

Helix Nebula – The Science Cloud

Introduction and Overview

Helge Meinhard / CERN

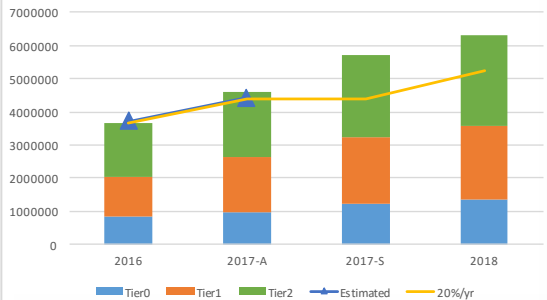
CHEP 2016 San Francisco – contribution # 397

Scale of LHC Data Tomorrow

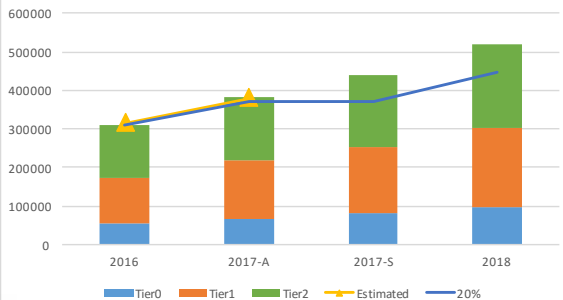
Scale of LHC Data Tomorrow

Short term: until end of Run 2

CPU

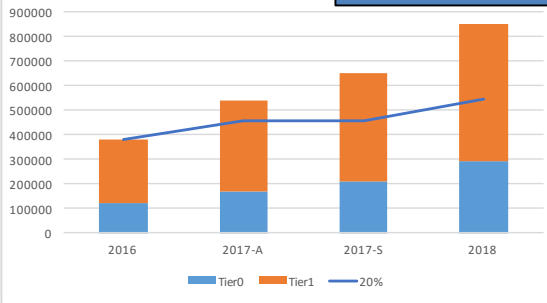


Disk



Estimated: Estimates made in 2014 for Run 2 up to 2017
20%: Growth of 20%/yr starting in 2016 ("flat budget")

