





European Open Science Cloud

3rd Digital ERA Forum – Brussels, 23 April 2015
Bob Jones, CERN

Why do we need a European Open Science Cloud?

Large Hadron Collider

CMS

LHCb

ATLAS

ALICE

LHC computing resources in 2014

- 3.4 billion CPU hours
- 100 PetaBytes storage

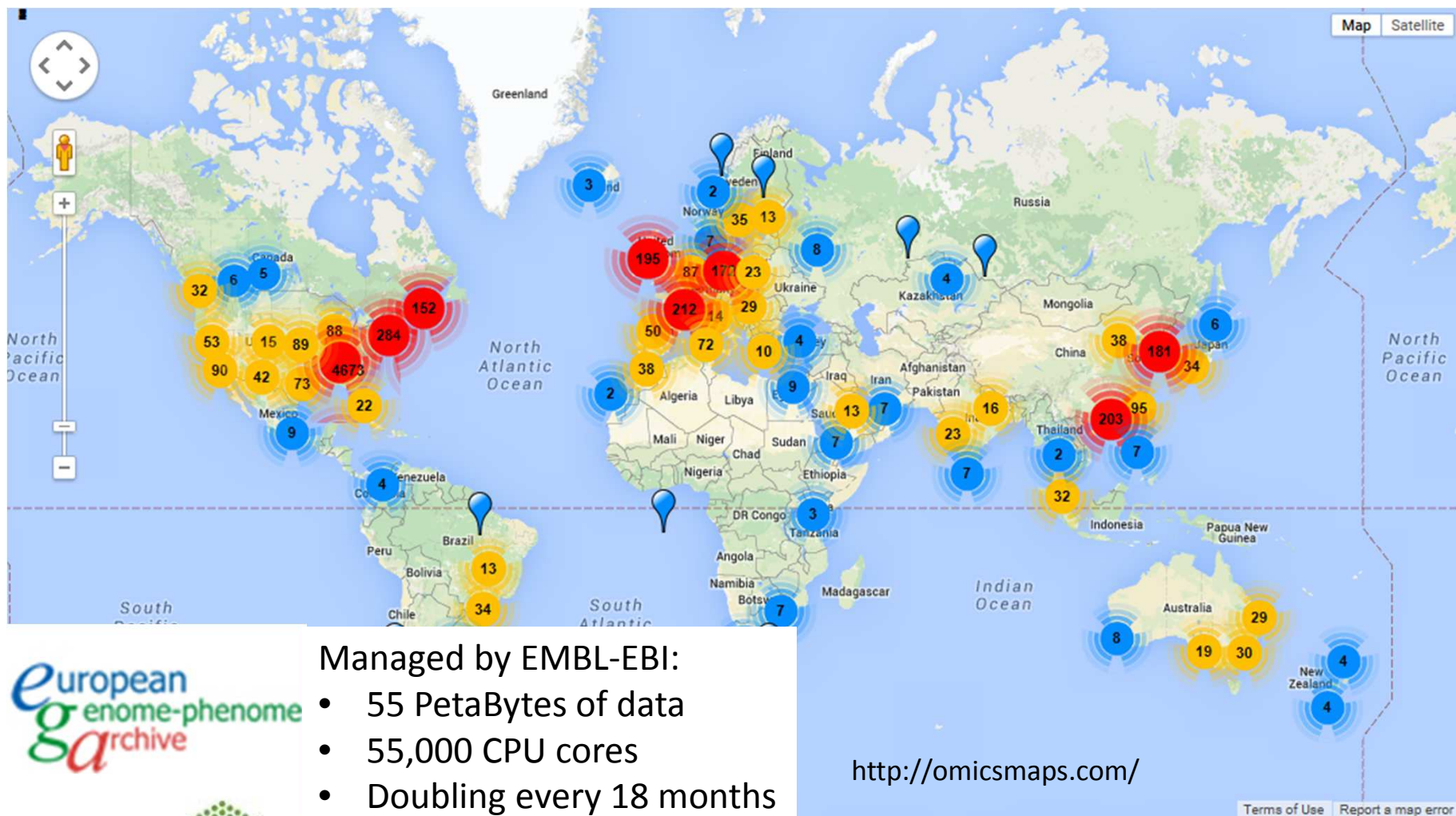
p

p

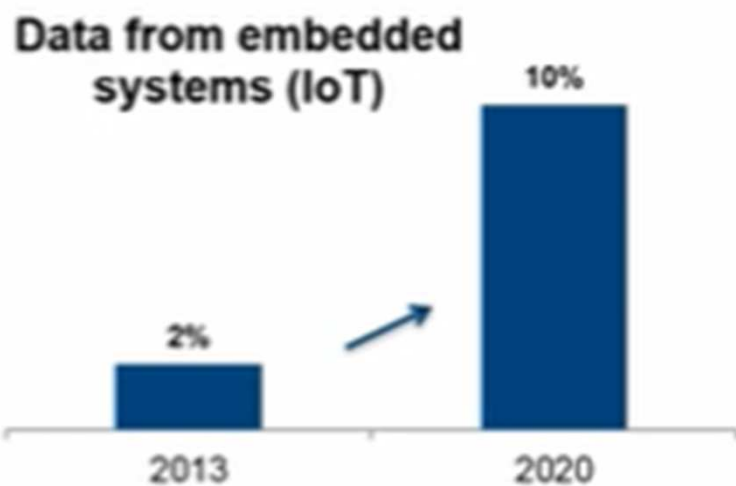


Bob Jones, CERN

Next Generation Genomics: World Map of High-throughput Sequencers

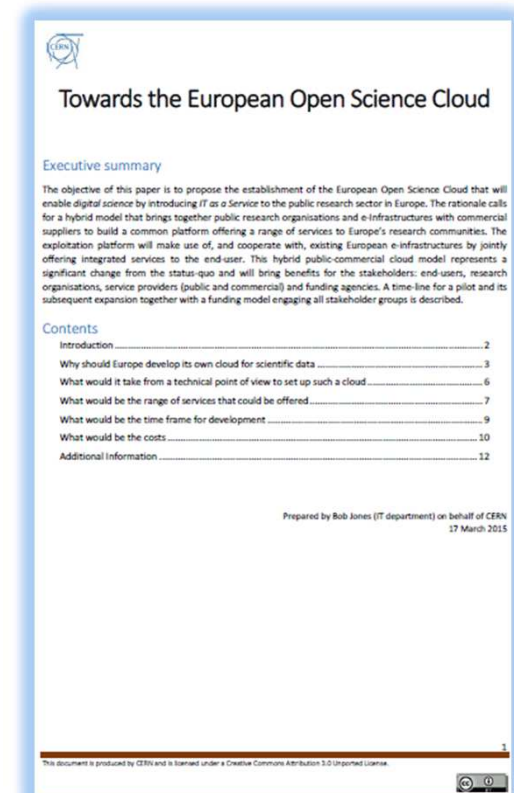


Growth of the Digital Universe from 2013 to 2020



What is the European Open Science Cloud?

