HNSciCloud Overview

29th November 2018

Bob Jones

CERN

IT department

Helix Nebula Science Cloud

- **Provide a common cloud platform for the European research community**

- Via a collective effort of 10 procurer Institutions forming the **Buyers Group**

Expressing the need to increase the analysis capability and capacity offered to their users to keep pace with the growth in scientific data
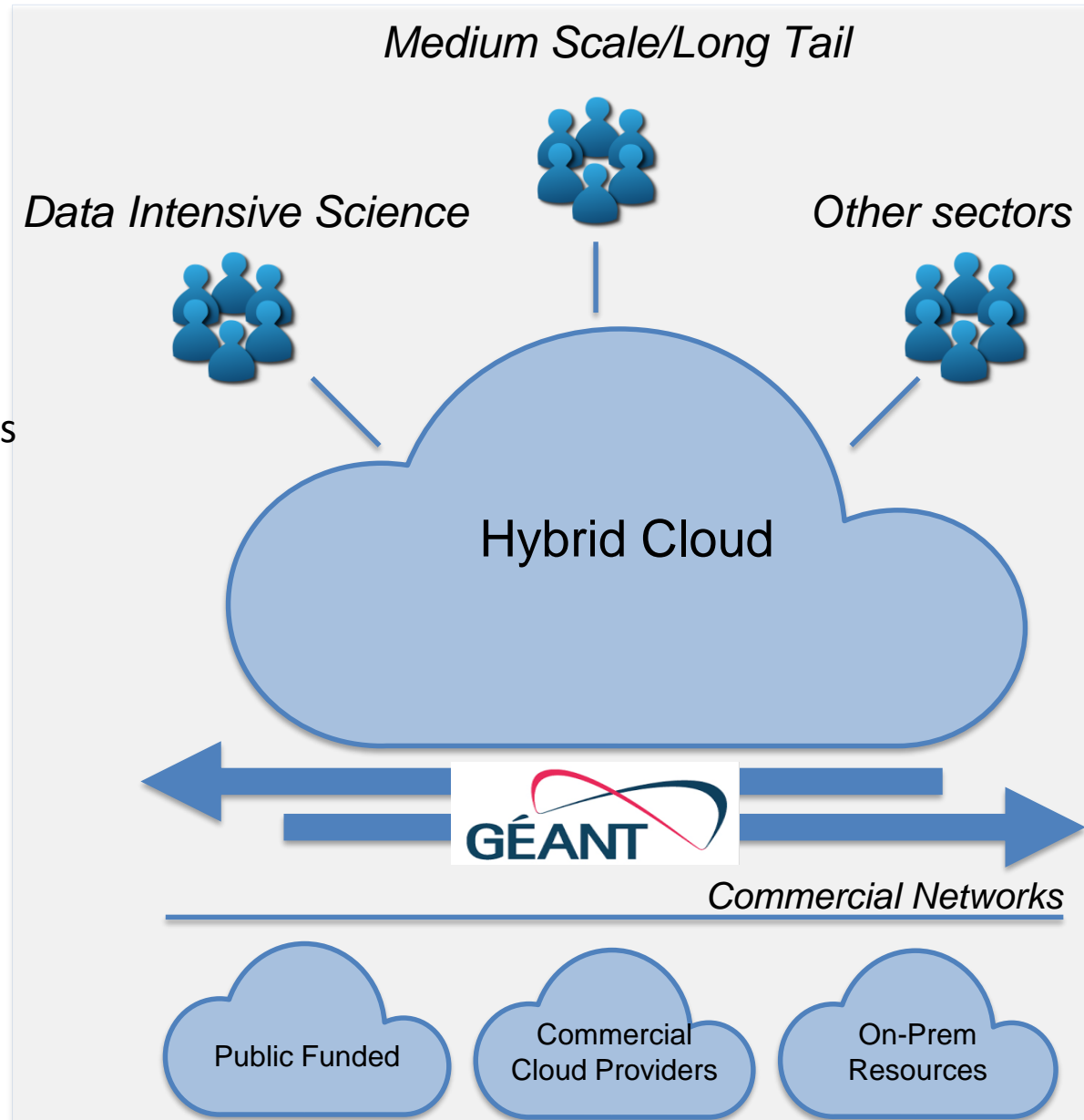
Helix Nebula Hybrid Cloud Model

