





# Helix Nebula Science Cloud

## 2<sup>nd</sup> ASTERICS-OBELICS Workshop

16-19 October 2017, Barcelona, Spain.

Bob Jones

CERN

Bob.Jones <at> cern.ch



Helix Nebula – The Science Cloud

Helix Nebula – The Science Cloud with Grant Agreement 687614 is a Pre-Commercial Procurement Action funded by H2020 Framework Programme



# HNSciCloud Joint Pre-Commercial Procurement



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

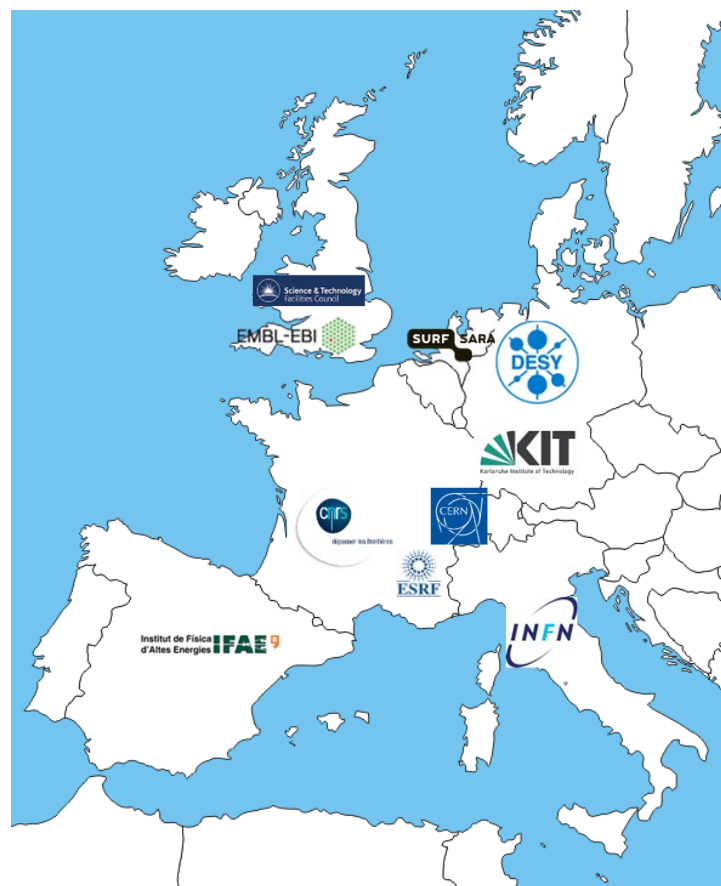
The group of procurers have committed

- Procurement funds
- Manpower for testing/evaluation
- Use-cases with applications & data
- In-house IT resources