- **Hybrid** – link public research organisations, e-Infrastructures & commercial cloud services
  - Use GEANT network to link Research Infrastructures, repositories (EUDAT, OpenAIRE), EGI, PRACE etc. to commodity commercial cloud services (multiple providers)
  - A cornerstone of the Open Science Commons*
- **Trust** - Researchers keep control of the cloud and their data
  - Guarantee a copy of all the data is kept on public resources
  - Ensure long-term preservation of the data
  - Insulate users from changes of service supplier and technology
- **Economy** - Must be cheaper than the '*build our own*' approach
  - Avoid separate 'silos' for each Research Infrastructure/Community
  - Profit from the economies of scale in commercial data centres

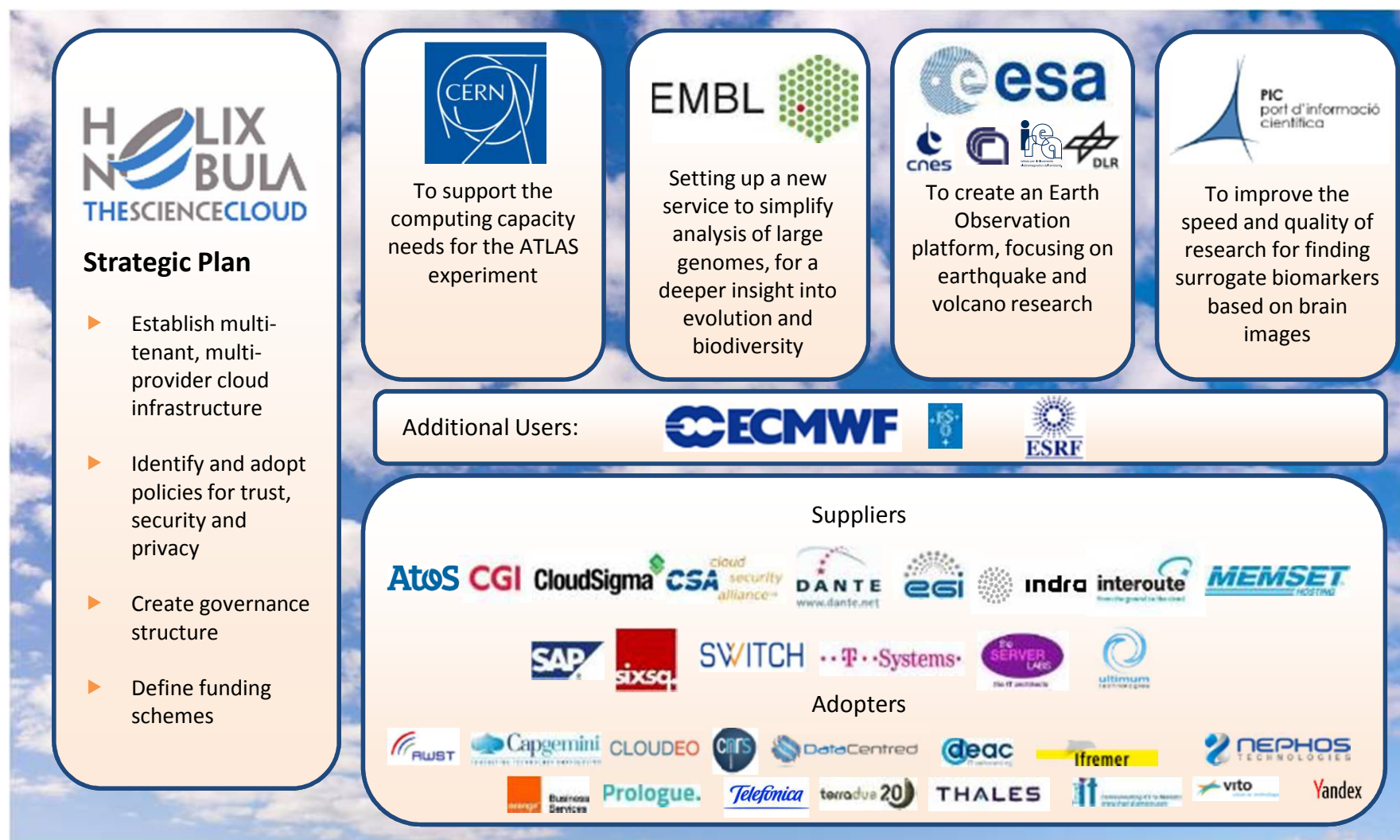


* <http://go.egi.eu/osc>

<http://dx.doi.org/10.5281/zenodo.16140>

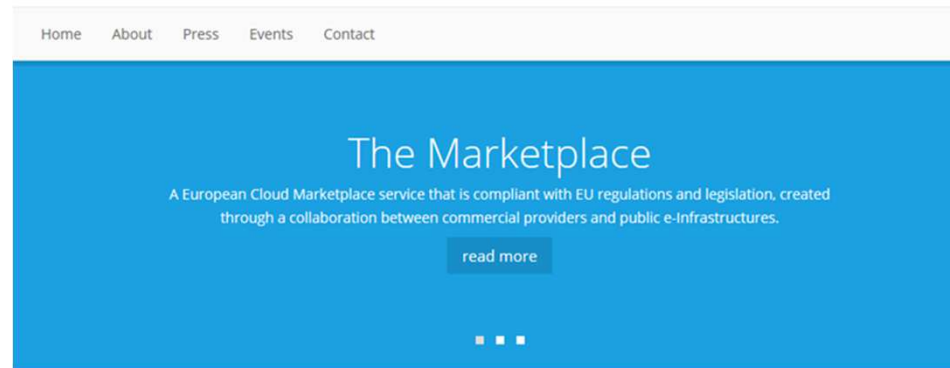


The Helix Nebula Initiative



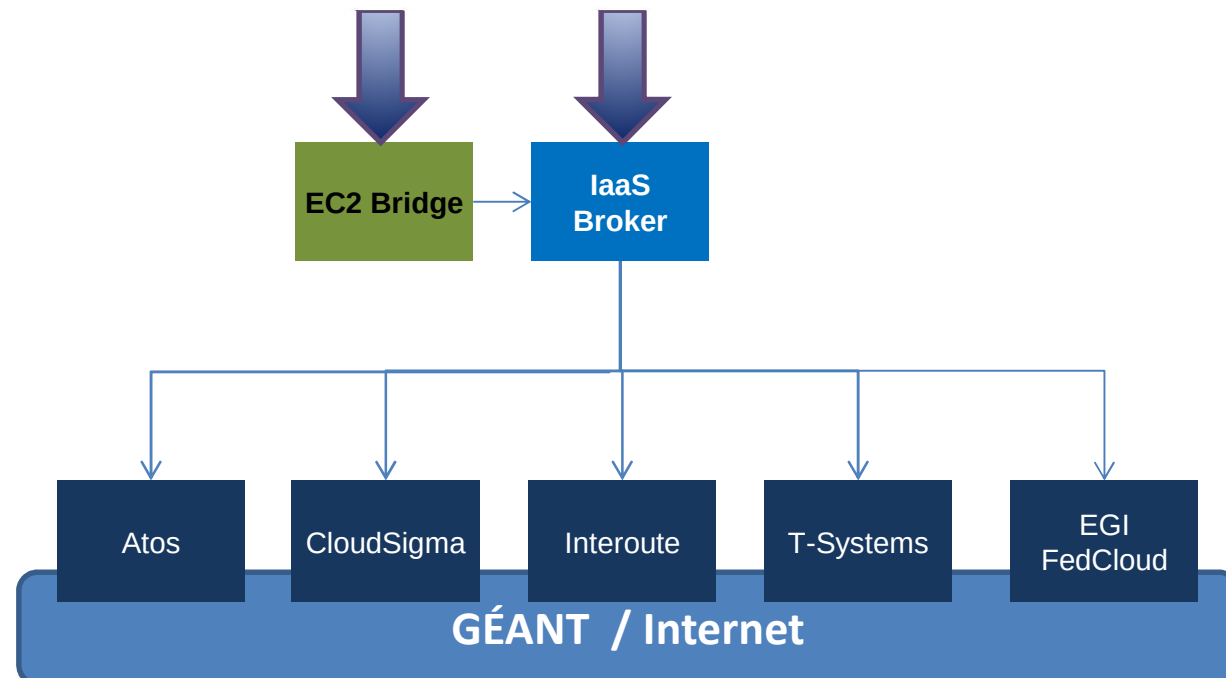


Branded-product # 1: HNX



Welcome to the Helix Nebula Marketplace

The Helix Nebula Initiative has been established by public and private organisations to build a multidisciplinary cloud platform for data intensive science.

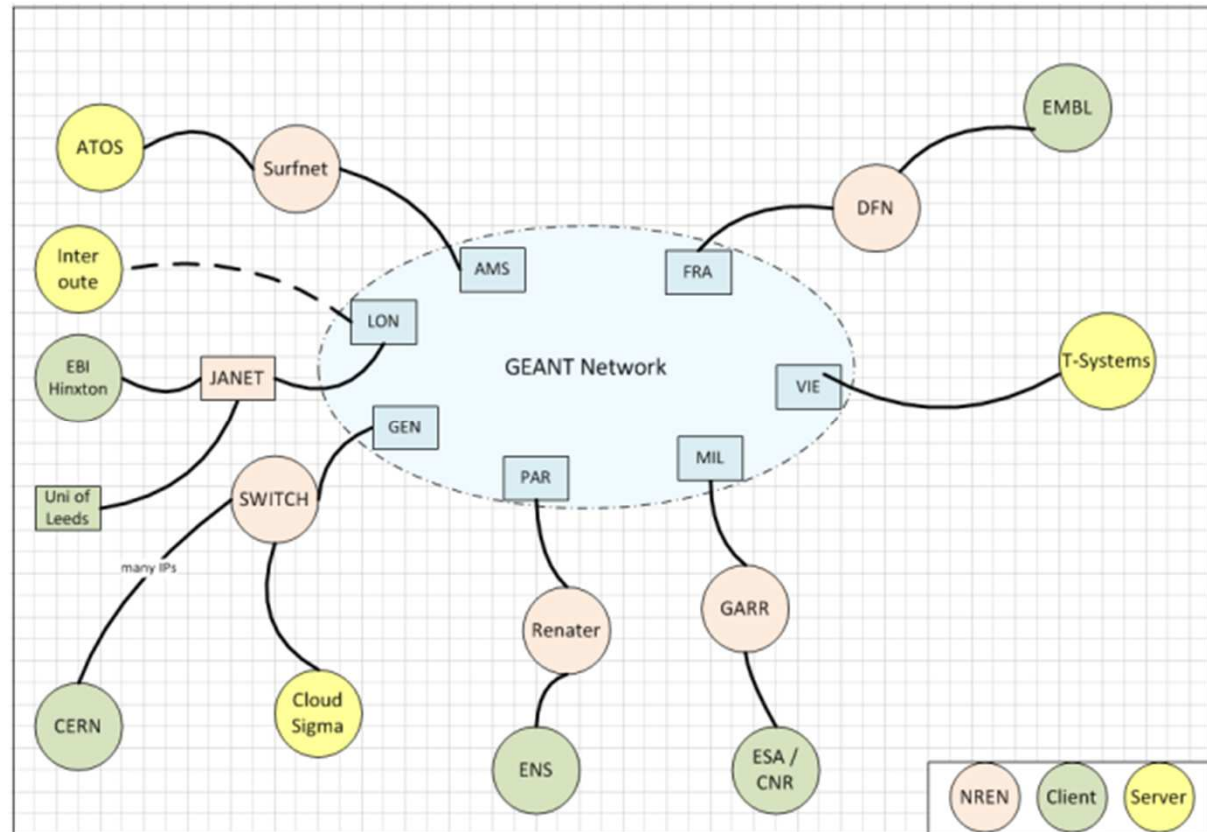


Building the hybrid cloud

Connecting commercial cloud providers to GÉANT/NRENs

GEANT Association offering free IP connectivity in GÉANT for research traffic during the pilot phase

