

IBERGRID



INFRASTRUCTURE STATUS

Jorge Gomes

on behalf of the IBERGRID collaboration



Distributed computing infrastructure

- a) Federates infrastructures from Iberian research and academic organizations (PT + ES) mainly focused:
 - Cloud Computing
 - Grid Computing
 - Data Processing
- b) Enables a joint participation in European initiatives including EGI and EOSC supporting research communities
 - Provides the regional operations coordination for the computing and data processing activities of several user communities including WLCG, ESFRIs and others
- c) Forum for common activities and sharing of knowledge.
 - Participation in EU and cross-border projects including both R&D and infrastructure oriented projects
 - Conferences and events (IBERGRID conference series)

Structure



REPÚBLICA
PORTUGUESA

Ministério da
Ciência, Tecnologia
e Ensino Superior



CSIC

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

FCT

Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA



IFCA

Instituto de Física de Cantabria

Tech coordination



Tech coordination

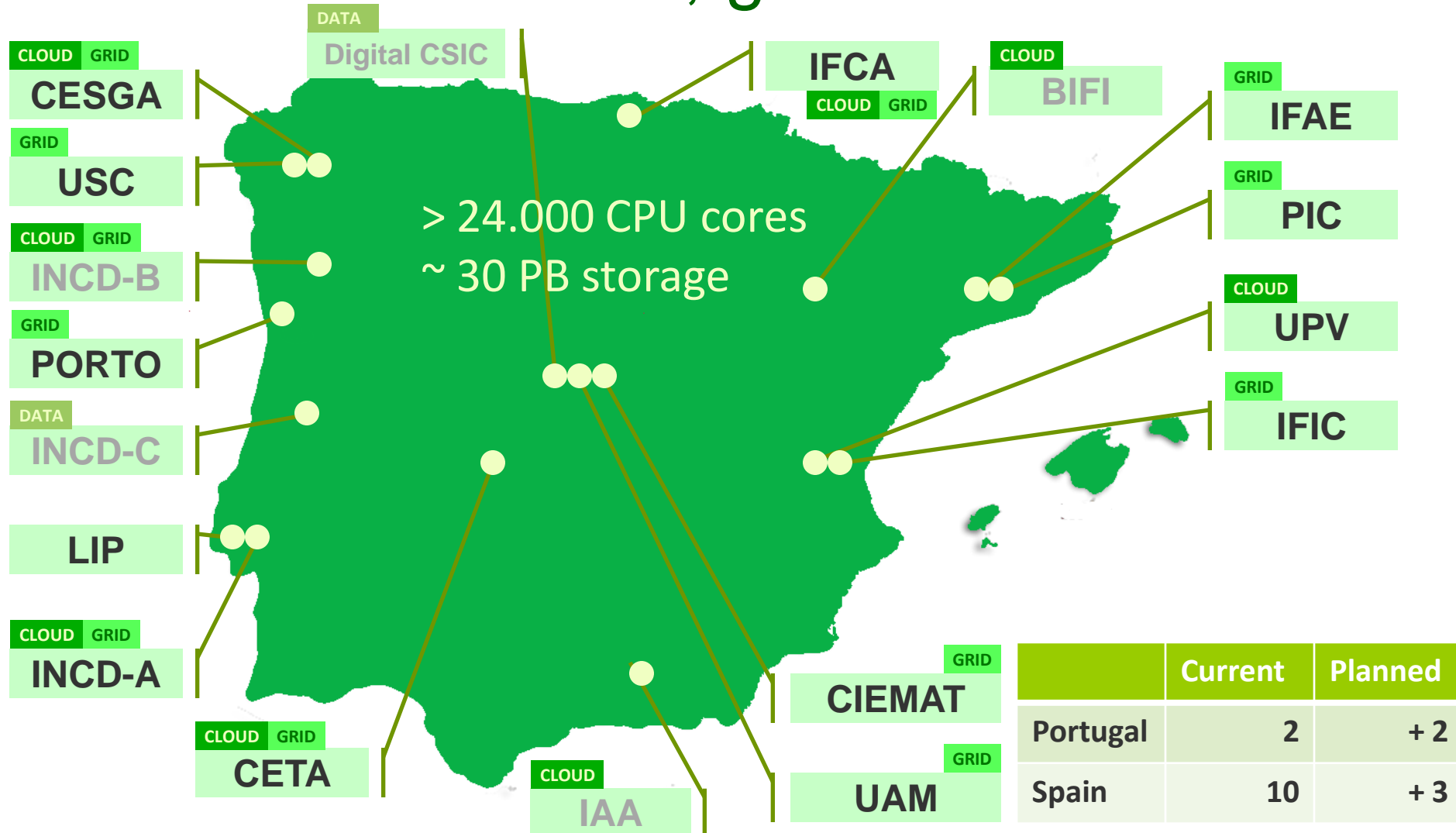
IBERGRID



Iberian summit of Valladolid in Nov 2018

- ❑ Support the IBERGRID conference organized annually in each of the two countries.
- ❑ Reinforce the collaboration between both countries to support the IBERGRID participation in international initiatives and infrastructures of distributed computing and digital repositories among which EOSC and EGI.
- ❑ Support the development and integration of thematic services of interest to the scientific community to be made available through the IBERGRID infrastructure.
- ❑ Support and promote the use of the IBERGRID infrastructure to support the participation in common scientific projects of strategic interest ,such as the participation at CERN, in ESFRIs and in the AIR Center.

