

EGI services for distribution and federation of data and computing

Tiziana Ferrari
Technical Director, EGI.eu
tiziana.ferrari@egi.eu

March 2014



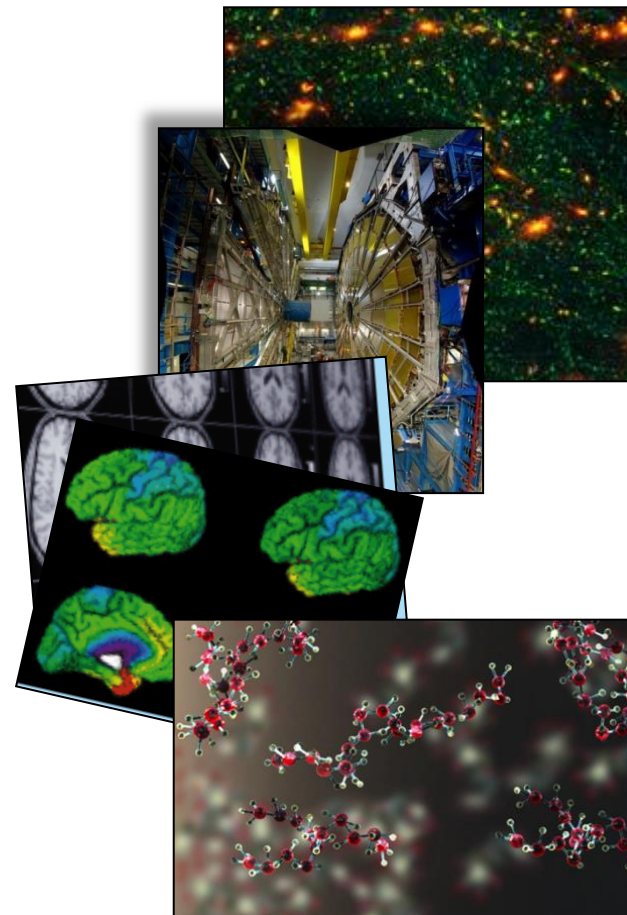
MISSION. To support researchers from all disciplines with the reliable and innovative ICT services they need to accelerate excellent science