NRENs have different commercial agreements (usually they apply a fee)



RedIRIS is connecting Atos in Tenerife

Example HNX procurement

CERN issued a price enquiry for a small fraction of the resources necessary to run the ATLAS experiment simulation software

- Several valid responses were received and Atos (data centre in Tenerife, Spain) was awarded the contract accessed via the SlipStream broker and provisioning engine (provided by SixSq SME) and delivered over the GEANT network
- >1 million CPU hours were successfully provided over a 1 month period with a 97% efficiency (measured in terms of Wallclock consumption) for CPU intensive (cloud storage used as a short-term cache) workload that confirmed
  - the use of commercial IaaS is technically feasible for simulation workloads
  - variations in RESTful APIs can be overcome with reasonable technical effort
  - independent monitoring of the quality of service and resources consumed is essential

Findings on Commercial cloud services

- Public organisations see value (rationalisation of IT services and reduction of costs) and opportunities (expand the impact of their work) in making use of commercial cloud services
- Testing of commercial cloud services by public research organisations is underway but currently has limited impact on production systems
- The public research sector and the exploitation of publicly funded research data is a potentially profitable market for commercial cloud service providers
- The procurement and use of commercial cloud services poses a number of legal, contractual and procedural questions
- A coordinated approach by public organisations will help structure the market and reduce the burden on individual organisations



Procurement Innovation for Cloud Services in Europe

TOP FIVE OBJECTIVES



Who is PICSE for?



Public Research
Organisations



Libraries

Stay Tuned!



www.picse.eu
info@picse.eu

- H2020 Coordination and Support Action
- Total budget: 500 K€
- Duration: 18 months
- Start date: 1 Oct. 2014
- Partners:



CSA cloud
security
alliance®



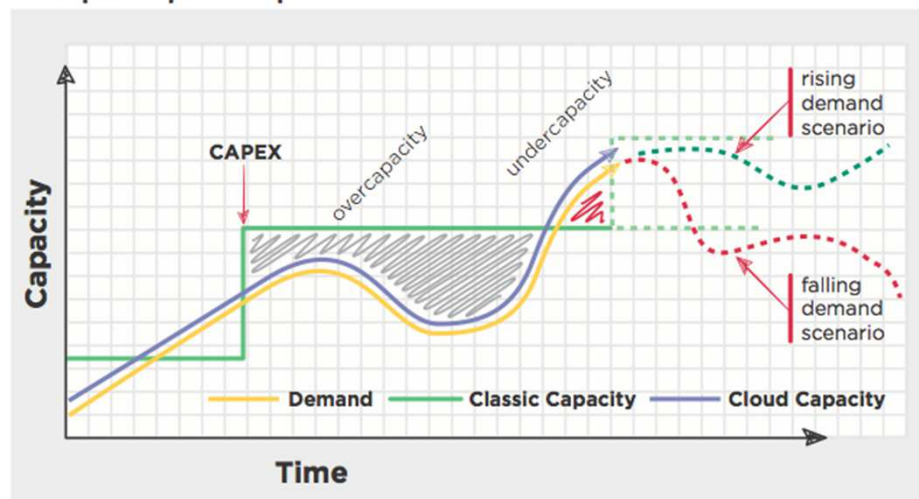
Trust-IT Services Ltd
Communicating ICT to markets

The funding model

Funding agencies supporting research communities make investments in e-infrastructure that must be justified

- Moving to a *pay-per-use* model represents a significant change
- Funders support specific user groups but others may also want to access the resources