Joins PT + ES cloud, grid and data:



Core services

- Authentication
Authorization
and virtual
organization
membership

INCD

TOPBDII

LFC

VOMS

LIP

web

wiki

LFC

VOMS

Accounting

TOPBDII

CESGA

ARGUS

LFC

VOMS

TOPBDII

IFCA

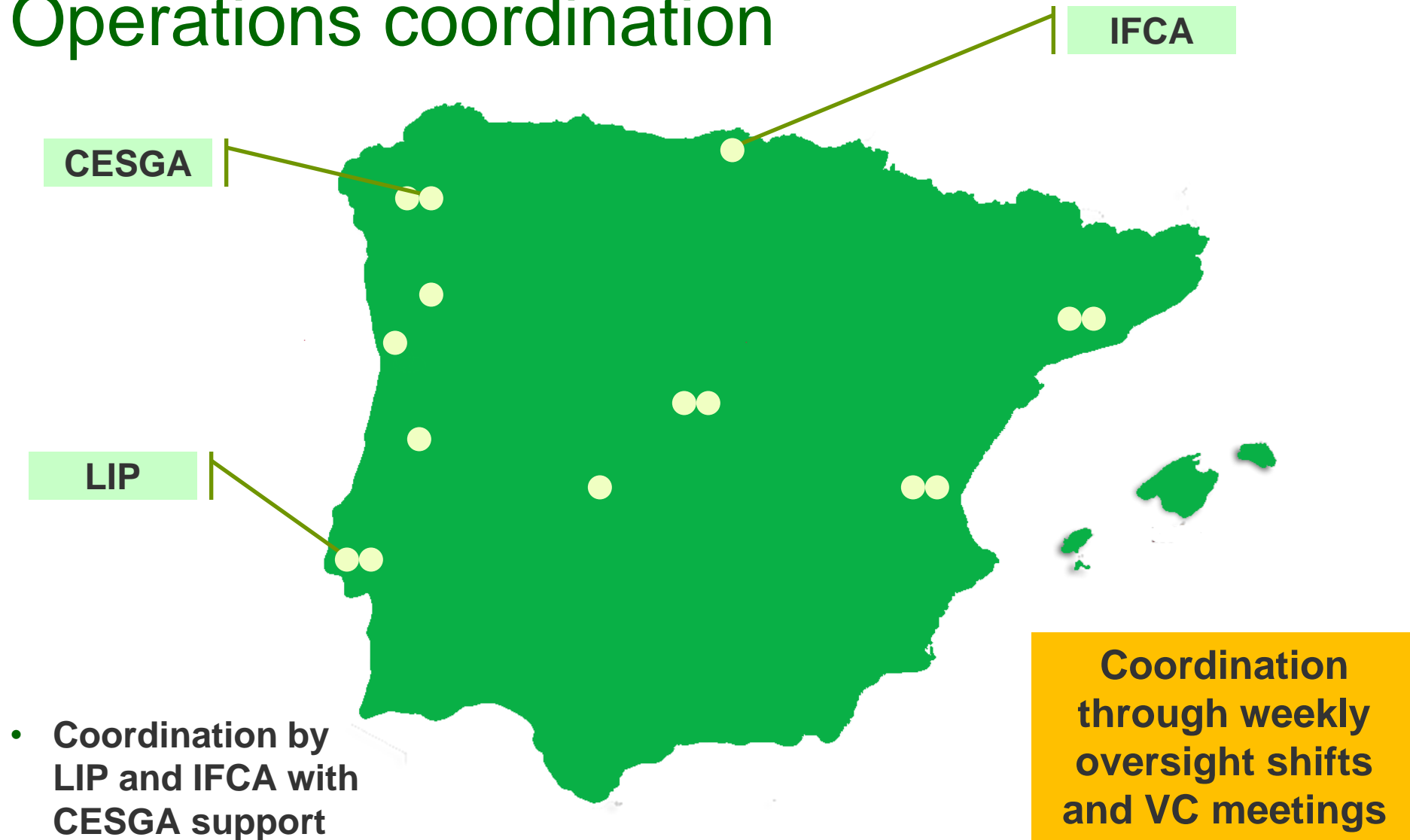
TOPBDII