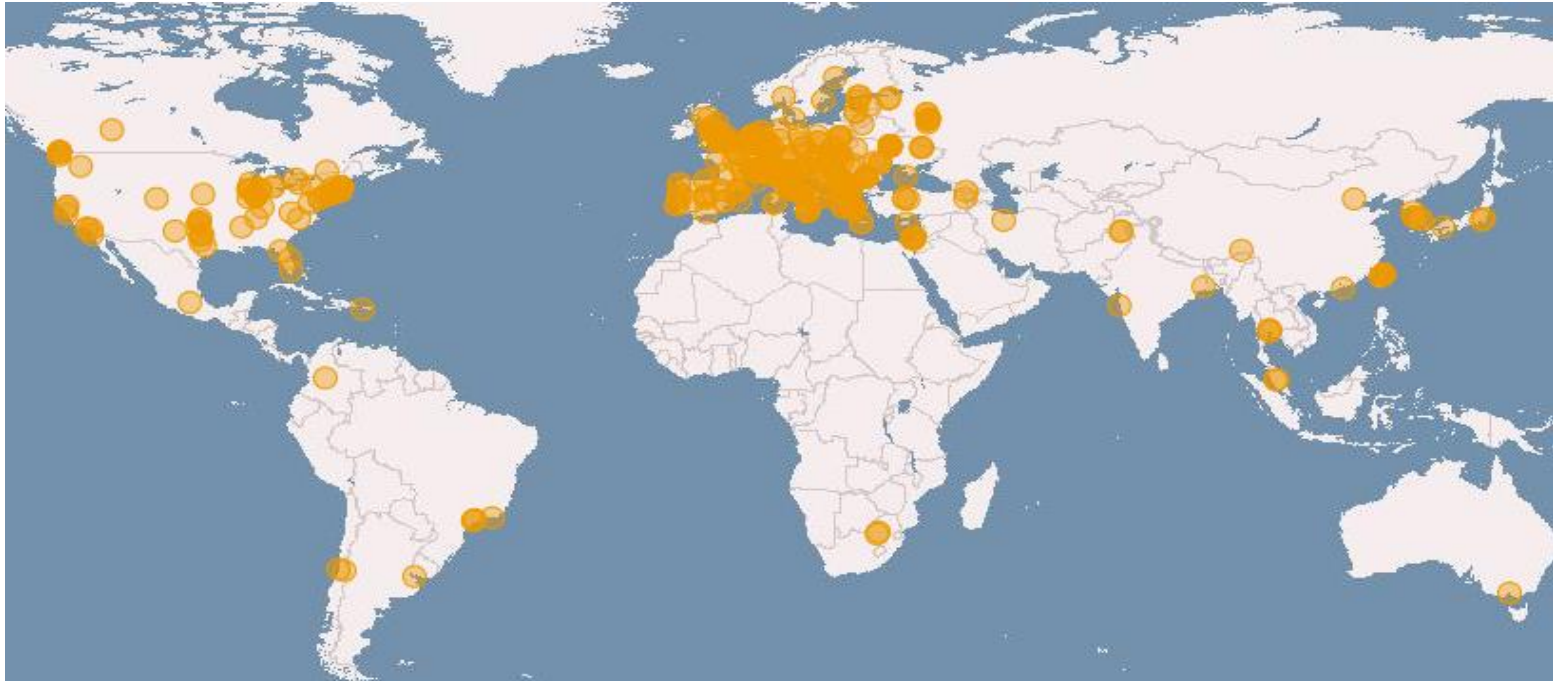
2.3 Million computing hours for earth science (2013-2014)

- Natural and physical sciences
- Medical and health sciences
- Engineering and technology

EC **EGI-InSPIRE** project (2010-2014)

<http://www.egi.eu/case-studies/>





- Distributed and federated data and computing facilities
- Grid and Cloud compute platform
- 340 data centres in 34 National Grid Initiatives/EIROs
- 435,000 logical CPU cores
- Supporting science for over 10 years
- > 200 research projects
- 190 PB disk, 180 PB tape
- 1.6 M jobs/day
- > 99.6% reliability

- Standards based access to heterogeneous distributed data and compute services for Grid and Cloud platforms
 - VALUE: uniform access, sharing, no lock-in, on-demand scale out capabilities,
- Production quality federation of services from
 - Publicly funded national e-Infrastructures
 - Institutional infrastructures
 - free at point of delivery, pay per use (in roadmap)
 - Commercial providers (Helix Nebula Market)
 - free at point of delivery + pay per use



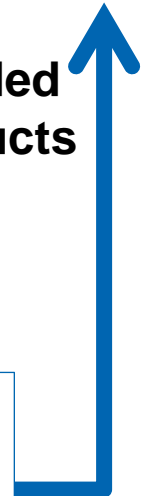
In 2-3 years Sentinels
8 TB data per day

Data Flow



Value added
products

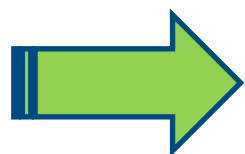
- Public Authorities
- policy and decision makers
- Downstream Industry
- Research and education



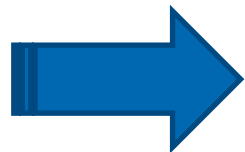
EGI assets for Copernicus data exploitation
High-Throughput data analysis facility
Federation of cloud infrastructures
Data management services, open data repository (in
roadmap)
Consultancy



