt 2018: CPU: 48%, disk: 74%, tape 90%

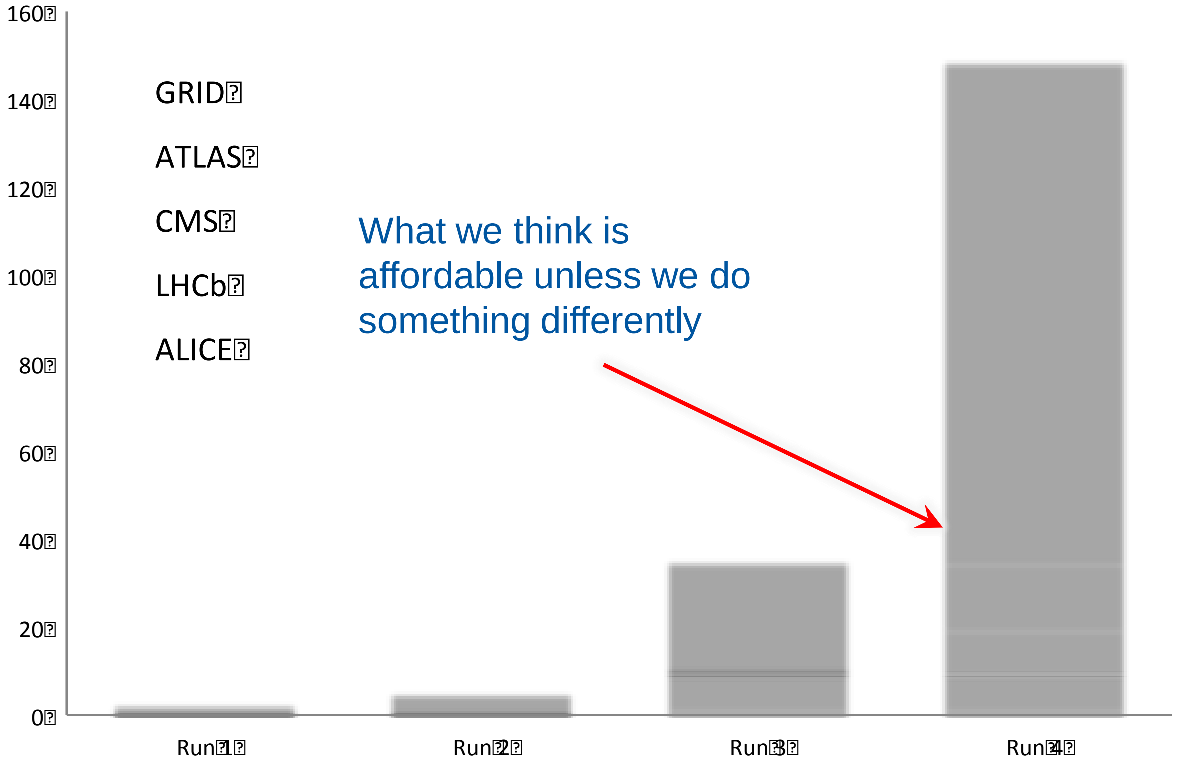
Scale of data tomorrow …

10 Year Horizon

2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	……	2030?

Run 1 | **LS1** | **Run 2** | **LS2** | **Run 3** | **LS3** | **HL-LHC** | **FCC?**





2010 2015 2021 2024

2010 2015 2021 2024

Data: ~25 PB/year → 400 PB/year

Compute: Growth > x50

Technology revolutions are needed

The WLCG Strategy Document

- The HL-LHC computing challenge: provide the computing capacity needed for the LHC physics program, managing the cost
- The WLCG strategy document is a specific view of the CWP, prioritizing R&Ds relevant to the HL-LHC computing challenge
- The prototyped solutions will be the foundation of the WLCG TDR for HL-LHC, planned for 2020. Timing to be re-considered?
- This is a presentation of the content of the strategy document
    - http://cern.ch/go/Tg79