CIEMAT

VOMS

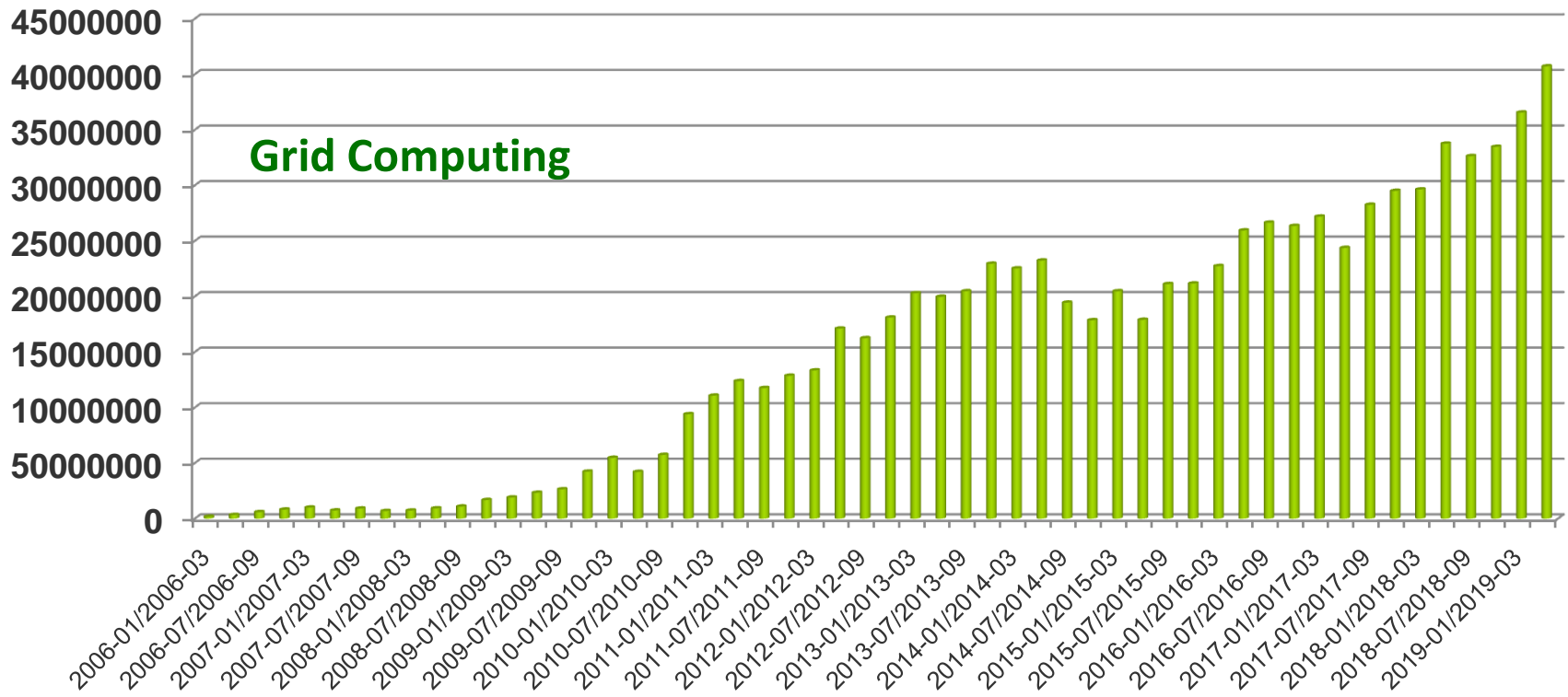
Replication for
critical infrastructure
services

Operations coordination



272 million jobs executed since 2006
1040 million processing hours

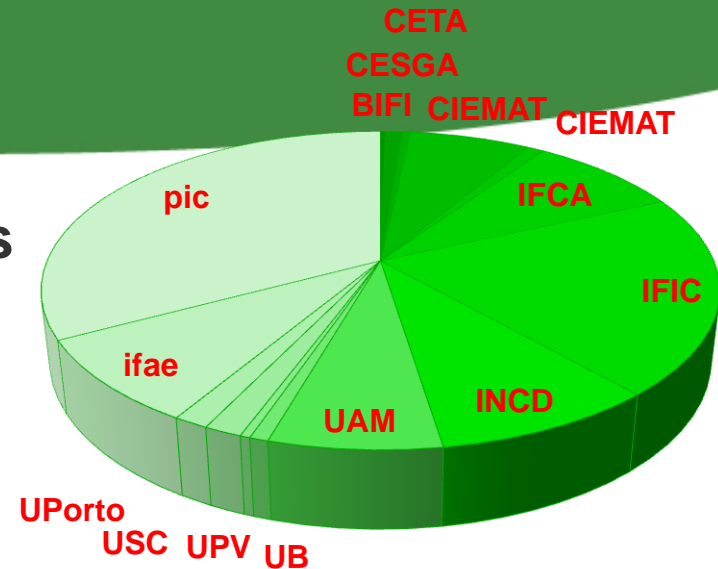
NGI_IBERGRID normalized CPU time (HS06)