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

Co-funded via H2020 Grant Agreement 687614

**Total procurement budget >5M€**

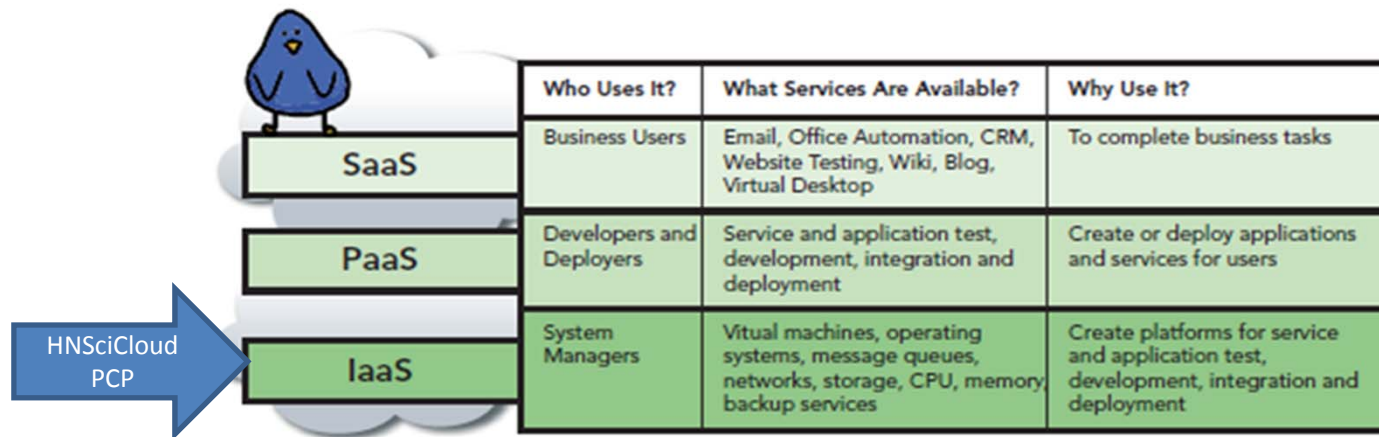


# What is being procured

A hybrid cloud platform for the European research community



Combining services at the IaaS level to support science workflows



Source: CloudComputing for Govies, DLT Solutions, David Blankenhorn, Van Ristau and Caron Beesley

The R&D services to be developed are to be integrated with Resources in data centres operated by the Buyers Group  
GEANT network



# Challenges



Innovative IaaS level services integrated with procurers in-house resources and public e-infrastructure as part of a hybrid cloud to support a range of scientific workloads

## *☛ Compute and Storage*

- ☛ support a range of virtual machine and container configurations including HPC working with datasets in the petabyte range

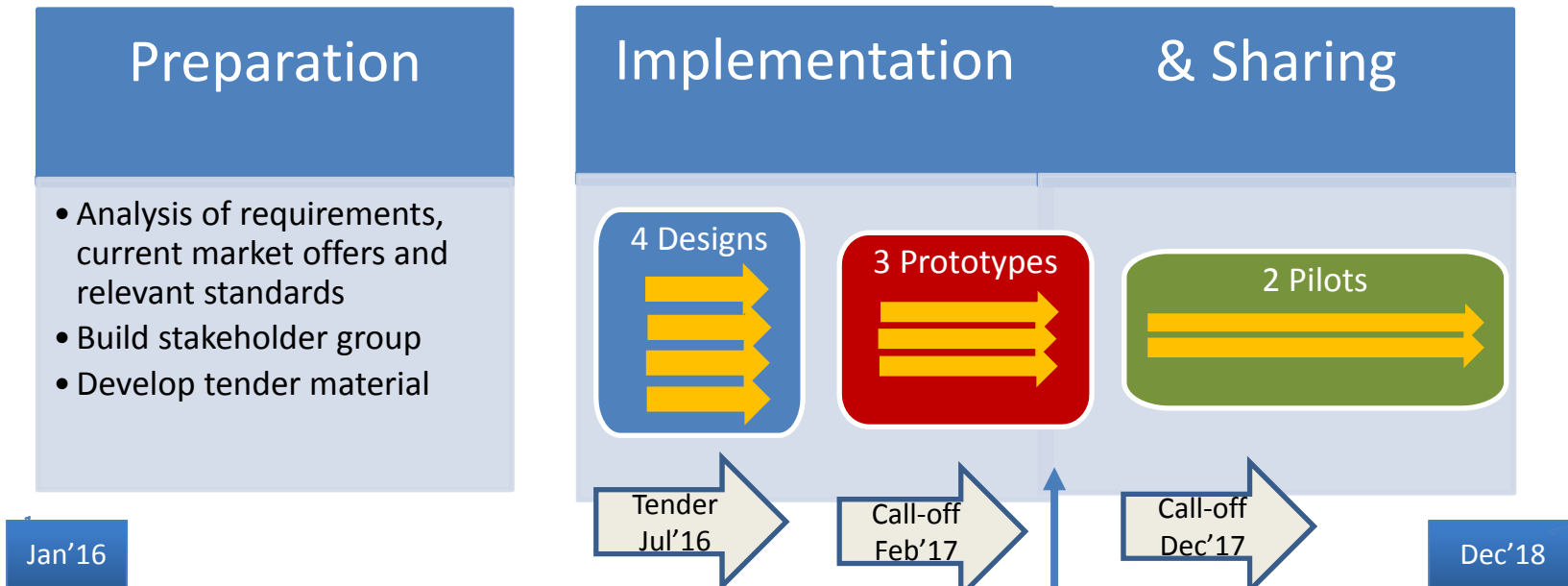
## *☛ Network Connectivity and Federated Identity Management*

- ☛ provide high-end network capacity via GEANT for the whole platform with common identity and access management

## *☛ Service Payment Models*

- ☛ explore a range of purchasing options to determine those most appropriate for the scientific application workloads to be deployed