WLCG Strategy - Outline

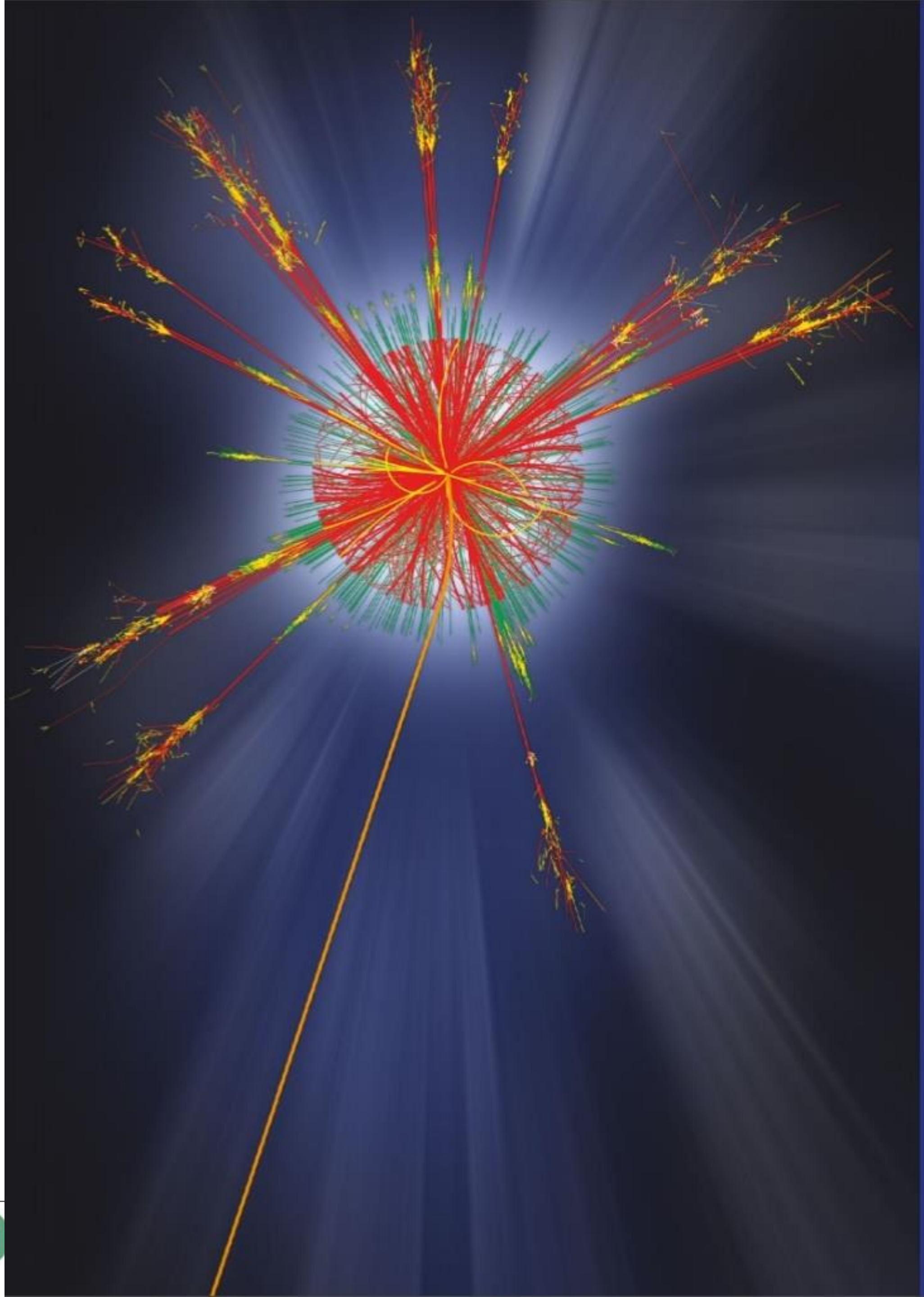
The strategy develops around five main themes …

1. Software performance
2. Algorithmic improvements / changes (e.g. generators, fast MC, reconstruction)
3. Reduction of data volumes
4. Managing operations cost
5. Optimizing hardware costs

It defines an R&D program with rough timelines, organized in sections:

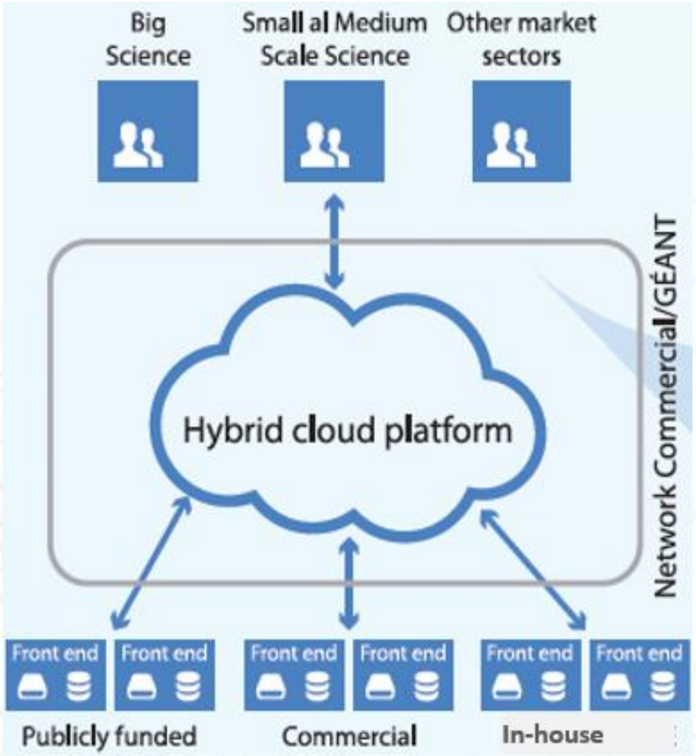
- The HL-LHC challenge, hardware trends and a cost model
- Computing Models
- Experiments Software
- System Performance and Efficiency
- Data and Processing Infrastructures
- Sustainability
- Data Preservation and Reuse