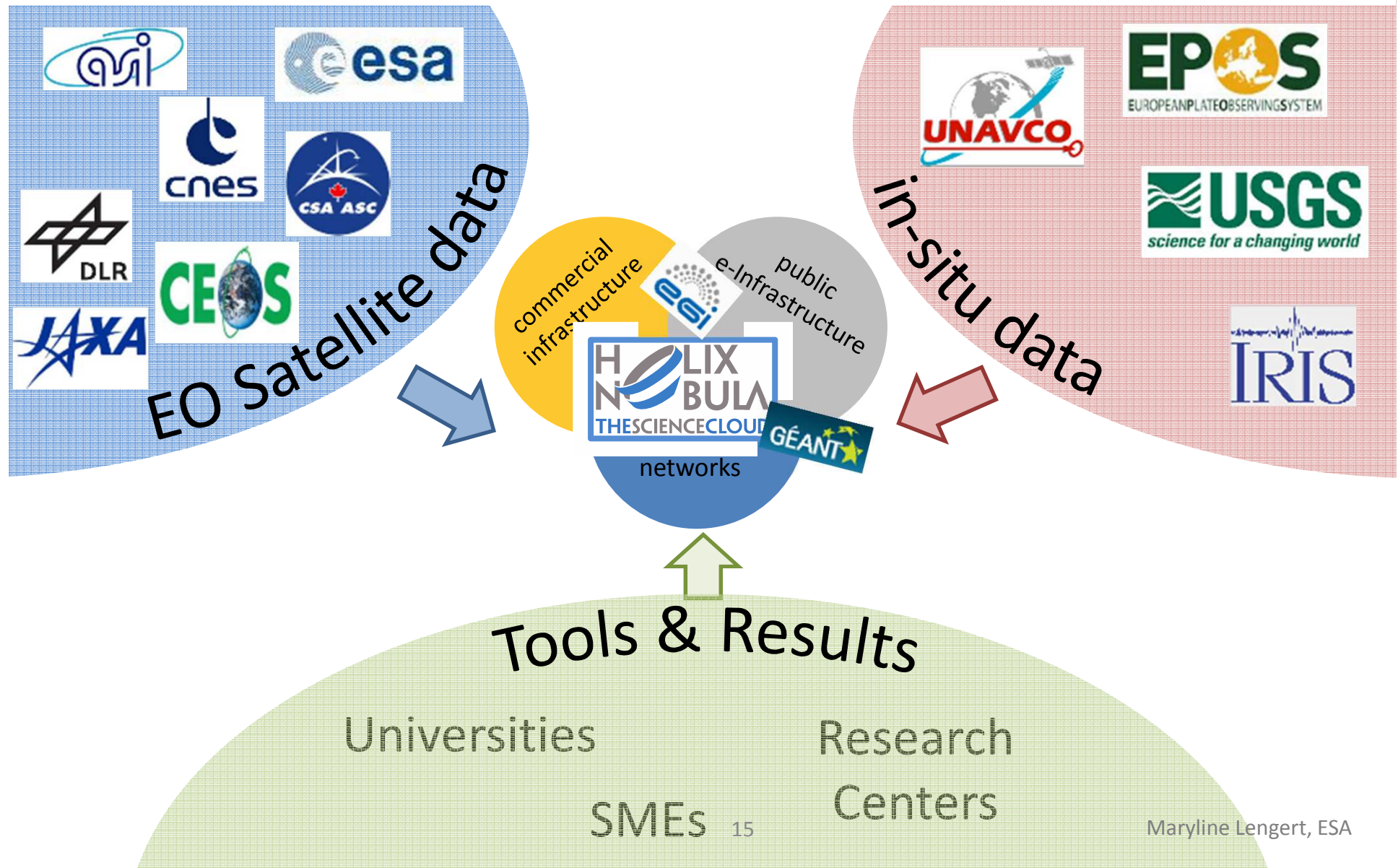
Example: Capex to Opex



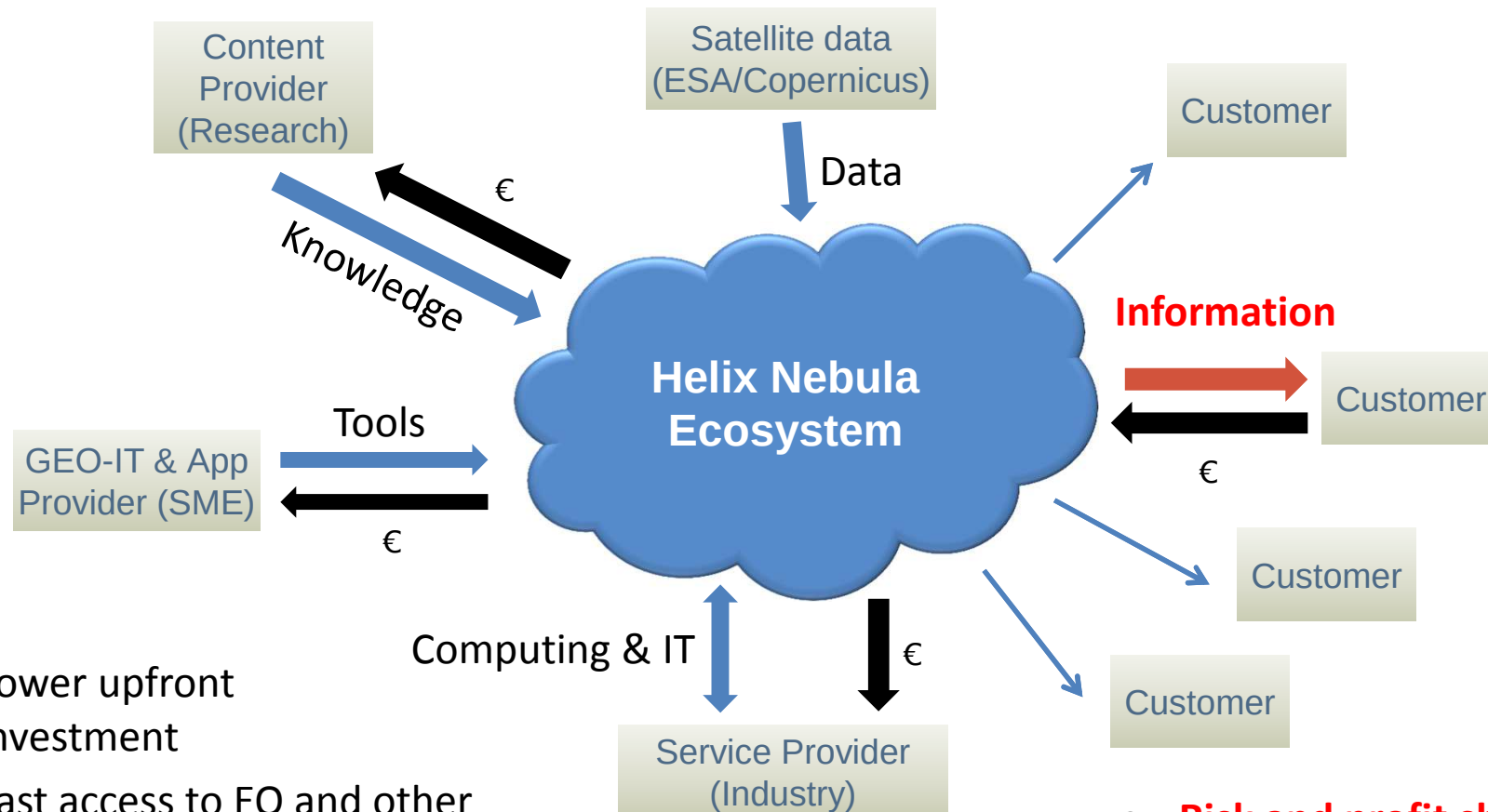
Source: Cloud: Economics, Sept 2011 by Rackspace Knowledge Center

Factors such as the pattern of demand, TCO and transformation costs need to be factored into any financial analysis of a potential cloud computing solution

Supersites Exploitation Platform: Infrastructure concept for science use



Information as a Service: Science interfacing with private sector



- Lower upfront investment
- Fast access to EO and other geodata resources
- Disruptive technology