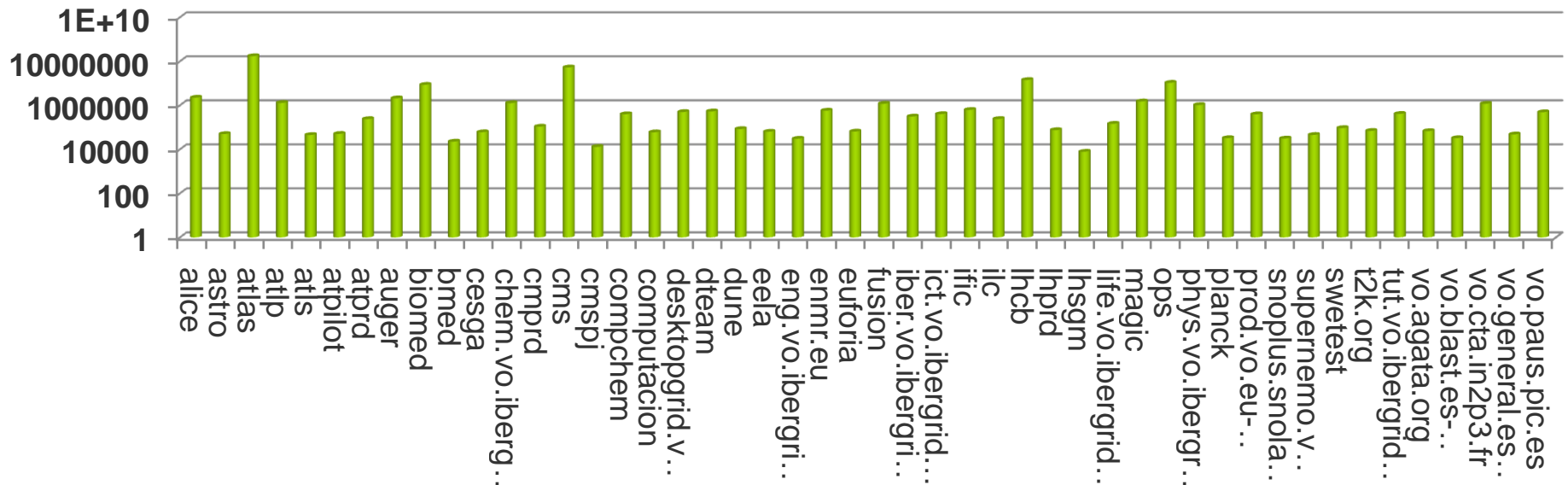
Grid Jobs

- By site and VO
- From 2006 to 2019
- Removing small usage

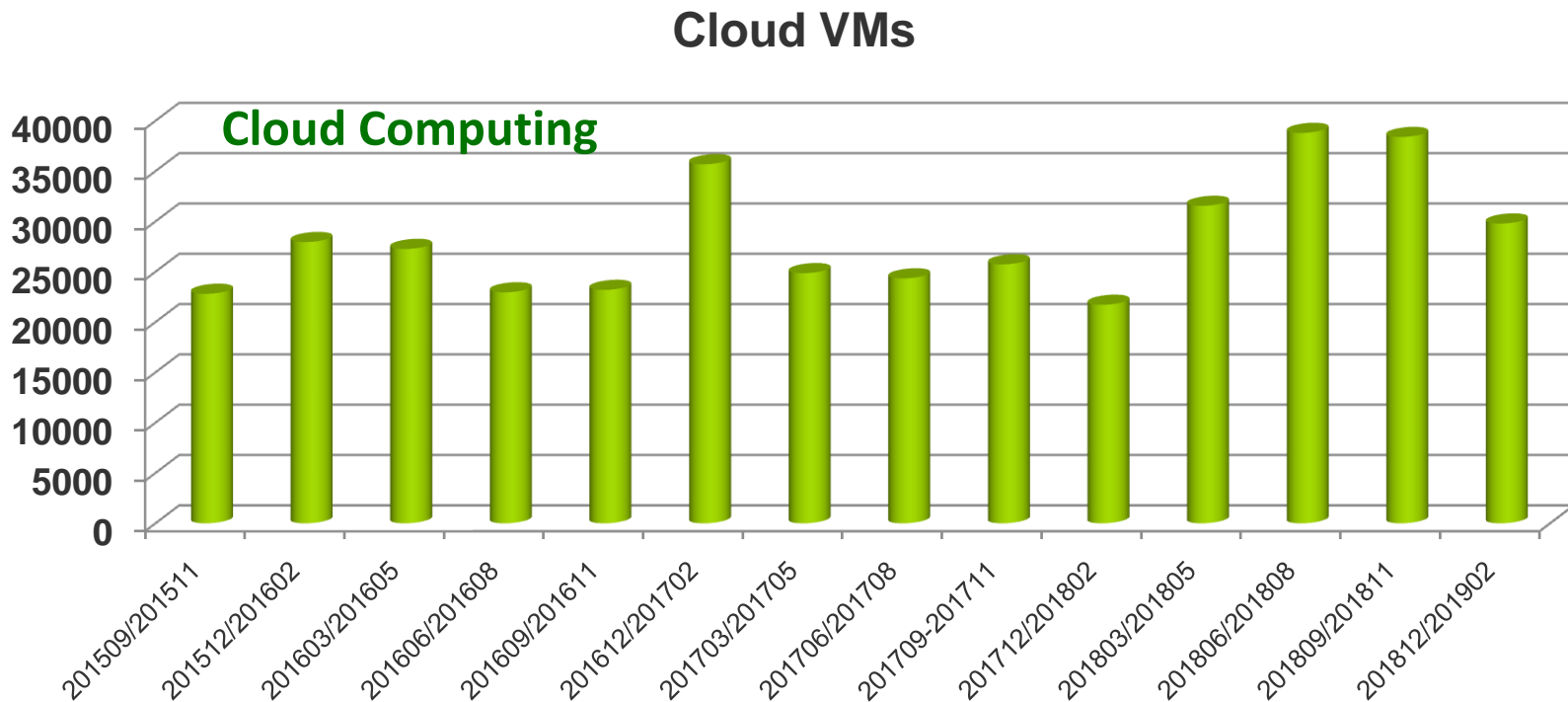
Jobs



Grid Jobs per VO



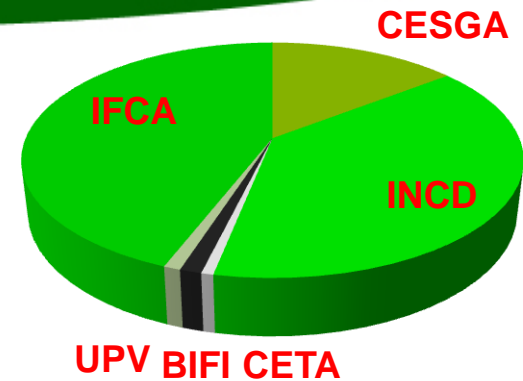
Instantiation of cloud VMs 395 thousand since September 2015