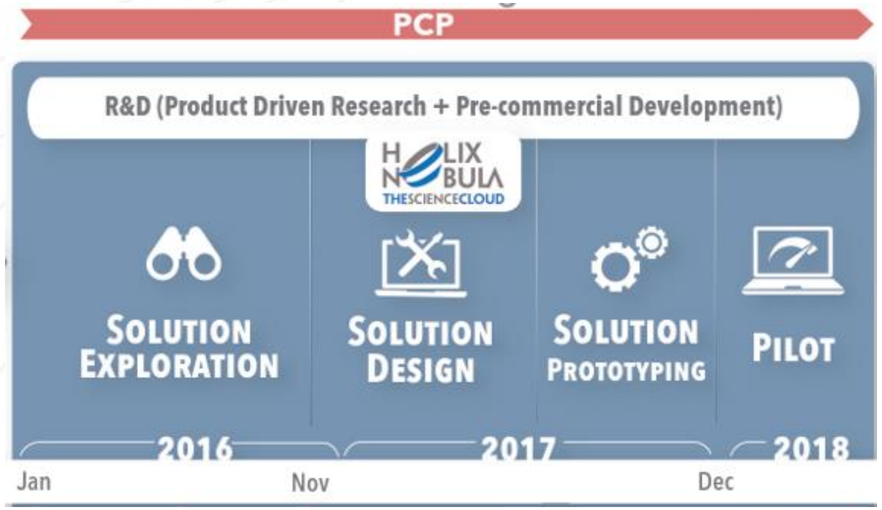
The goal is to demonstrate to the funding agencies that we are in control of the HL-LHC cost, while exploiting the full potential of the physics program



HELIX NEBULA THESCIENCECLOUD

Sharing Open Science Services

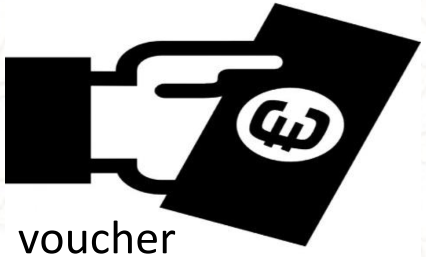




HELIX NEBULA THESCIENCECLOUD

Business models

Pay-as-You-Go	Term Subscription
👍 Pay only for services consumed Adjusts to business requirements No commitment	👍 Discounted pricing/improved ROI Predictable expense
👎 Higher price Expenses can be unpredictable	👎 Payment upfront Committed to a specific term Properly scope and forecast requirements

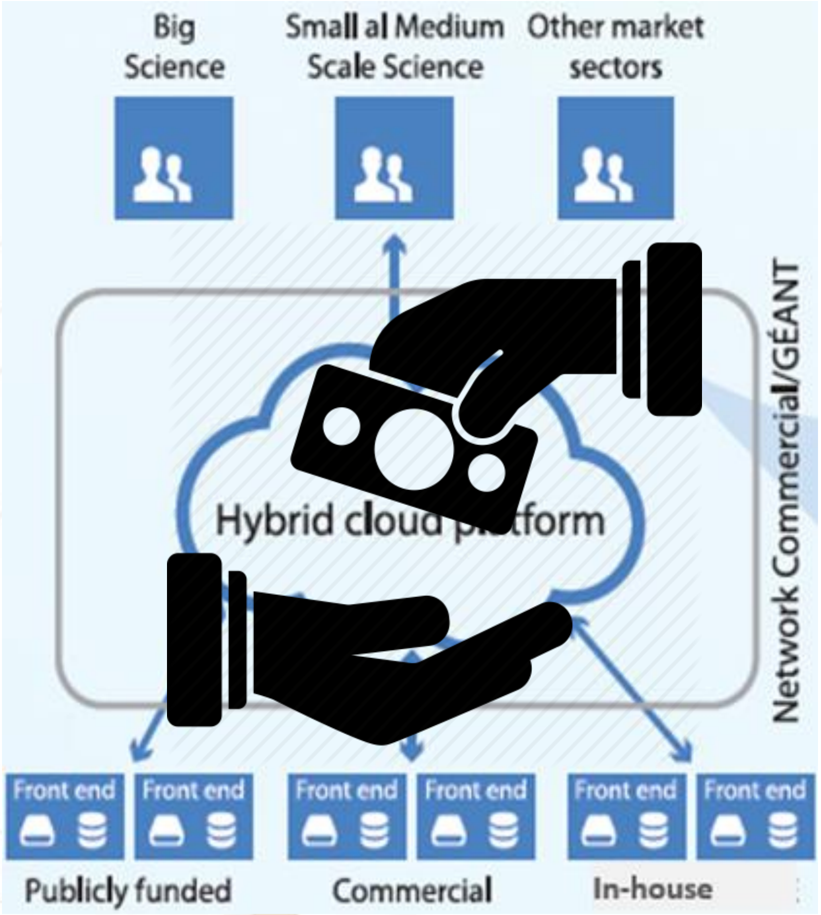


voucher

For long tail of science,
new & exploratory usage,
SLA breach compensation

Data Controller vs.
Data Processor

Need to repatriate data



Thank you for your attention

"The task of the mind is to produce future"

Paul Valéry