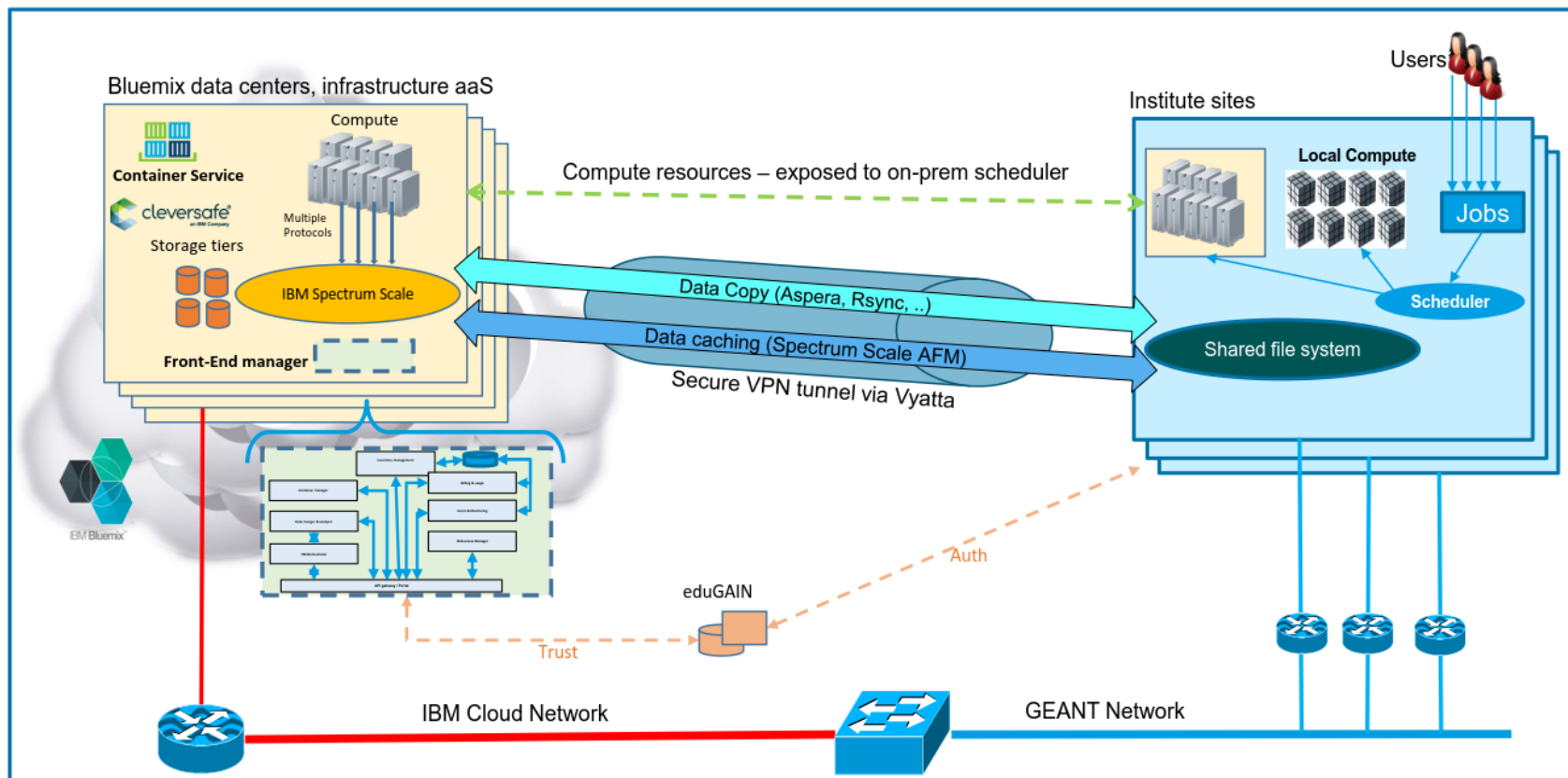
# HNSciCloud project phases



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



# Main components - High level view



# T Systems

## PROTOTYPE - DATA MANAGEMENT

ONE DATA

openstack.