- European federation of community clouds
 - Cloud Compute: deploy/manage virtual machines (VM)
 - Cloud Storage: store/access/retrieve digital objects
 - Image Repository: store/retrieve certified VM images
- Integrate heterogeneous cloud technologies
 - OpenStack, OpenNebula, CloudStack,...
- Integrate with commercial providers
 - Within EGI or through the Helix Nebula Marketplace



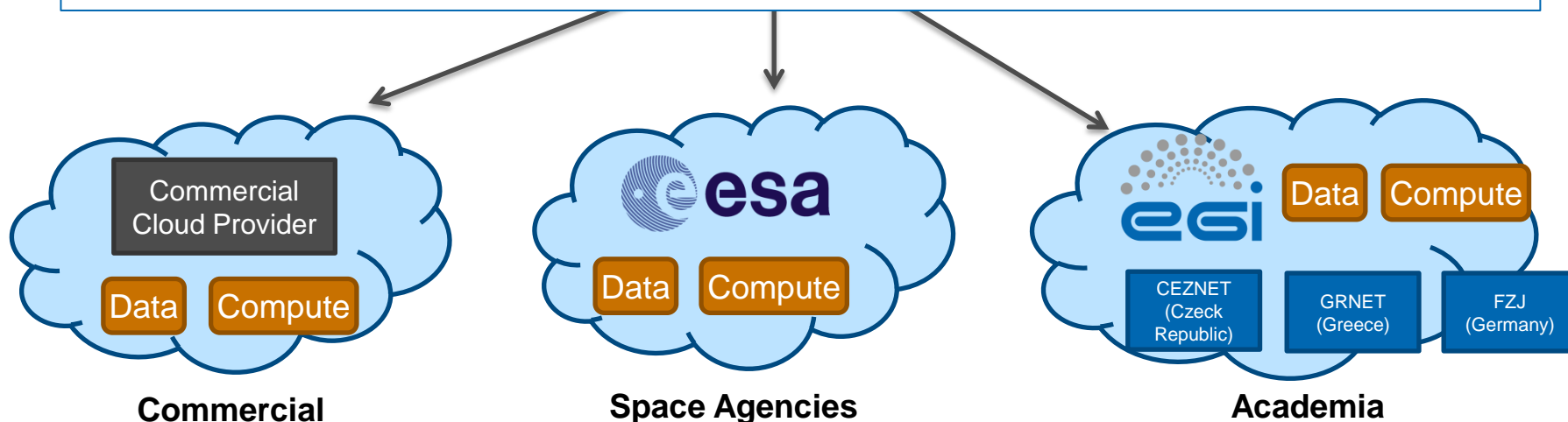
2014: 13000 cores, 5.7 PB



- Easy deployment of own/customised services
- Efficiency by co-locating big data + cloud computing