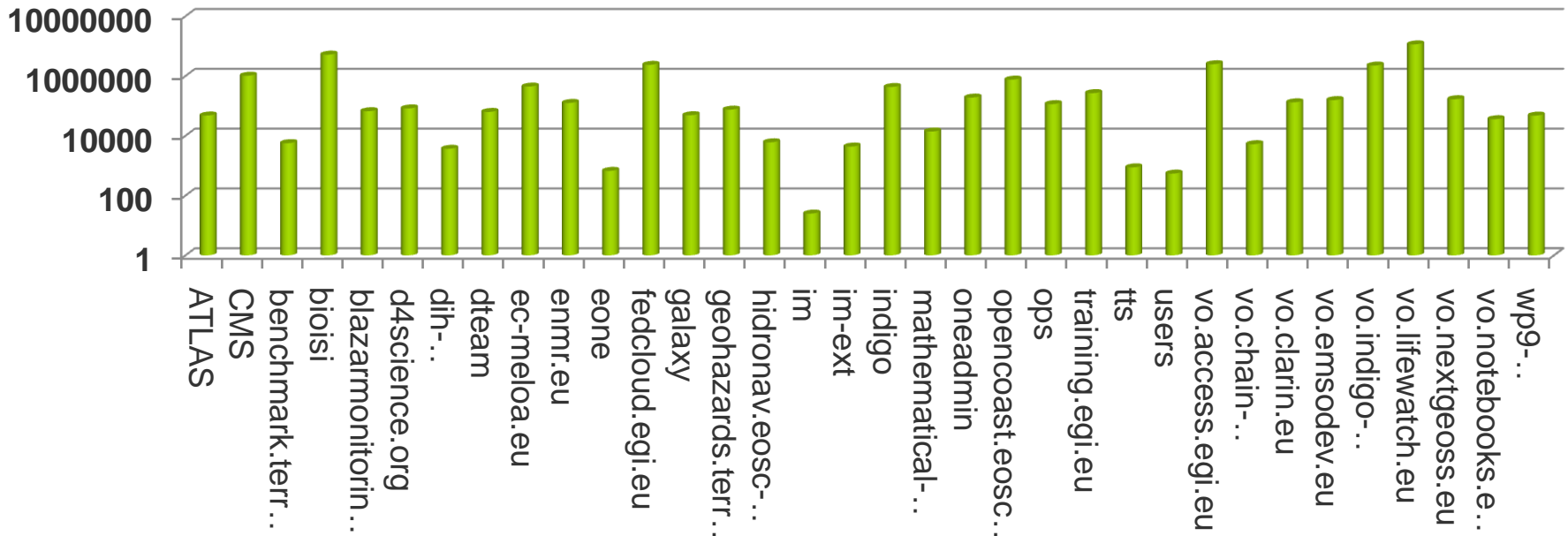
Cloud Hours

- By site and VO
- From 2015 to 2019
- > 25 million hours

Cloud
Hours

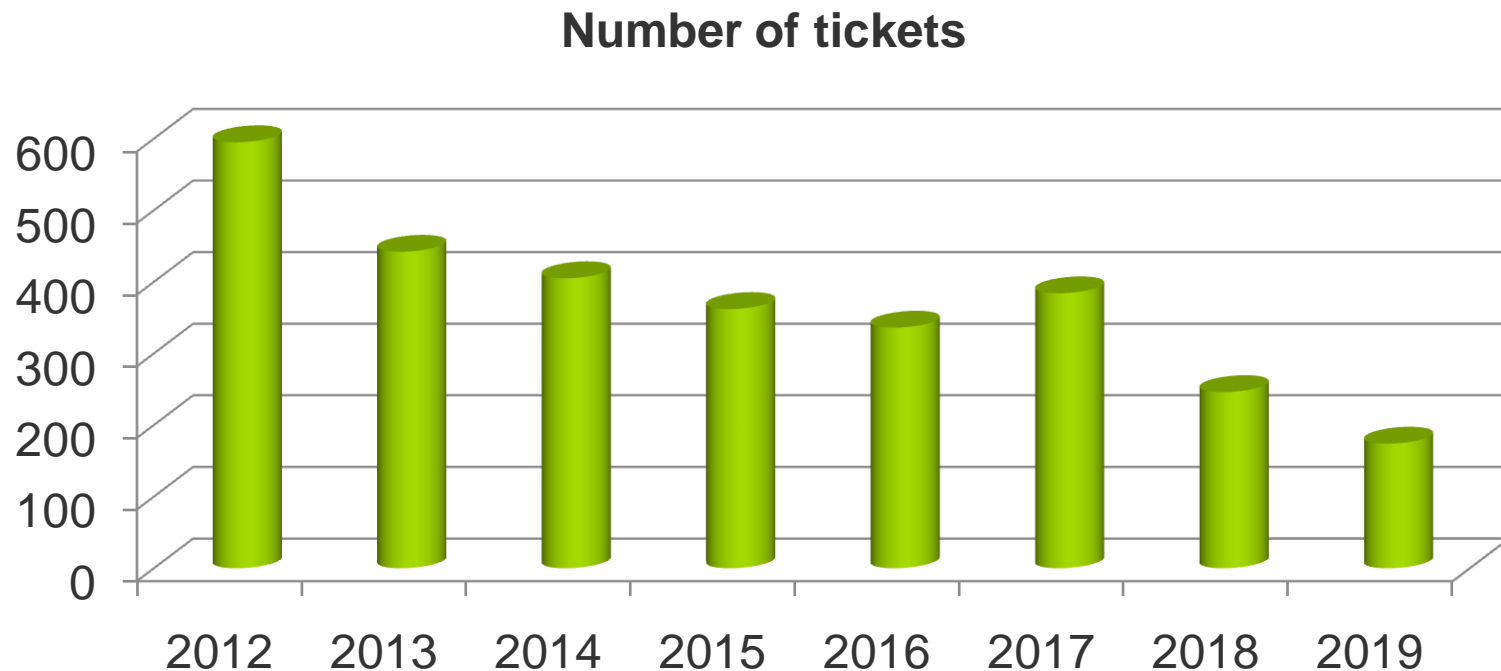


Cloud Hours per VO

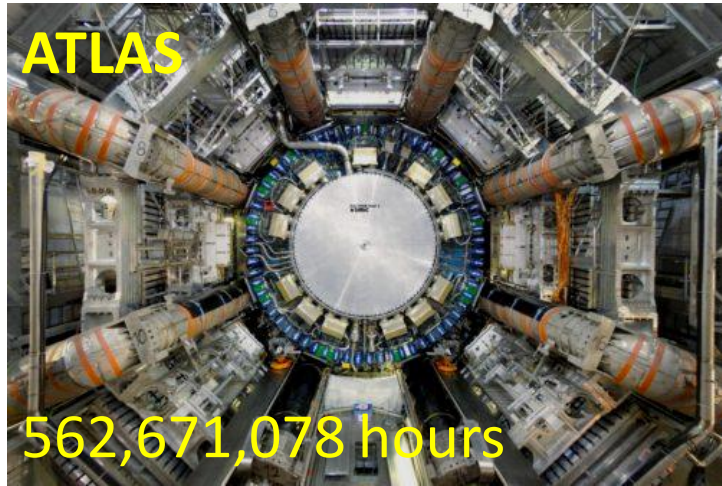


Infrastructure support

- Approximately 250 in 2018
- Helpdesk: EGI GGUS

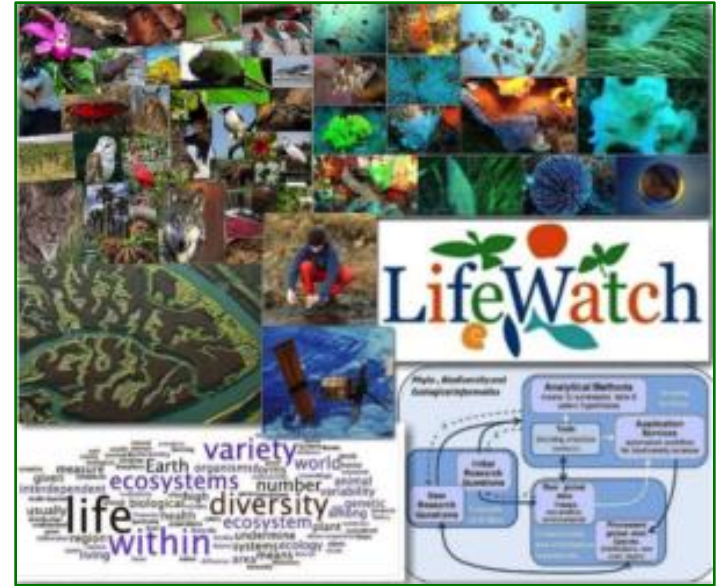


CERN LHC simulation and data analysis



LifeWatch ERIC

- LifeWatch ERIC, is a distributed Research e-Infrastructure to advance biodiversity research and to address the big environmental challenges and support knowledge-based strategic solutions to environmental preservation.
- The services currently available for the biodiversity community are also available for wider re-use by other scientific communities.
- In the Iberian area the activity of Lifewatch ERIC is articulated via **IberLife**, which in turn relies on the support of **Ibergrid** for the deployment and operation of core services.



16,733,691 hours on IBERGRID cloud resources

GBIF

- The Global Biodiversity Information Facility (GBIF) is an international network and research infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open access to data about all types of life on Earth.



- GBIF nodes in Portugal and Spain are maintained in the scope of **LifeWatch** and **IberLife** and these activities are supported by **IBERGRID**.
 - dados.gbif.pt
 - datos.gbif.es

Supported by IBERGRID Cloud resources

WeNMR and MoBrain

- WeNMR provides an e-Infrastructure platform and Science Gateway for **structural biology**. WeNMR serves all relevant **INSTRUCT** communities in line with the ESFRI roadmap.
- WeNMR is part of West-Life virtual research environment for structural biology.
- The current challenges in the post-genomic era call for virtual research platforms that provide the worldwide research community with both user-friendly tools, platforms for data analysis and exchange, and an underlying e-infrastructure.