POWERED

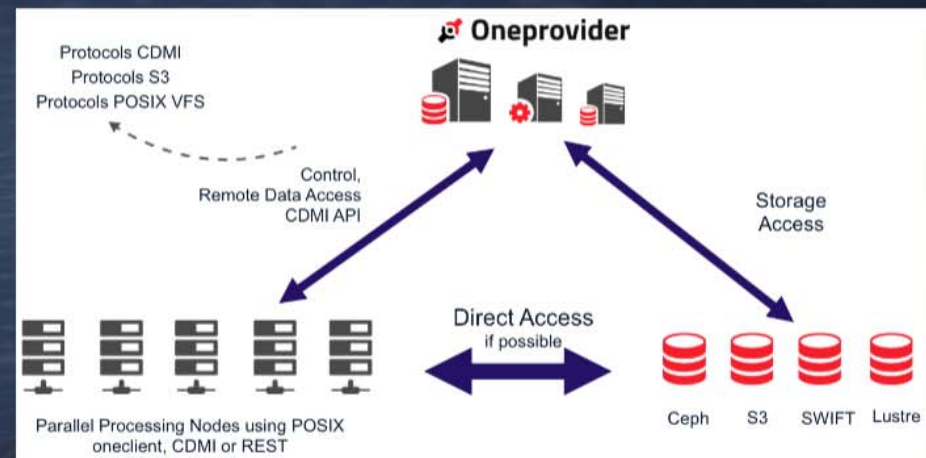
### NEW OPENSTACK RELEASE (MITAKA)

#### STORAGE FEDERATIONS AND MULTI-TENANCY

- Ease of access to distributed data in hybrid cloud
- Secure data sharing between users