Bringing together:

- Research Organisations
- Data Providers
- Publicly funded e-infrastructures
- Commercial cloud providers

With:

Procurement and Governance suitable for the dynamic cloud market

Helix Nebula Science Cloud
Joint Pre-Commercial Procurement

Procurers: **CERN, CNRS, DESY, EMBL-EBI, ESRF, IFAE, INFN, KIT, STFC, SURFSara**

Experts: Trust-IT & EGI.eu

Resulting IaaS services support use-cases from many research communities

Deployed in a hybrid cloud mode:
- procurers data centres
- commercial cloud service providers
- GEANT network and EduGAIN Federated Identity Management.

Co-funded via H2020 Grant Agreement 687614
Total procurement budget >5.3M€

Challenges

Innovative IaaS cloud services integrated with procurers in-house resources to support a range of scientific workloads

- *Compute and Storage*
  - Support a range of virtual machine and container configurations including HPCaaS, working with datasets in the petabyte range, accessible transparently

- *Network Connectivity and Federated Identity Management*
  - Provide high-end network capacity via GEANT for the whole platform with common federated identity and access management

- *Service Payment Models*
  - *Explore* a range of purchasing options to determine those most appropriate for the scientific application workloads, including *vouchers* or other means of easy integration in the organisations procurement models
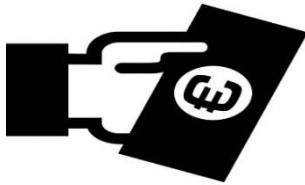
Service Payment Models

Single tender jointly funded by the procurers and the EC

Free at the point of use for end-users nominated by procurers

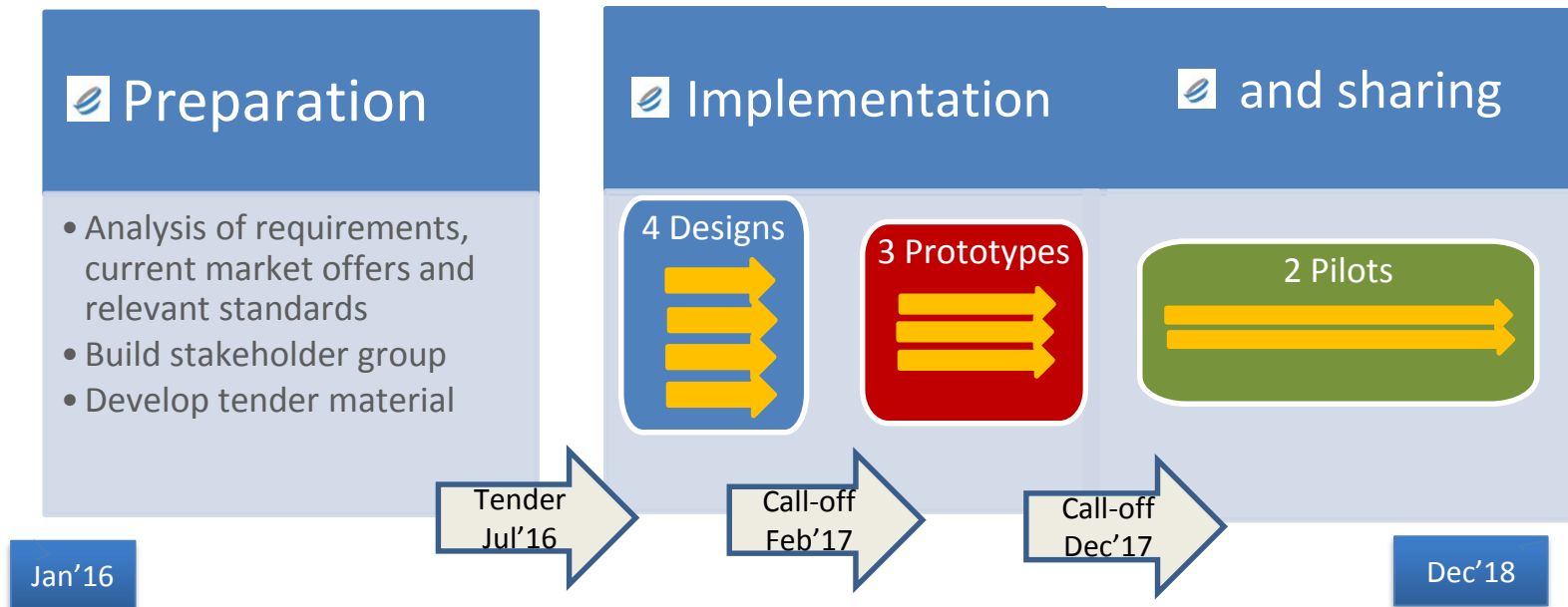Vouchers for long tail of science, new & exploratory usage, etc.