AUGER

- On the vast Pampa Amarilla in western Argentina, the Pierre Auger Observatory is studying the highest-energy particles in the Universe, so-called cosmic rays.
- Cosmic rays with low to moderate energies are well understood, while those with extremely high energies remain highly mysterious. By detecting and studying these rare particles, the Pierre Auger Observatory is tackling the enigmas of their origin and existence.



12,714,524 hours in IBERGRID

OPENCoastS

- OPENCoastS – On-demand Operational Coastal Circulation Forecast Services
- Provides on-demand circulation forecast systems as-a-service for the European Atlantic coasts.
- OPENCoastS generates forecasts of water levels, 2D velocities and wave parameters over the spatial region of interest for periods of 72 hours, based on numerical simulations of all relevant physical processes.



Integrated into IBERGRID and EGI as an EOSC thematic service
Collaboration LIP, LNEC, INCD, UNICAN, CNRS, CSIC

Biomed

- large-scale international Virtual Organization health and life sciences. Currently, it is divided into 3 sectors:
 - a) medical imaging
 - b) bioinformatics
 - c) drug discovery.
- The VO is openly accessible to academics, and to private company for non-commercial purposes.



19,901,904 hours in IBERGRID

Communities on-boarding

- EMSO is a large-scale research infrastructure of seafloor & water-column observatories, set up to monitor long-term environmental processes and their interactions.
- MELOA (Multi-purpose/Multi-sensor Extra Light Oceanography Apparatus) project that is developing, WAVY drifter units, for in-situ measurements of marine environments.
- BIOISI understand and address biological questions using integrative – Systems – approaches, joining biology, physics and computational sciences.



Resources on-boarding

- **BIFI** in Zaragoza integrating cloud capacity
- **CETA CIEMAT** in Trujillo reinforcement of cloud capacity with GPUs
- **DIGITAL.csic** data repositories from CSIC
- **IAA** in Granada integration of cloud resources for the SKA
- **INCD-B** site in Minho integrating grid and cloud capacity in collaboration with FCT.
- **INCD-C** site in Coimbra for long-term data storage.



INDIGO-DC *Software Collaboration Agreement*

IBERGRID teams are involved in the development of software to support innovative services for researchers in the European Open Science Cloud framework.

CSIC, LIP and UPVLC are signatories of the **INDIGO-Datacloud Software Collaboration Agreement**, a key Technology provider for the EOSC ecosystem.



INDIGO - DataCloud
Better Software for Better Science

The **INDIGO-DC Software Stack** Catalogue is a catalogue of open source software components that follow the Architecture defined by the INDIGO-DataCloud project, funded by the European Union under the Horizon 2020 Framework Program with Grant Agreement 653549.

This Collaboration Agreement targets **to sustain and further develop the INDIGO-DataCloud** architecture, the original INDIGO-DataCloud Software Catalogue and as well as the „INDIGO brand” beyond the lifetime of the INDIGO-DataCloud Project, through a not-for-profit, liability-free mutual....

Software Quality Management



A set of Common Software Quality Assurance Baseline Criteria for Research Projects

Abstract

The purpose of this document is to define a set of quality standards, procedures and best practices to conform a Software Quality Assurance plan to serve as a reference within the European research ecosystem related projects for the adequate development and timely delivery of software products.

Copyright Notice

Copyright © Members of the INDIGO-DataCloud, DEEP Hybrid-DataCloud and eXtreme DataCloud collaborations, 2015-2020.

Authors

Pablo Orviz Fernández (IFCA - CSIC), Álvaro López García (IFCA - CSIC), Doina Cristina Duma (INFN-CNAF), Giacinto Donvito (INFN-Bari), Mario David (LIP), Jorge Gomes (LIP).

Acknowledgements



The INDIGO-DataCloud, DEEP-Hybrid-DataCloud and eXtreme-DataCloud projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 653549, 777435 and 777367 respectively.