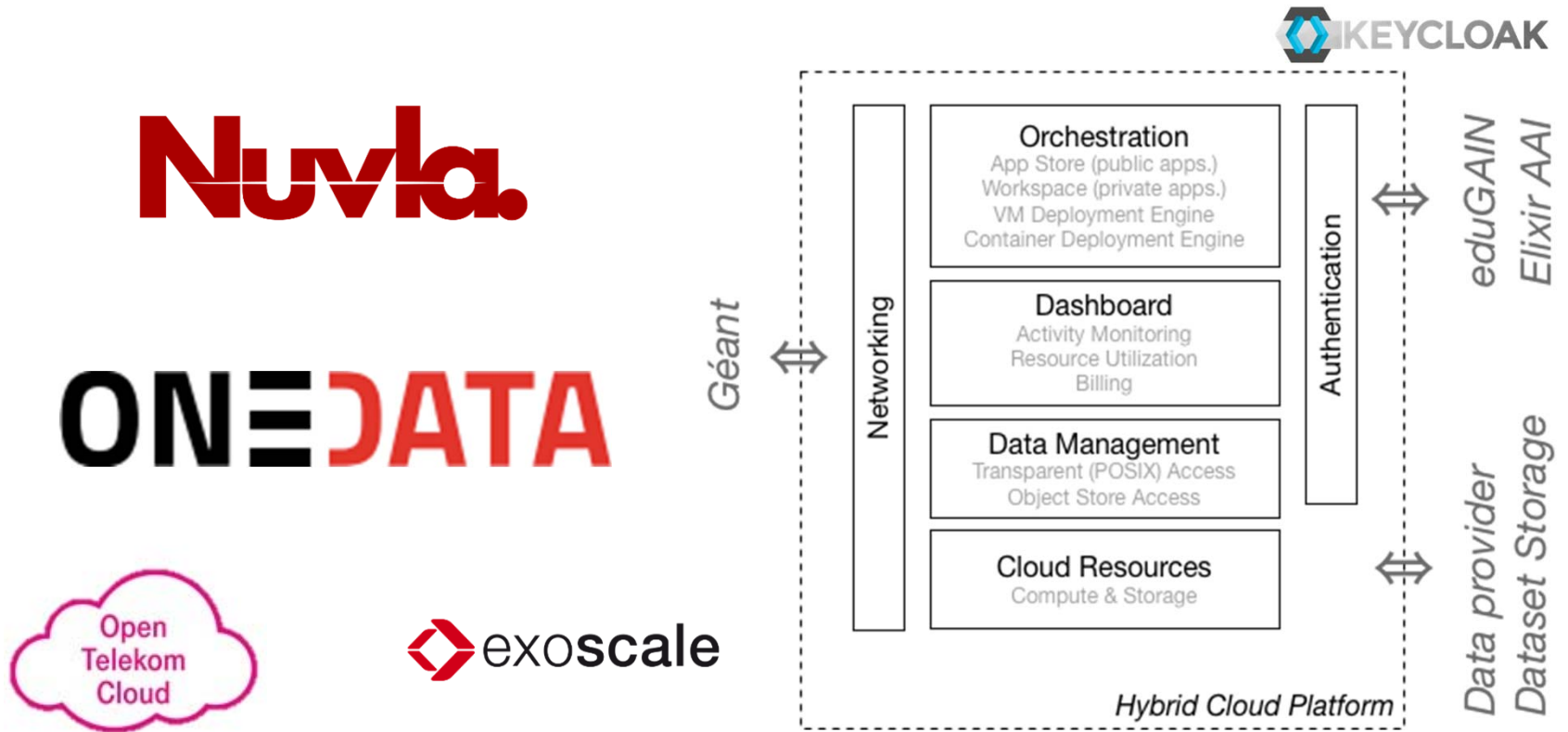
#### UNIFIED TRANSPARENT ACCESS TO HYBRID-CLOUD STORAGE RESOURCES

- Transparent data access across hybrid-cloud deployment based on POSIX and CDMI

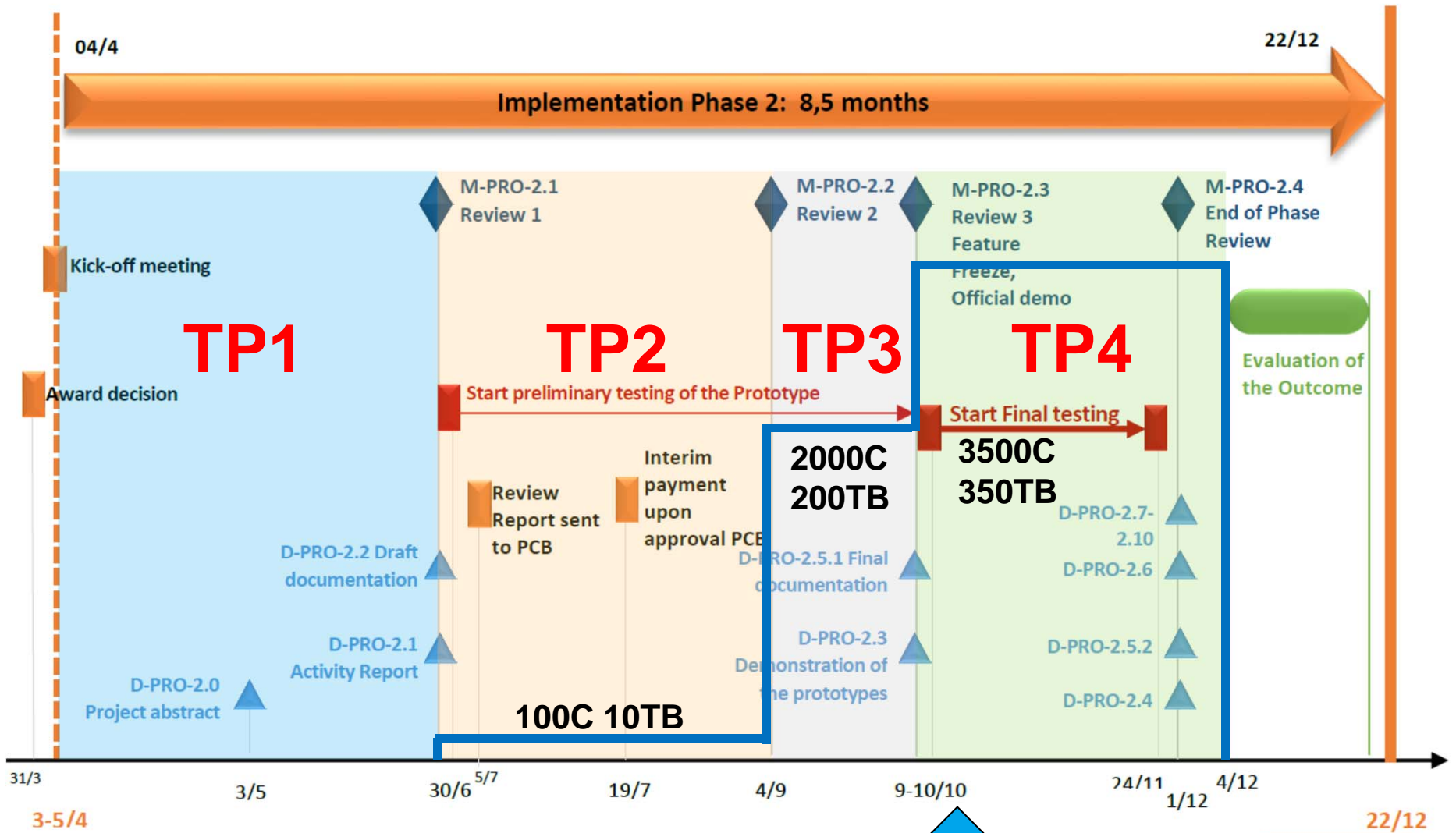




## Overview of Prototype Solution



# Test Phases & Resource Ramp-Up



# Test Suite



## CERN

- Edugain AAI
- CERN Benchmark Suite
- S3 Endpoints
- PerfSONAR Network
- VM Provisioning and Personalisation
- Grid Jobs in Docker
- Dockerized deployment of EOS+CERNBOX+SWAN

## CNRS

- IAAS access via EduGain and local accounts

## DESY

- HDF5\_IO

## EMBL

- CLOUD\_BENCHMARK
- ELIXIR\_AAI
- data\_transparency

## ESRF

- FDMNES

## IFAE

- monitoring-accounting
- image creation and contextualization

## INFN

- BelleII\_SIM
- COMPASS/TGEANT\_IO
- INFN/GPFS\_AFM
- Dynfarm

## KIT

- static-batch-extension
- dynamic-batch-extension
- AFM-deployment-test

## STFC

- batch-system-extension
- jobs-on-kubernetes