- **Risk and profit sharing**
- **Sustainability**
- **Data Value Chain**

Vertical Market Big Data Heatmap



Based on mean scores assigned by survey respondents

Proposed Joint Pre-Commercial Procurement

The group of buyers have committed

- >1.6M€ of funds
- Manpower
- Applications & Data
- In-house IT resources

To procure innovative cloud services

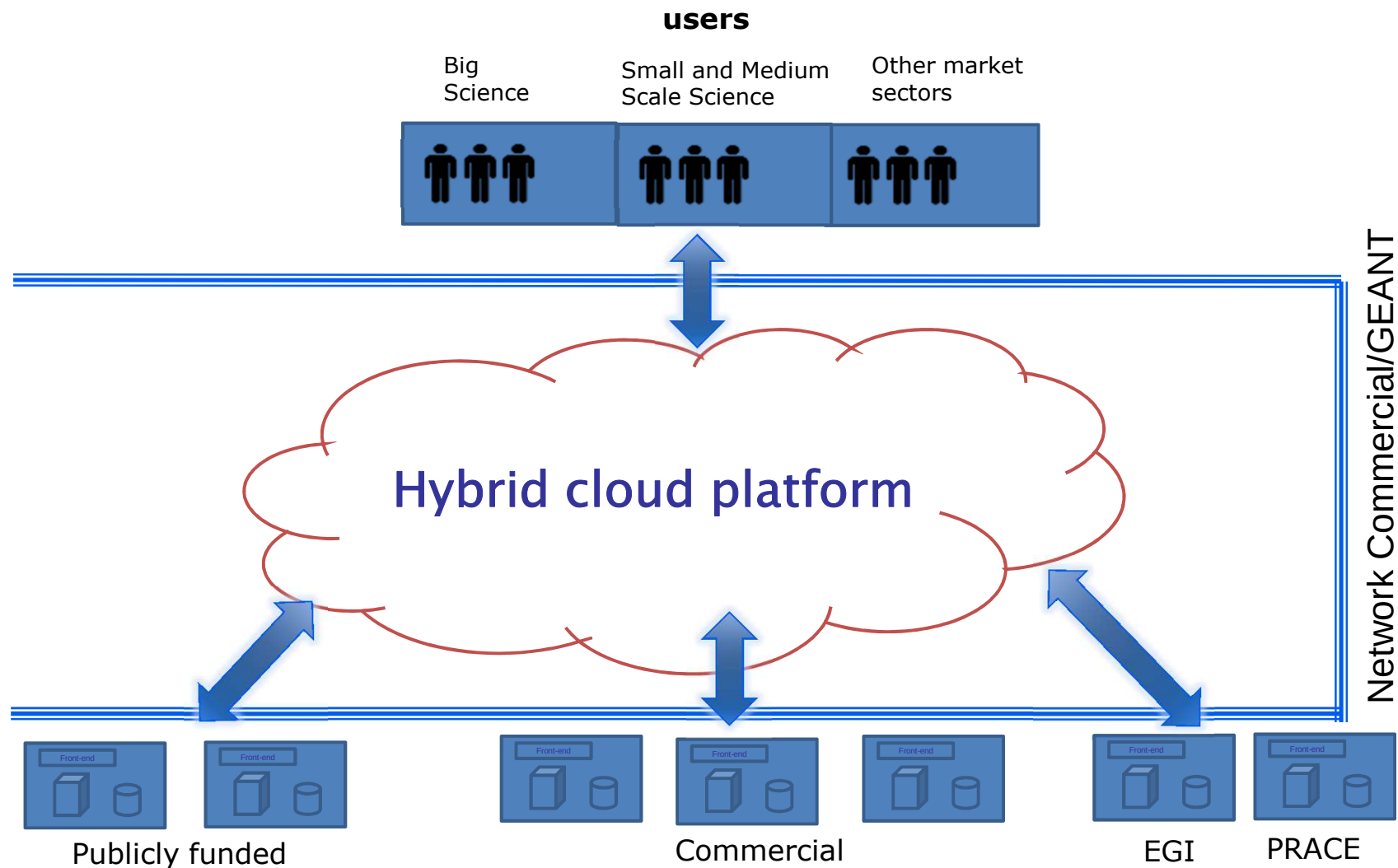
Integrated into a hybrid cloud model:

- Commercial cloud services
- e-Infrastructures
  - GEANT network
  - eduGAIN Fed. Id mgmt.
  - EGI Fed Cloud
- Potentially host data services from EUDAT, INDIGO-Datacloud, etc.
- In-house IT resources

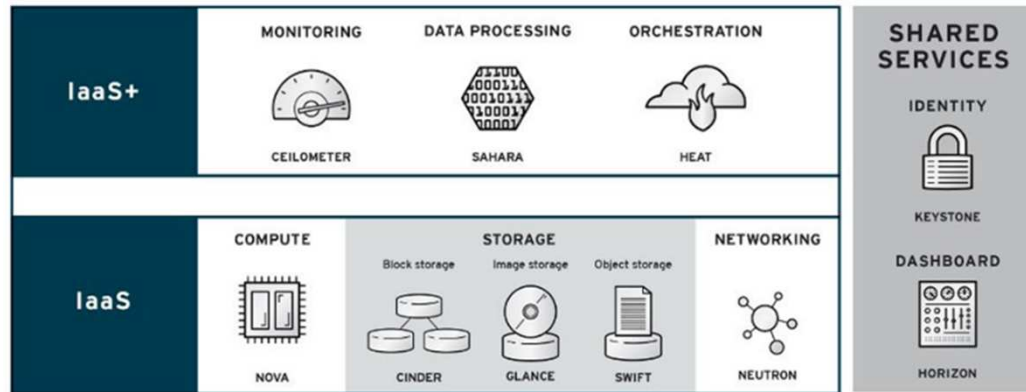
Made available to end-users including BBMI, DARIAH, ELIXIR, EISCAT_3D, EPOS, INSTRUCT, LifeWatch, LHC, etc.



Helix Nebula Hybrid Cloud Model



Open standards



OpenStack software controls large pools of compute, storage, and networking resources throughout a datacenter, managed through a dashboard or via the OpenStack API.