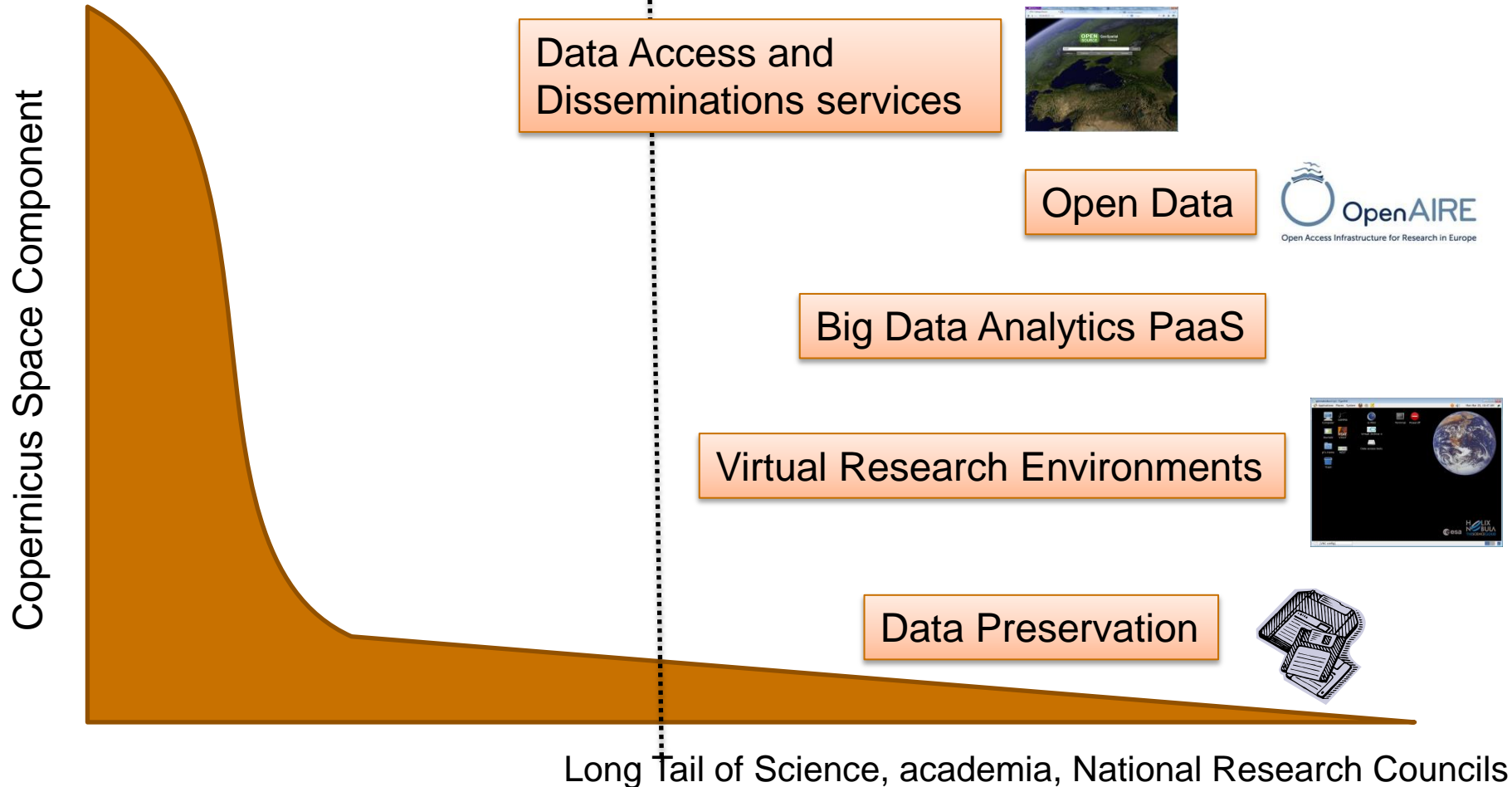
Services for an European federation of Cloud and Data resources, for Cloud Compute, Cloud Storage, Image Repository

- Based on **Open Standards** and common **Core Services** to ensure interoperability
 - Common **AAI, Accounting, Monitoring, Image and service Marketplace...**
 - **OCCI, CDMI, OVF, ...**
- Integrate agencies and commercial providers
 - Within EGI or through the **Helix Nebula Marketplace**



Core products

Value Added products



- **Data Access and Dissemination PaaS service**

- Provide community with a ready-to-use service for collecting, storing, cataloguing, searching and disseminating EO data.
- Based on the ENVRI tools and the FedCloud Storage and Infrastructure services.
- PoC on-going with the ENVRI/EISCAT-3D community.



- **Big Data Analytics**

- High-Throughput-Computing PaaS.
- PoC completed in September with ESA G-POD service.
- Future extension (in roadmap) to support Hadoop and Rasdaman PaaS



- **Data Preservation & Open Data Services (in roadmap)**

- Data preservation services, in collaboration with APARSEN, for scientists and SMEs, to store, preserve and keep accessible scientific data after the end of the research project
- Repositories for Open Data, linking scientific publications with the input data and high level products, federated in OpenAIRE