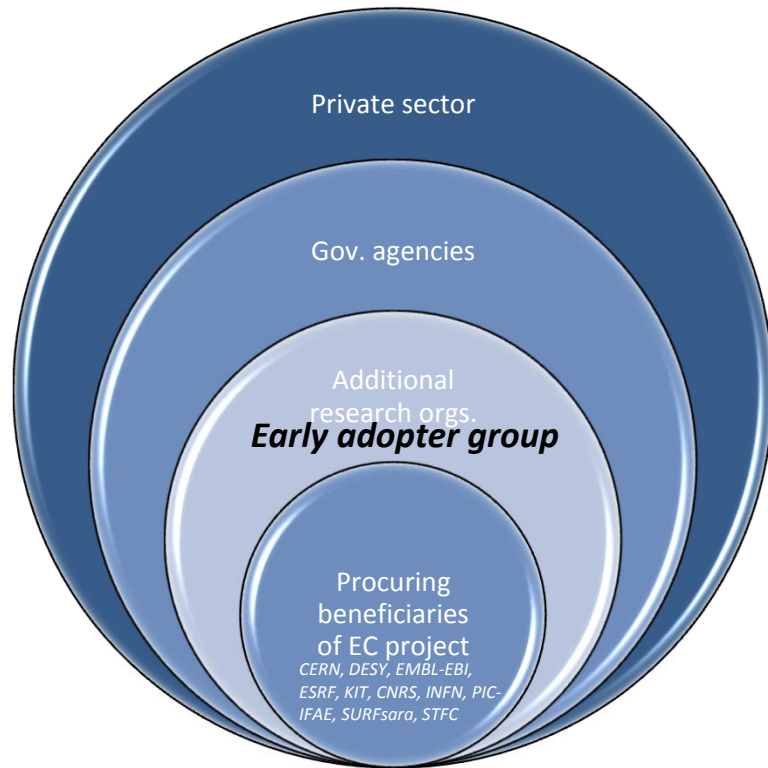
## SURFsara

- BBMRI\_data
- dCache\_io
- kubertest

# Pilot Phase 2018

- ☞ Kick-off event 6-8 Feb 2018 (INFN Bologna)
- ☞ Progressively scale-up IaaS resources
  - ☞ totalling 20k cores 2PB storage
- ☞ Scalability testing
- ☞ Deployment of applications
- ☞ Assess procurement/payment models

# Growing the buyers group



The initial group of buyers is the set of research orgs. that committed their resources at the start and became beneficiaries of the H2020 project

Now starting to grow the buyers group by including more publicly funded research organisations: **Early adopter group**

## Benefits

- Access commercially supported cloud services selected and tested by HNSciCloud research community procurers
- No need to perform your own tender
- Ability to fund the use of the services via your regional, national or EC projects
- Use the legal framework and access conditions established by HNSciCloud



## European Open Science Cloud

### Current and future EU support to the EOSC and EC way forward from EOSC Summit to Roadmap for Implementation

Carmela Asero,

European Commission, DG Research & Innovation

e-IRG workshop - Tallinn, 3 October 2017



EUROPEAN COMMISSION  
DIRECTORATE-GENERAL FOR RESEARCH & INNOVATION

The Director-General

Brussels, 30 July 2017

### **EOSC Declaration**

RECOGNISING the challenges of data driven research in pursuing excellent science;

GRANTING that the vision of European Open Science is that of a research data commons, widely inclusive of all disciplines and Member States, sustainable in the long-term;

CONFIRMING that the implementation of the EOSC is a process, not a project, by its nature iterative and based on constant learning and mutual alignment;

UPHOLDING that the EOSC Summit marked the beginning and not the end of this process, one based on continuous engagement with scientific stakeholders, the European Commission;

PROPOSES that all EOSC stakeholders consider sharing the following intents and will actively support their implementation in the respective capacities:

#### **Data culture and FAIR data**

- **[Data culture]** European science must be grounded in a common culture of data stewardship, so that research data is recognised as a significant output of research and is appropriately curated throughout and after the period conducting the research. Only a considerable cultural change will enable long-term reuse for science and for innovation of data created by research activities: no disciplines, institutions or countries must be left behind.
- **[Open access by default]** All researchers in Europe must enjoy access to an open-by-default, efficient and cross-disciplinary research data environment supported by FAIR data principles. Open access must be the default setting for all results of publicly funded research in Europe, allowing for proportionate limitations only in duly justified cases of personal data protection, confidentiality, IPB concerns, national security or similar (e.g. 'as open as possible and as closed as necessary').
- **[Skills]** The necessary skills and education in research data management, data stewardship and data science should be provided throughout the EU as part of higher education, the training system and on-the-job best practice in the industry. University associations, research organisations, research libraries and other educational bodies play an important role but they need substantial support from the European Commission and the Member States.
- **[Data stewardship]** Researchers need the support of adequately trained data stewards. The European Commission and Member States should invest in the education of data stewards via career programmes delivered by universities, research institutions and other trans-European agents.
- **[Rewards and incentives]** Rewarding research data sharing is essential. Researchers who make research data open and FAIR for reuse and/or reuse and reproduce data should be rewarded, both

## **EC gathering endorsement and commitments on the EOSC declaration**

- EOSC Declaration won't be modified
- Action list/commitment its evolutive part
- Roadmap will follow (and will have future updates)

*By endorsing the principles of the EOSC Declaration, stakeholders signal their intention to be involved in the making of the EOSC (eg. by taking specific action, by joining the Executive Board, by providing inputs via the annual stakeholder forum, or again by joining consortia , to implement the EOSC via Horizon 2020).*

### Objectives:

- Identifying 'doers' out of wide range of stakeholders
- Promote accountability



**DRAFT**

## **WP 2018-2020: Call INFRAEOSC**

**Topic 1:** Access to commercial services through the EOSC hub (2018)

**Topic 2:** Prototyping new services (2019)

**Topic 3:** *Integration & consolidation of pan-European access mechanisms to public e-infrastructures & commercial services through the EOSC hub (2020)*