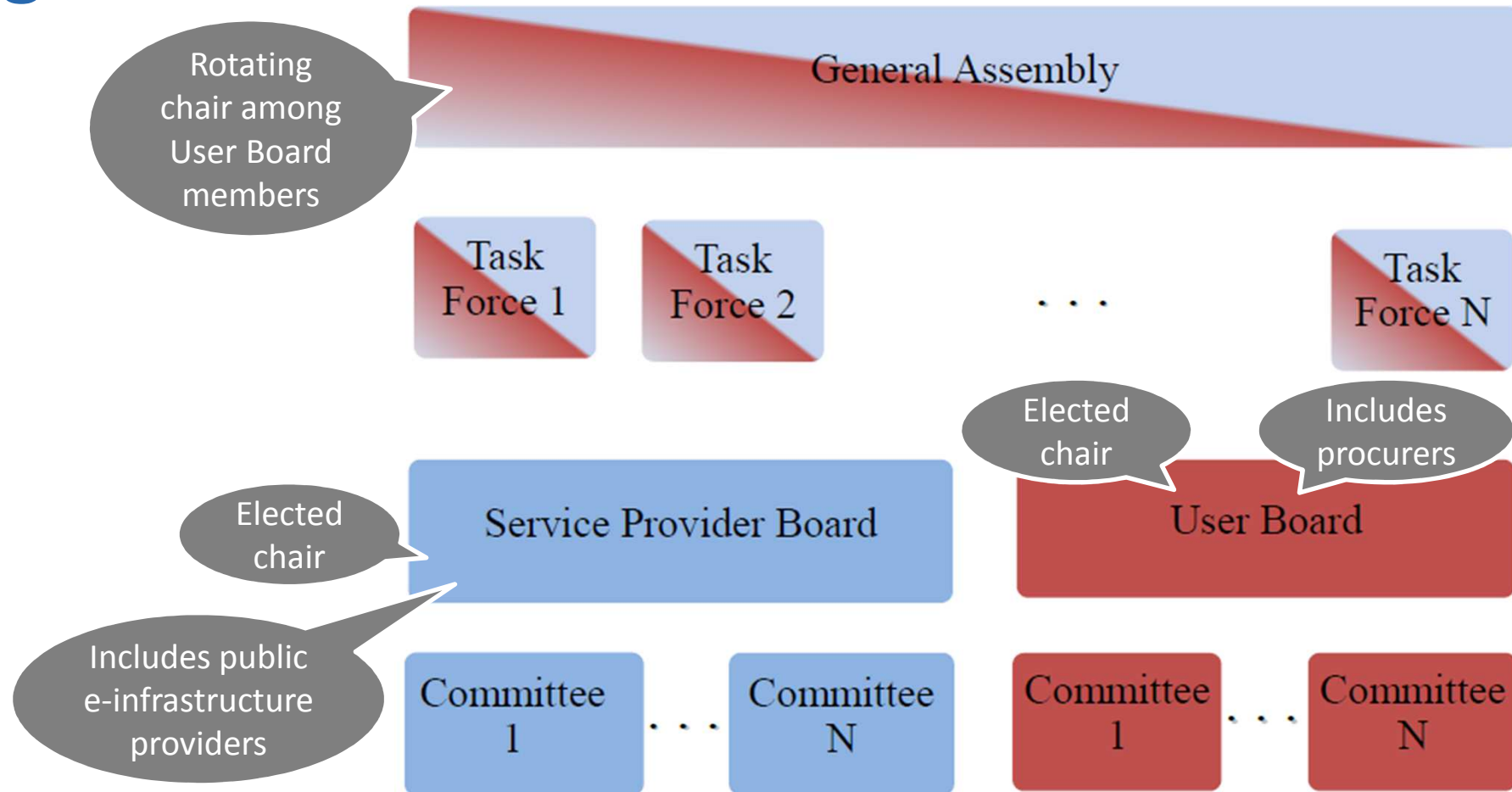
OpenStack based Public cloud services providers

Bob Jones, CERN



The OpenStack community

The Helix Nebula Initiative governance model



Next general assembly
24-26 June 2015
CERN, Geneva

General Assembly Chair organisation provides the secretariat resources
All members sign the HNI Membership Participation Agreement
No membership fees