Monitoring Service quality according to Service Level Agreements

Role of Data Controller vs. Data Processor

Need to repatriate data at end of contract

HNSciCloud project phases

Preparation

- Analysis of requirements, current market offers and relevant standards
- Build stakeholder group
- Develop tender material

Jan'16

Implementation   and sharing

4 Designs

3 Prototypes

2 Pilots

Tender
Jul'16

Call-off
Feb'17
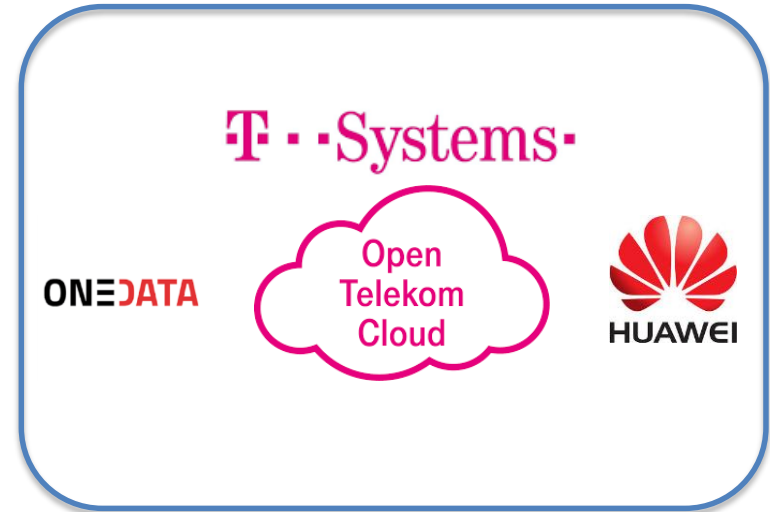
Call-off
Dec'17

Dec'18

Each step is competitive - only contractors that successfully complete the previous step can bid in the next