**Topic 4:** Connecting ESFRI RI through Cluster projects (2018)

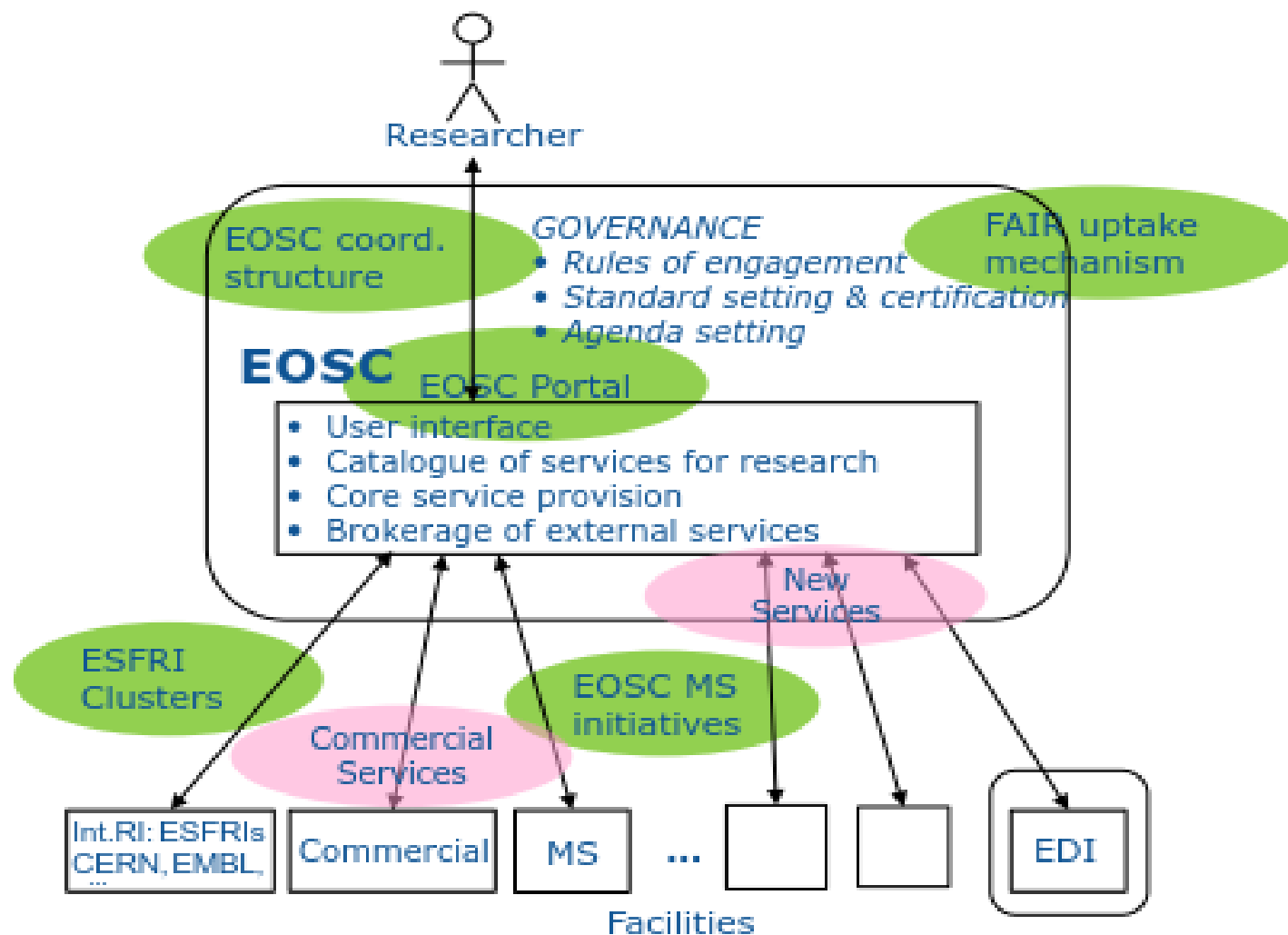
**Topic 5:** Support to the EOSC governance (2018-2019)

**Topic 6:** Enhancing the EOSC portal *and connecting thematic clouds* (2019-2020)



# Mapping of 2018-2019 actions

DRAFT





# EOSC – the way forward

## **EOSC Declaration (Sept 2017)**

- key input for the Roadmap
- for agreement and poss. specific commitments by stakeholders
- basis for follow-up discussion with MS (ERAC SWG 5+)

## **EOSC Stakeholders Forum (Nov 2017)**

- open to ALL categories represented at the Summit & endorsing EOSC Declaration
- by end August, EC to define application procedure & working modalities
- initially based on EOSCpilot project (Stakeholder Engagement Event 28-29 Nov)

## **EOSC Roadmap (Dec 2017)**

- **Governance structure** , incl. mandate & selection procedure for Executive Board
- **Broad (federated) architecture**, incl. EOSC core services
- **Financing**

# Summary

- ☛ 3 hybrid cloud prototypes (IBM, RHEA, T-Systems) have been procured, deployed and tested by a group of 10 research organisations
- ☛ 2/3 prototypes will be selected for large-scale pilots in 2018
- ☛ The *Early adopter group* will allow more research organisations to procure the resulting services
- ☛ HNSciCloud is contributing to the *European Open Science Cloud (EOSC)*
- ☛ There are opportunities for ESFRIs to participate in the 2018-2020 work programme, make use of commercial cloud services and engage with EOSC