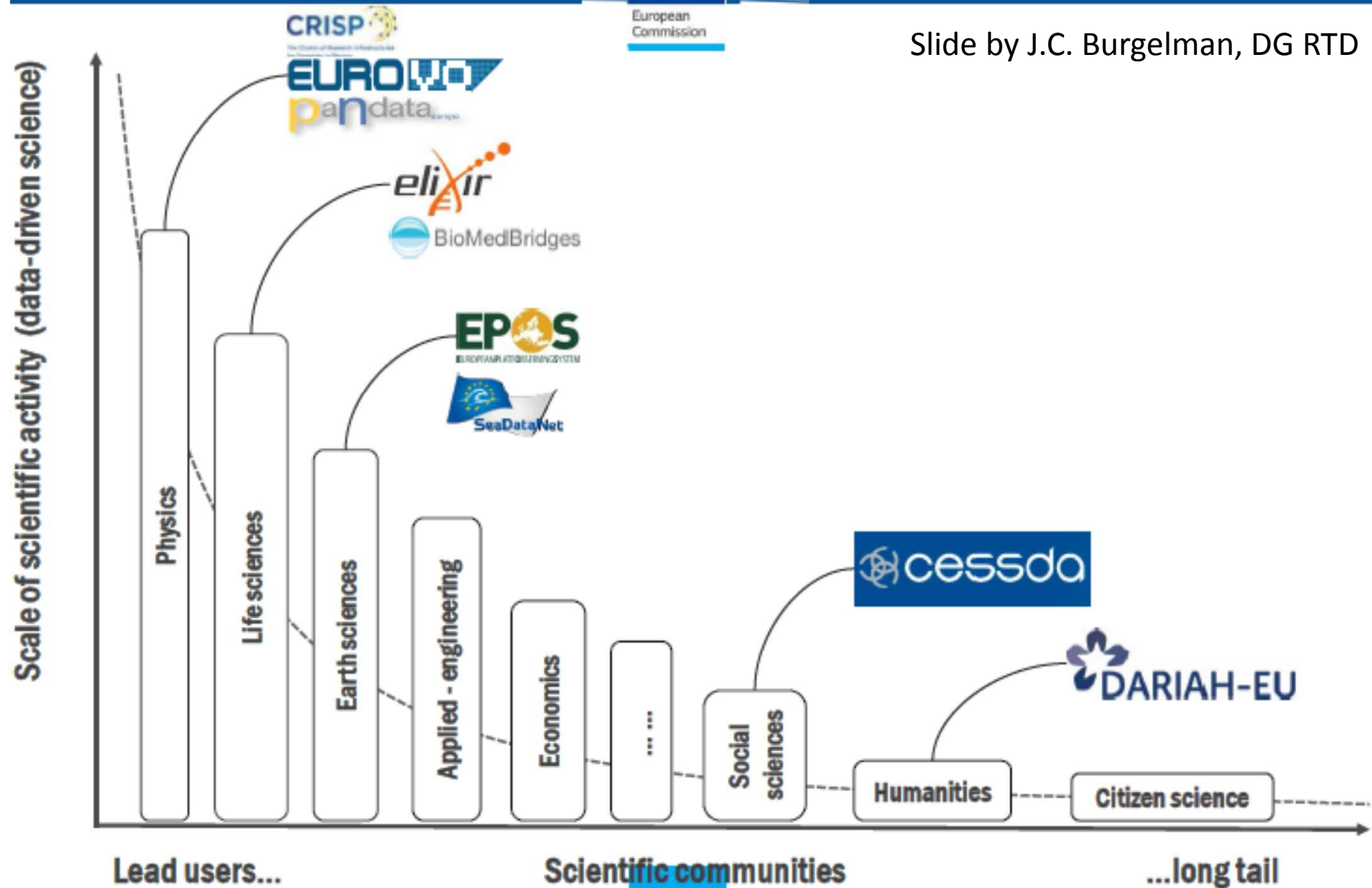
How to reach full-scale

- Grow from 5% to 100%
  - Engage more research infrastructures/communities & cloud service providers
  - Increase commodity service capacity at the IaaS level
  - Offer higher-level features (Platform as a Service, Software as a Service, etc.)
  - Expand coverage to broader public sector (eGovernment) and private sector (industries that use research data – energy, pharmaceutical, insurance etc.)
  - Interact with other regions (Africa, Asia, Latin America, North America)
- Leverage the 2016-2017 work-programme
  - Work with DG CNECT to engage e-infrastructures and cloud providers
  - Work with DG RTD to engage Research Infrastructures
  - Develop s/w technologies that can scale to the Exabyte range at the infrastructure, data and application layers
  - Continue investment in the public e-infrastructures during the same period

European Research Cloud



Slide by J.C. Burgelman, DG RTD

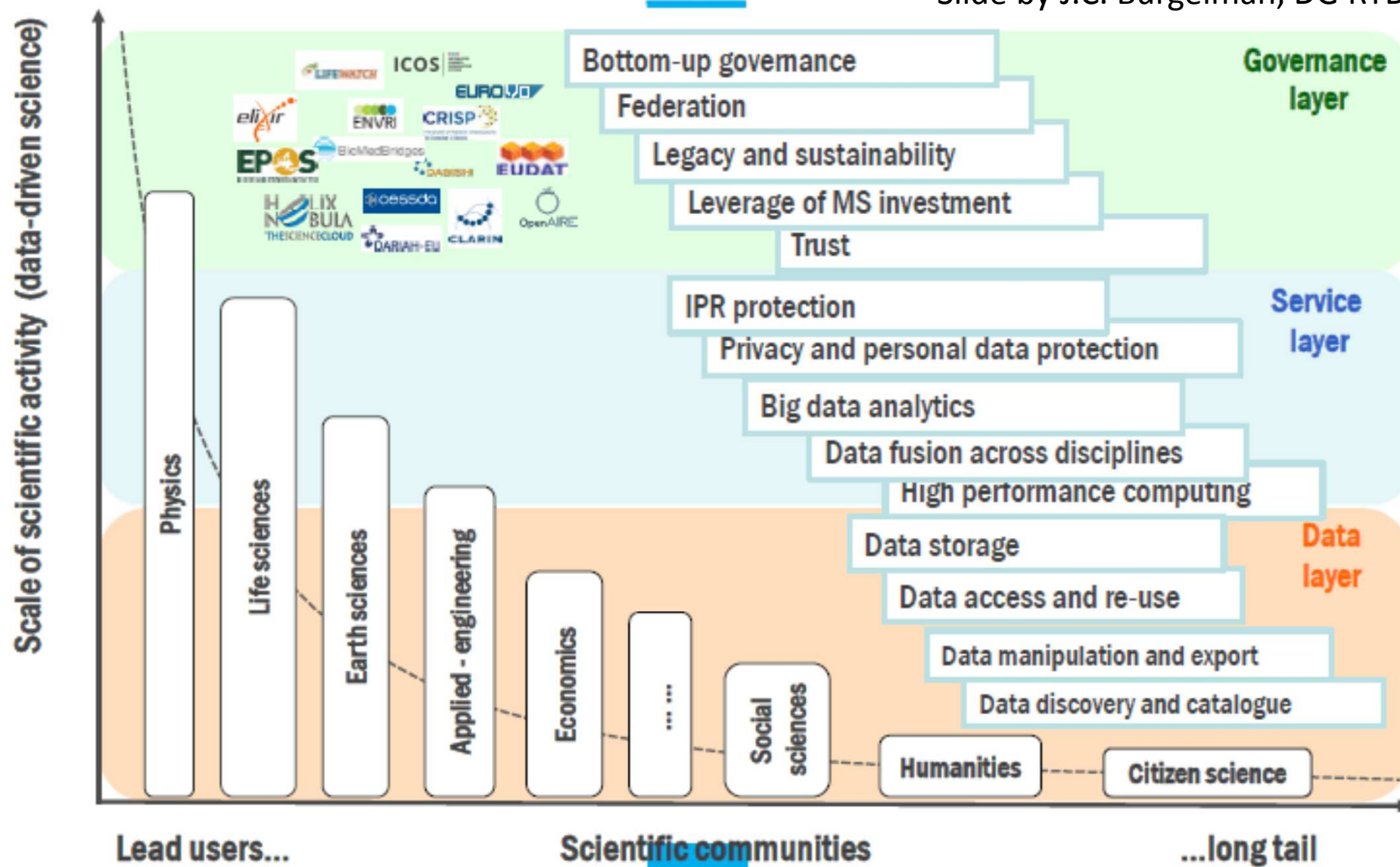


European Research Cloud



European
Commission

Slide by J.C. Burgelman, DG RTD



Closing remarks



- Cloud computing offers a cost effective model for IT resource provisioning
- A hybrid cloud model leverages the investments made in both the public and private sectors while ensuring trust and continuity
- Changes to the procurement process in the public research sector are necessary to benefit from a dynamic Digital Single Market
- Mature open source technologies exist but integration, policy and governance requires careful attention
- A European Open Science Cloud will promote public-private innovation to satisfy the needs of the research communities and increase the global competitiveness of Europe's cloud service suppliers