**Phases of the tender are defined by the Horizon 2020
Pre-Commercial Procurement financial instrument**

Pilot Cloud Platforms

- **T-Systems**

  IaaS based on OTC

- **RHEA**

  IaaS provided by Exoscale

Pilot Phase Timeline

10k/1PB

2 Pilots

5k/500TB

3.5k350TB

Scalability Testing   **End User Access**

Cores/Storage per Contractor
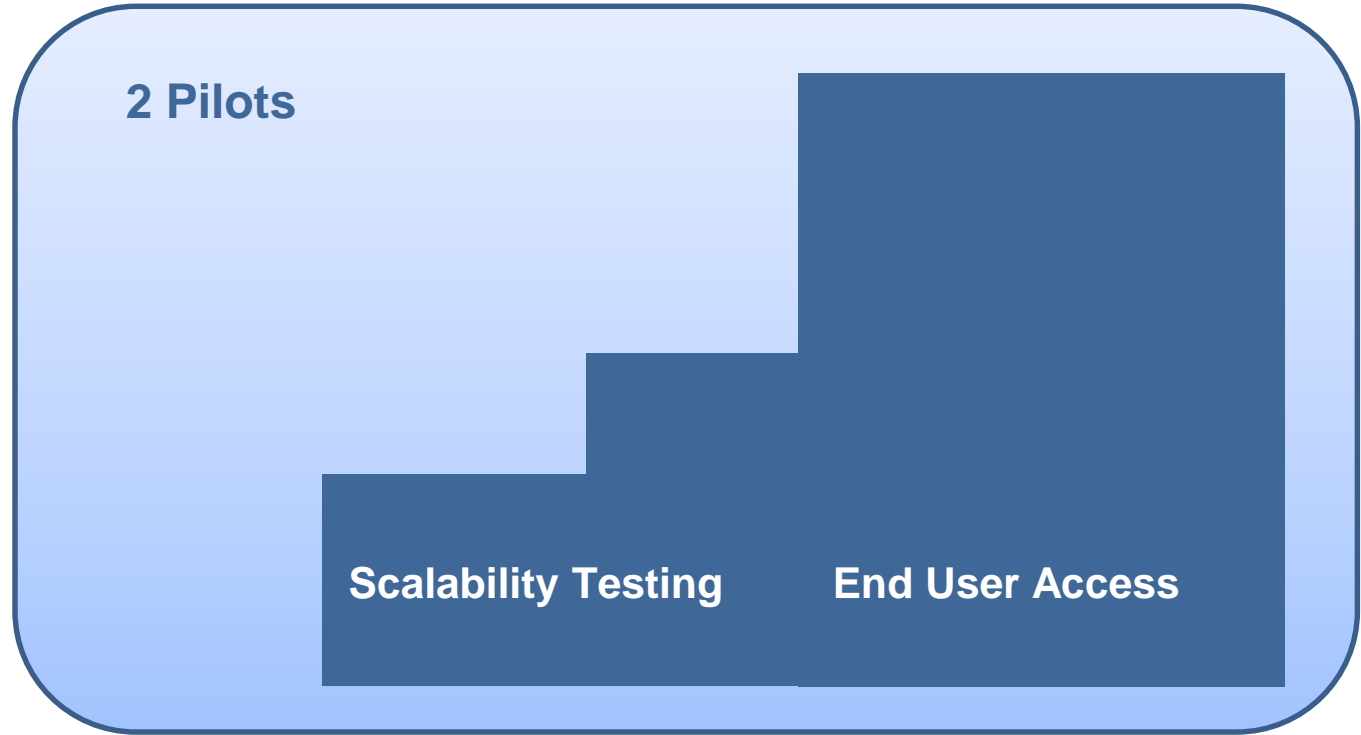
Call-off Dec'17

Network 40Gbps

Feb'18

Dec'18

HNSciCloud promoted as a working example by the EC High Level Expert Group

The HNSciCloud consortium would like to thank the following for their contributions to the project

Companies/Projects/Organisations

AARC(2)	Consultancy for federated AAI
Addestino	open market consultation
Cloud For Europe	PCP experience and lessons learned
DG CNECT	Dirk Van Rooy and Lieve Bos for their support and guidance
eInfraCentral	Registered pilot services in the central service catalogue
ELIXIR	Use-case and AAI support
EOSC-Hub	Integration of voucher scheme with EGI Application on Demand service
EOSC-Pilot	Support for several of the use-cases deployed
ESA	Hosted event during the preparation phase
GEANT	Established the network connectivity between contractors and buyers group premises
IBM	Participation in the design and prototype phases
INDRA	Participation in the design phase
Oxford Academics	Gilles Valet & students of the Postgraduate Programme in Project Engineering for risk analysis
SLALOM & SLA-Ready	Provided basis for the common Service Level Agreement
StrategicBlue	Review tender material and financial benchmarking
WLCG	Sponsoring a major use-case for the physics community, operational security challenge

Use cases to be presented		
Pan-Cancer large scale genomics analyses in cancer studies	Life sciences	Tony Wildish (EMBL)
CrystFEL Serial Femtosecond Crystallography	Photon/Neutron Sciences	Sergey Yakubov (DESY)
FDMNES simulation of x-ray spectra	Photon/Neutron Sciences	Rainer Wilcke (ESRF)
MAGIC	Astronomy	Jordi Casals (IFAE)
DODAS-CMS on T-Systems resources (INFN)	High Energy Physics	Giuseppe La Rocca (EGI.eu)
Dynamic Batch System Extension into the HNSciCloud	High Energy Physics	Preslav Konstantinov (KIT)
LOFAR ASTRON Low Frequency Array	Astronomy	Martin Brandt (SURFsara)
Consolidated batch queue for the LHC experiments	High Energy Physics	Ben Jones (CERN)
Interactive Data Analysis for End Users on HNSciCloud	High Energy Physics	Jakub Moscicki (CERN)
Ceph S3 Cloud Integration Tests	High Energy Physics	Roberto Valverde (CERN)
AI-assisted prostate cancer pathology with deep learning	Life sciences	Marco Capuccini (Univ. Uppsala)
A Deep Learning tool for fast simulation	High Energy Physics	Sofia Vallecorsa (CERN)