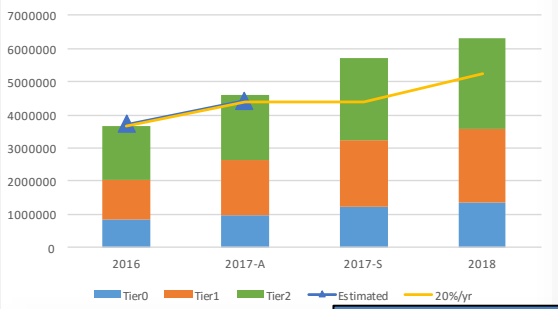
Tape



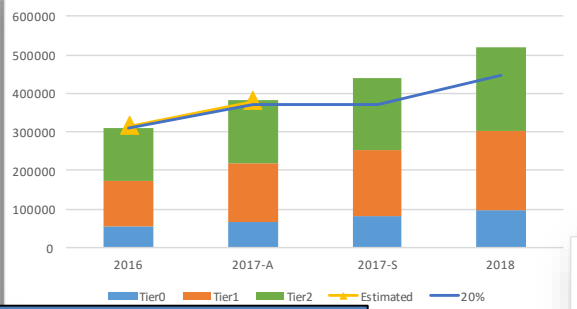
Scale of LHC Data Tomorrow

Short term: until end of Run 2

CPU

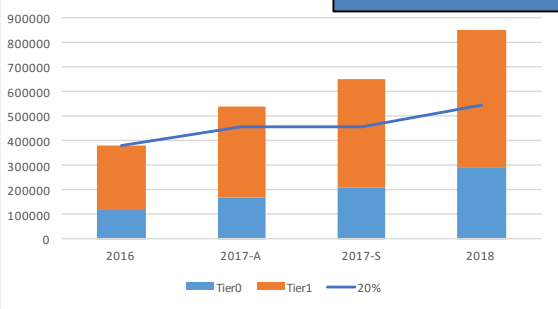


Disk

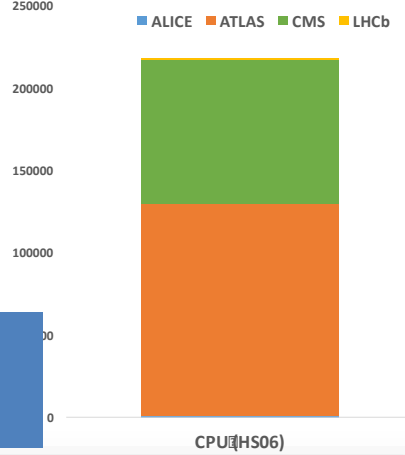


Estimated: Estimates made in 2014 for Run 2 up to 2017
20%: Growth of 20%/yr starting in 2016 ("flat budget")

Tape

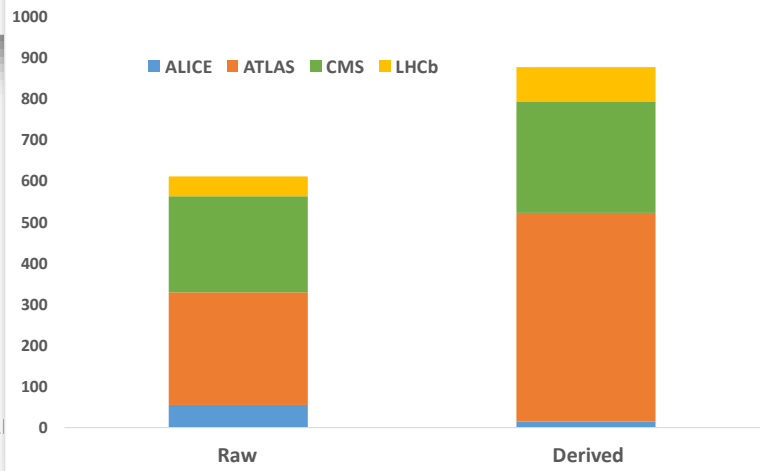


CPU Needs for 1st year of HL-LHC (kHS06)



2026:
HL-LHC

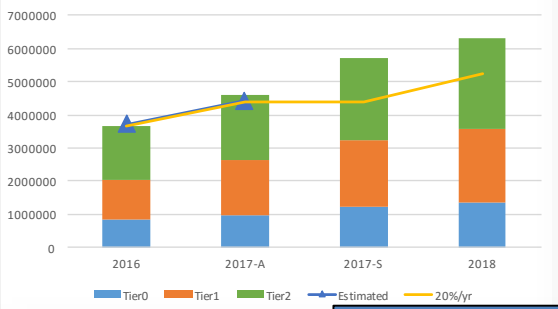
Data Estimates for 1st year of HL-LHC (PB)



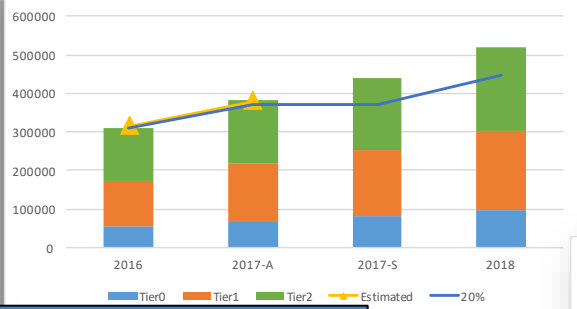
Scale of LHC Data Tomorrow

Short term: until end of Run 2

CPU

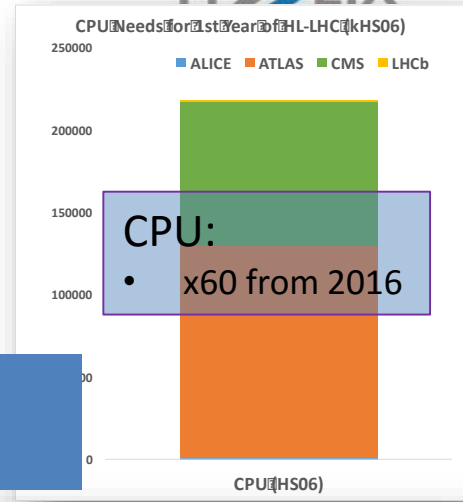
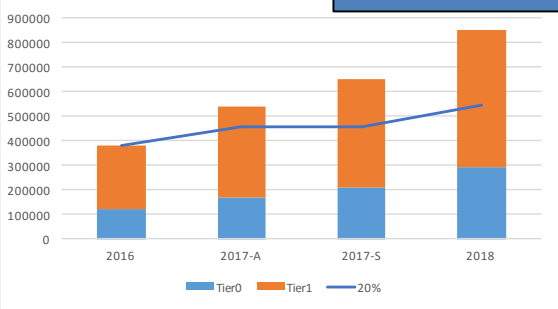


Disk



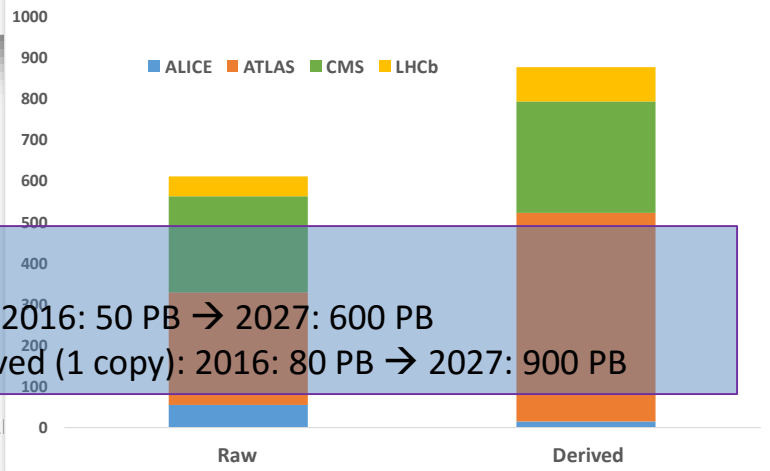
Estimated: Estimates made in 2014 for Run 2 up to 2017
20%: Growth of 20%/yr starting in 2016 ("flat budget")

Tape



2026:
HL-LHC

Data Estimates for 1st year of HL-LHC (PB)

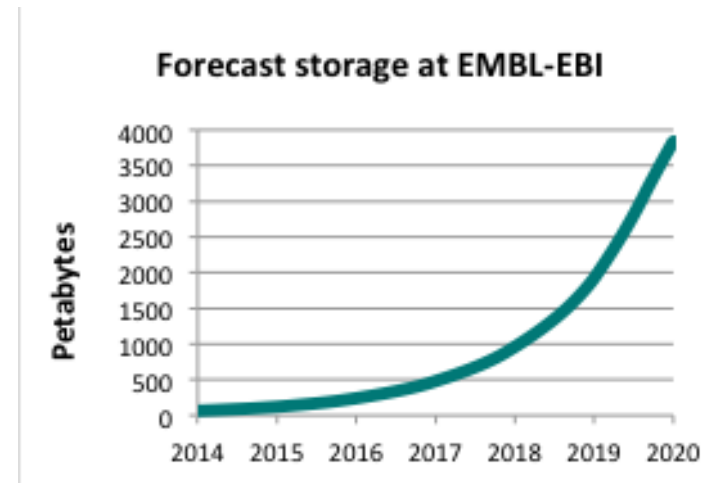


Data:

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

Future Requirements

- ☞ Not only LHC, but a number of particle physics projects with high data rates
- ☞ Not only particle physics, but also other physics fields (e.g. astronomy)
- ☞ Not only physics, but also other sciences (e.g. life sciences, material science)
 - ☞ E.g. EBI: Data doubles every 12 months



Scaling up: Public Clouds (1)

- Additional resources, perhaps later replacing on-premise capacity
- Potential benefits:
 - Economy of scale
 - More elastic, adapts to changing demands
 - Somebody else worries about machines and infrastructure

Scaling up Further: Public Clouds (2)

- ☛ Potential issues:
 - ☛ Cloud provider's business models not well adapted to procurement rules and procedures of public organisations
 - ☛ Lack of skills for and experience with procurements
 - ☛ Market largely not targeting compute-heavy tasks
 - ☛ Performance metrics/benchmarks not established
 - ☛ Legal impediments
 - ☛ Not integrated with on-premise resources and/or publicly funded e-infrastructures

HELIX NEBULA Science Cloud

Joint Pre-Commercial Procurement

Procurers: CERN, CNRS, DESY, EMBL-EBI, ESRF,
IFAE, INFN, KIT, SURFSara, STFC
Experts: Trust-IT & EGI.eu

The group of procurers have committed

- >1.6M€ of procurement funds
- Manpower for testing/evaluation
- Use-cases with applications & data
- In-house IT resources

To procure innovative IaaS level cloud services integrated into a hybrid cloud model

- Commercial cloud services
- European e-Infrastructures

Services will be made available to end-users from many research communities

Co-funded via H2020 (Jan'16-Jun'18)

- Grant Agreement 687614

Total procurement commitment >5M€



User groups to be supported

- ☛ High Energy Physics
 - ☛ LHC experiments
 - ☛ Belle II
 - ☛ COMPASS

- ☛ Astronomy
 - ☛ CTA – Cherenkov Telescope Array
 - ☛ MAGIC
 - ☛ Pierre Auger Observatory

- ☛ Life Sciences
 - ☛ ELIXIR
 - ☛ Euro-BioImaging
 - ☛ Pan-Cancer
 - ☛ BBMRI
 - ☛ WeNMR

- ☛ Photon/Neutron science
 - ☛ PETRA III, European XFEL, 3DIX, OCEAN, OSIRIS

- ☛ Long tail of science



Technical Challenges

☞ Compute

- ☞ Integration of some HPC requirements

☞ Storage

- ☞ Caching at provider's site, if possible automatically (avoid managed storage)

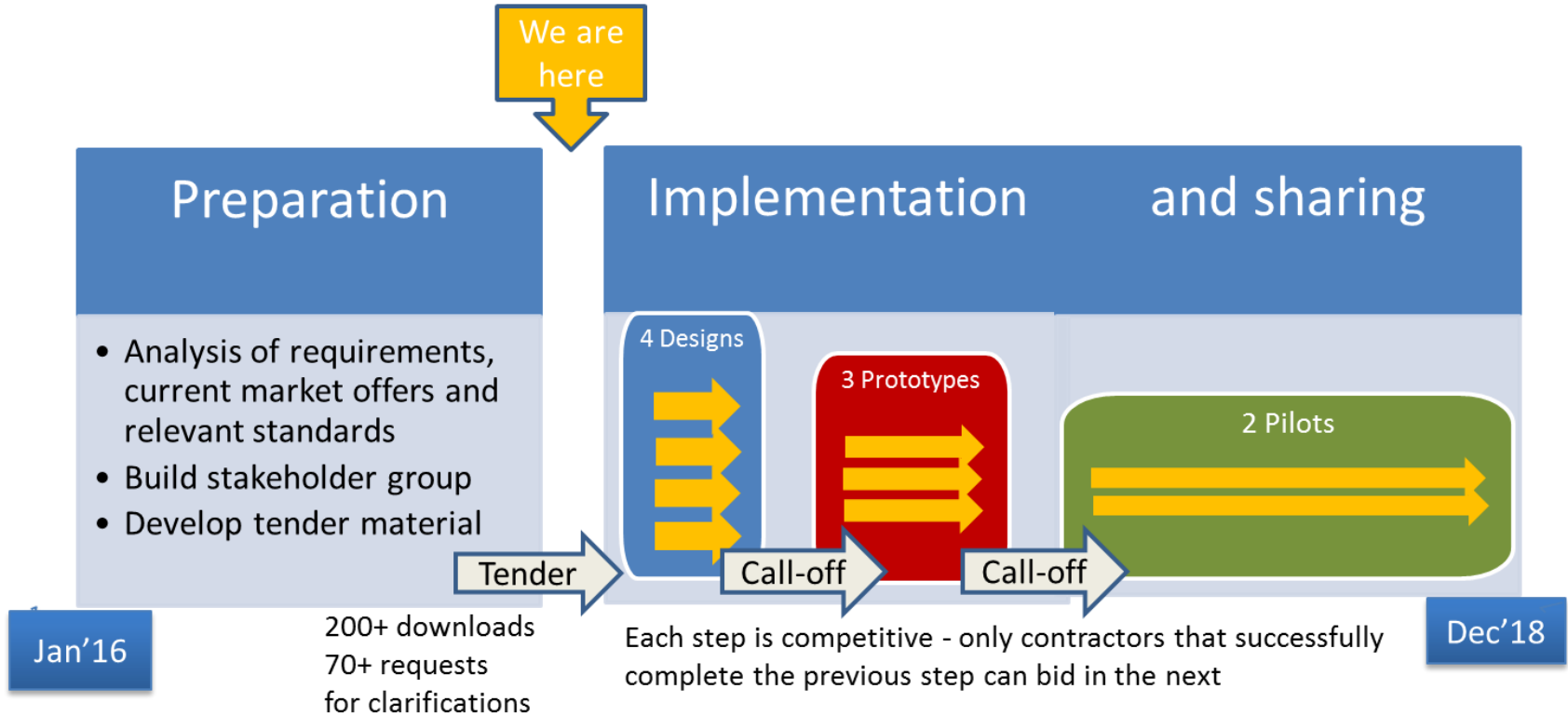
☞ Network

- ☞ Connection via GÉANT
- ☞ Support of identity federation (eduGAIN) for IT managers

☞ Procurement

- ☞ Match of cloud providers' business model with public procurement rules

HNSciCloud Project Phases



HNSciCloud – Current Status

- ☞ Official start of project: Jan 2016, duration: 30 months
- ☞ Tender announced in Jan 2016
- ☞ 17-Mar-2016: Open market consultation
- ☞ 21-Jul-2016: Tender issued (> 200 downloads, > 70 requests for clarification)
- ☞ 07-Sep-2016: Tender information day – design phase
- ☞ 19-Sep-2016: Deadline for tender replies
 - ☞ Sufficient number of valid tenders received
 - ☞ Evaluation by administrative and technical experts
- ☞ 07-Oct-2016: Award decision, contracts
- ☞ 02-Nov-2016: Kick-off meeting with contractors

Summary

- ☞ Commercial cloud services are expected to play an increasing role in the computing models of scientific Research Infrastructures as part of a hybrid cloud platform
- ☞ Such a hybrid cloud platform has the potential to serve many high-profile research projects
- ☞ **Helix Nebula Science Cloud** is a Pre-Commercial Procurement project with a budget of more than 5M€ that is co-funded by the European Commission
 - ☞ The objective is to produce a hybrid cloud platform for the European research community
- ☞ Changes to the procurement process in the public research sector are necessary to benefit from a dynamic Digital Single Market and should be supported by the platform
- ☞ A hybrid cloud poses a number of technical challenges that are being addressed by **Helix Nebula Science Cloud**
- ☞ **Helix Nebula Science Cloud** is the first in a foreseen series of EC co-funded projects which will contribute to the European Cloud Initiative