See: <http://digital.csic.es/handle/10261/160086>

DEEP-Hybrid-DataCloud

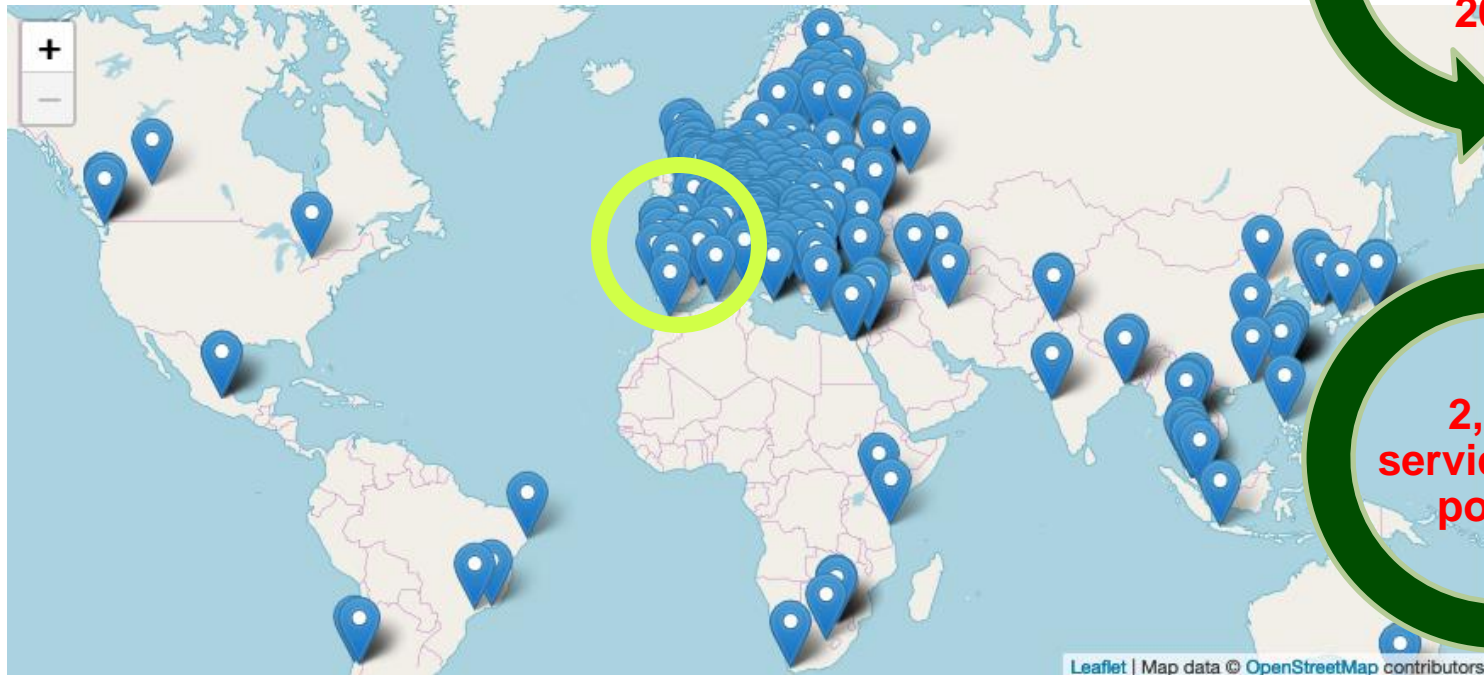
Supporting **artificial intelligence (machine learning and deep learning)**, parallel post-processing of very large data, and analysis of massive online data streams over distributed e-Infrastructures in the **European Open Science Cloud (EOSC)**.

- Support **intensive computing** techniques that require specialized **HPC** hardware, like **GPUs** or **low-latency** interconnects, to **explore large datasets**
- **DEEP as a Service** enabling the transparent training, sharing and serving of Deep Learning models both locally or on hybrid cloud system
- A DevOps approach for the project software and applications quality



EGI

- ❑ Federated e-infrastructure
- ❑ IBERGRID is a regional infrastructure in EGI
 - Shares services and capacity through EGI
 - Uses EGI services for integration and support



**4.4 Billion
CPU core
wall time
(2018)**

**> 1 Million
computing
cores in
2019**

**> 740 PB
disk & tape**

**2,915
service end-
points**

What is the EOSC?

A trusted and open virtual environment for the scientific community with seamless access to services addressing the whole research data life cycle.

EOSC

Is potentially available to 1.7m EU researchers and over 70 million professionals in science

Leverages on past infrastructure investment (10b per year by MS, two decades EU investment)

EOSC Services and Resources eosc-portal.eu



Networking



Compute



Storage



Sharing &
Discovery



Data
Management



Processing &
Analysis



Security &
Operations







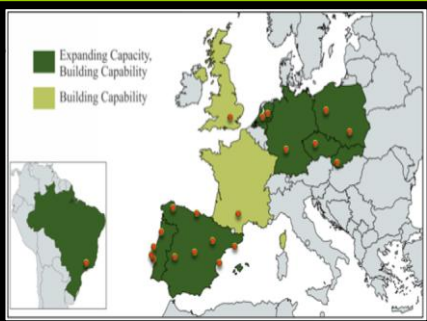







Training &
Support

IBERGRID in EGI and EOSC-hub

- Previously in EGI
 - Middleware rollout (LIP/IFCA)
 - Middleware criteria definition & validation (IFCA/CESGA/LIP)
 - EGI accounting (CESGA)
 - EGI support (LIP/IFCA)
- Now in EOSC-hub
 - Configuration and Change Management, Release and Deployment Management
 - EGI accounting portal
 - Processing and Orchestration
 - Common service requirements
 - Stakeholder Engagement Programme
 - Technical user requirements
 - Services and Software
 - OPIE, IM, udocker
 - LifeWatch, OPENCoastS



EU Funded Projects to implement EOSC

Acronym				
Region				
Coordinator				

- The EC has funded four projects to organize EOSC harmonization activities across Europe
- In the period 2019-2022 they will work together to implement the EOSC vision.

EOSC-synergy

- ❑ Infrastructure integration
- ❑ Fostering EOSC Service Quality
- ❑ Services for Research & Infrastructures
- ❑ Scientific Data Repositories Exploitation and FAIR principles
- ❑ Policy harmonization
- ❑ Training and self-learning





Vicente Hernandez

Javier Tobío



Gaspar Barreira



Further information

- <http://www.ibergrid.eu/>

Thank You

Cofinanciado por:

