IBERGRID

Prompting an EOSC in Practice in the Iberian area

Isabel Campos
isabel.campos@csic.es

Institute of Physics of Cantabria - CSIC
& EOSC High Level Expert Group

ES-NGI

CSIC
Consejo Superior de Investigaciones Científicas

IBERGRID

The objective of the EOSC is to give the Union a global lead in research data management and ensure that European scientists reap the full benefits of data-driven science, by offering '1.7 million European researchers and 70 million professionals in science and technology a virtual environment with free at the point of use, open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines'.
(https://ec.europa.eu/research/openscience)

EUROPEAN OPEN SCIENCE CLOUD - EOSC

Content of the presentation

- Reflexions from the EOSC HLEG
- How is IBERGRID positioned towards EOSC
- Innovation: Opportunities and Challenges

IBERGRID

The EOSC High Level Expert Group

Members
Isabel Campos Plasencia (CSIC, ES)
George Komatsoulis (Microsoft, USA, **Observer**)
Andreas Mortensen (EPFL, CH)
Silvana Muscella (Trust-IT, **Chair)**
Toivo Raïm (Estonian Ministry, ST)
François Robida (BRGM, FR)
Linda Strick (Fraunhofer, DE)
Klaus Tochtermann (German Libraries, DE)
Ziga Turk (University of Ljubljana, SLO)
Ross Wilkinson (Australian Library Ser., AU, **Observer**)

Mission:

Advise the EC on the measures needed to implement the European Open Science Cloud.

IBERGRID

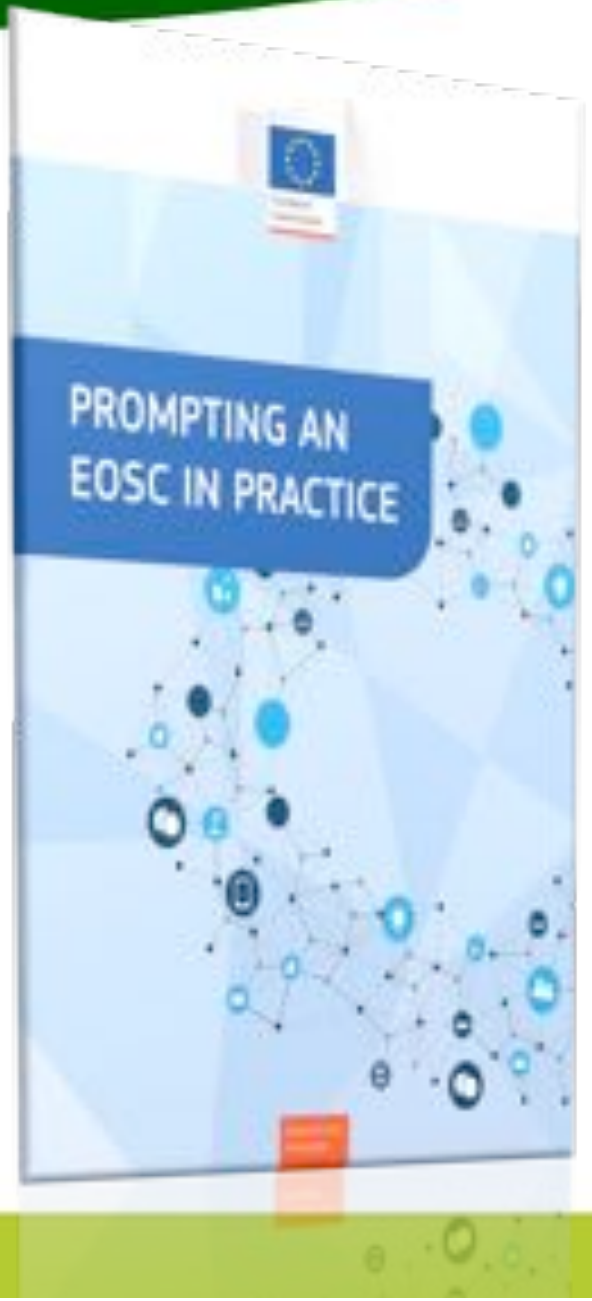
Framework of the EOSC HLEG

The HLEG work takes place through direct funding from the EC

- Collaborates with ongoing Horizon 2020 projects contributing to the establishment of the EOSC: EOSC-Pilot, EOSC-Hub

With the **Commission expert group on Turning FAIR data into reality** and the **Open Science Policy Platform.**

- **Summary report of the Stakeholder** workshops (June 2017 and October 2018);
- **Interim report** – May 2018;
- **Final Roadmap Report:** advice on implementation of the EOSC preparatory phase Dec. 2018



IBERGRID

The rational

THE EOSC
MINIMUM VIABLE ECOSYSTEM

IBERGRID

Designing EOSC as a Minimum Viable Ecosystem

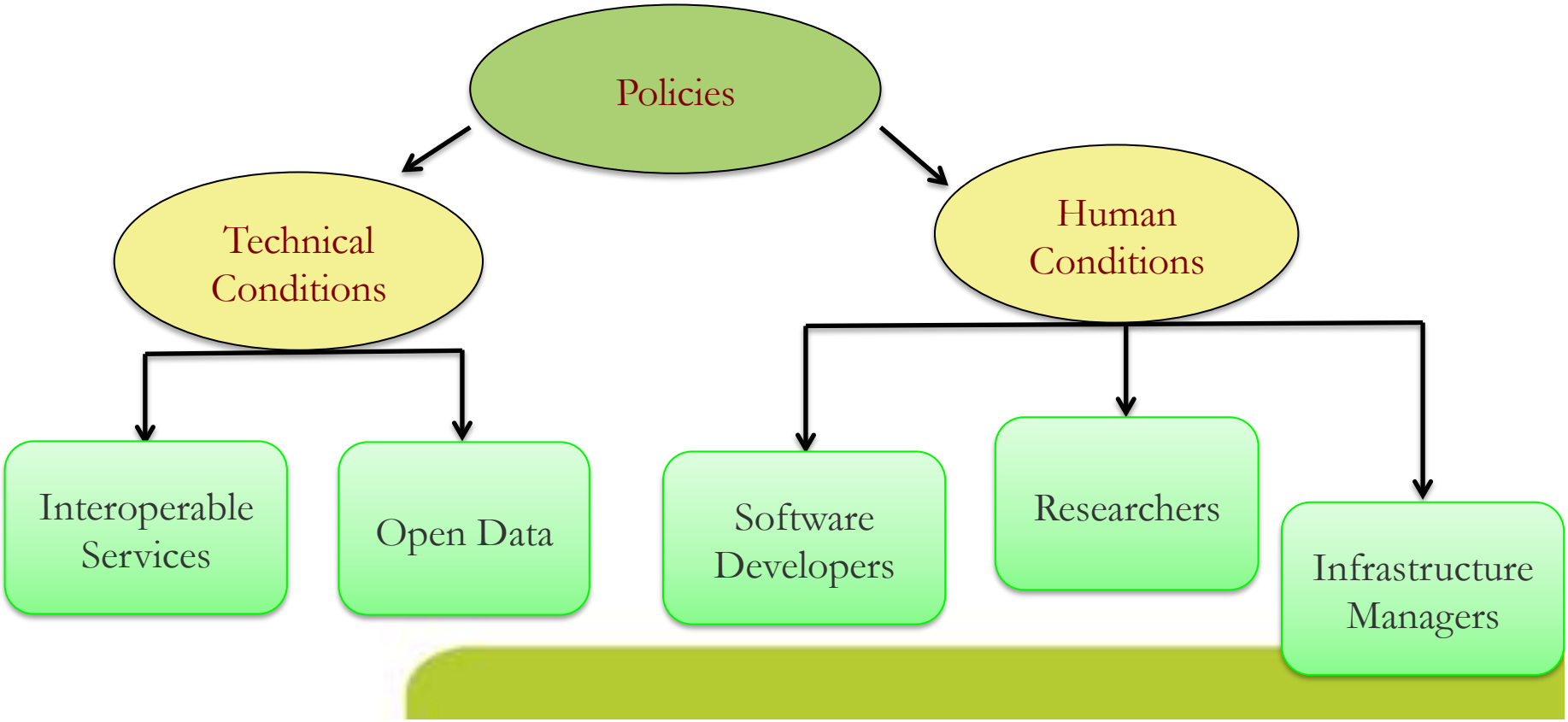
EOSC aims to put Europe at a global lead regarding scientific data infrastructures

The provision of infrastructures, technology development, and human resources to support it will take place in a **very heterogeneous landscape**.

Addressing this challenge requires the **definition of a smallest common denominator: the EOSC Minimum Viable Ecosystem (MVE).**

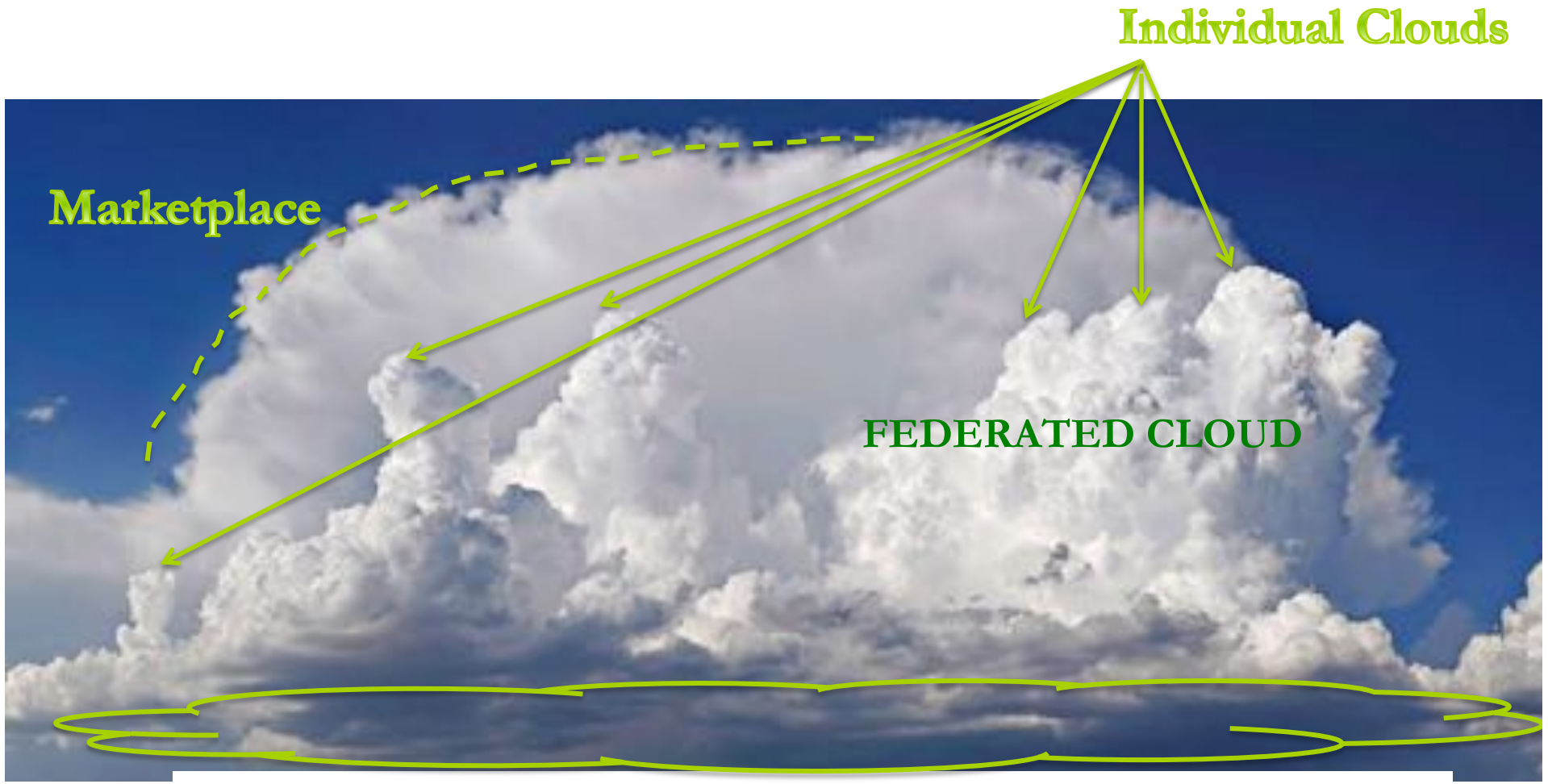
Minimum Viable Ecosystem

A Minimal Viable Ecosystem (MVE) will emerge if Technical and Human / Sociological conditions are met.



IBERGRID

EOSC is an effort to enable Hybrid Clouds



EOSC Infrastructure Managers Incentives

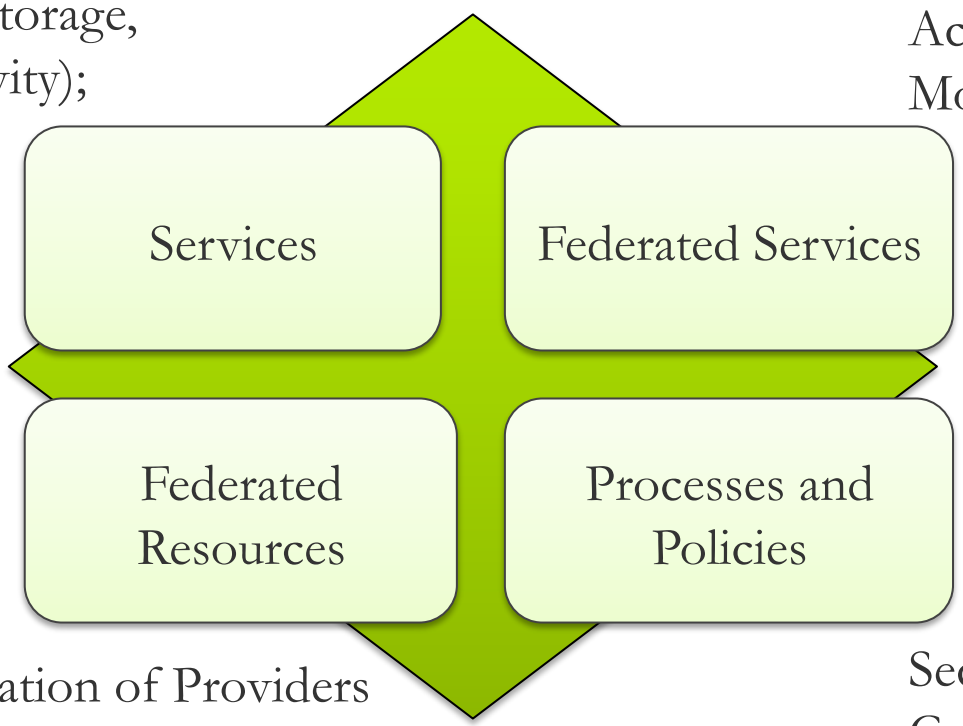
- Infrastructure **publicly funded by national agencies** and providers
  - Banking on the economy of scale to support international collaborations
  - Provide a "stable" infrastructure safe from money flow issues that affect Research funding lines.
  - Provide the environment to do IT independent research at State of the Art (piloting innovative services).

- Infrastructure **provisioned by Commercial clouds** (direct market offer, pre-procurements,…)
  - Direct market offer: could be a solution for "peaks of demand".
  - Pre-procurement procedures for "flat budget" organizations.

The Interoperability Challenge for EOSC

Applications & tools;
Baseline services (storage, compute, connectivity);
Training…

Marketplace
AAI
Accounting
Monitoring….



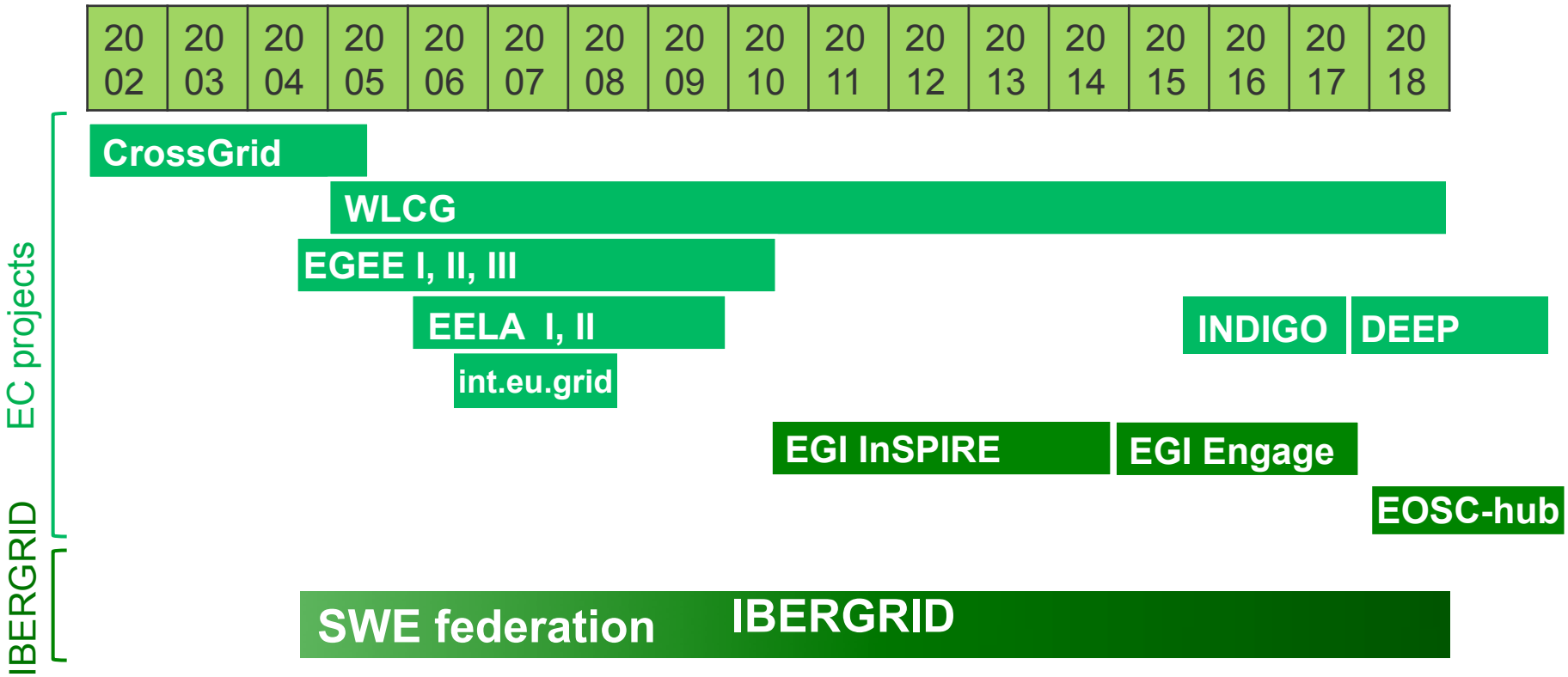
Lightweight certification of Providers
SLA negotiation…

Security regulations,
Compliance to standards,
Terms of use,
FAIR guidelines ...

IBERGRID towards the EOSC

IBERGRID BACKGROUND

IBERGRID history and origins



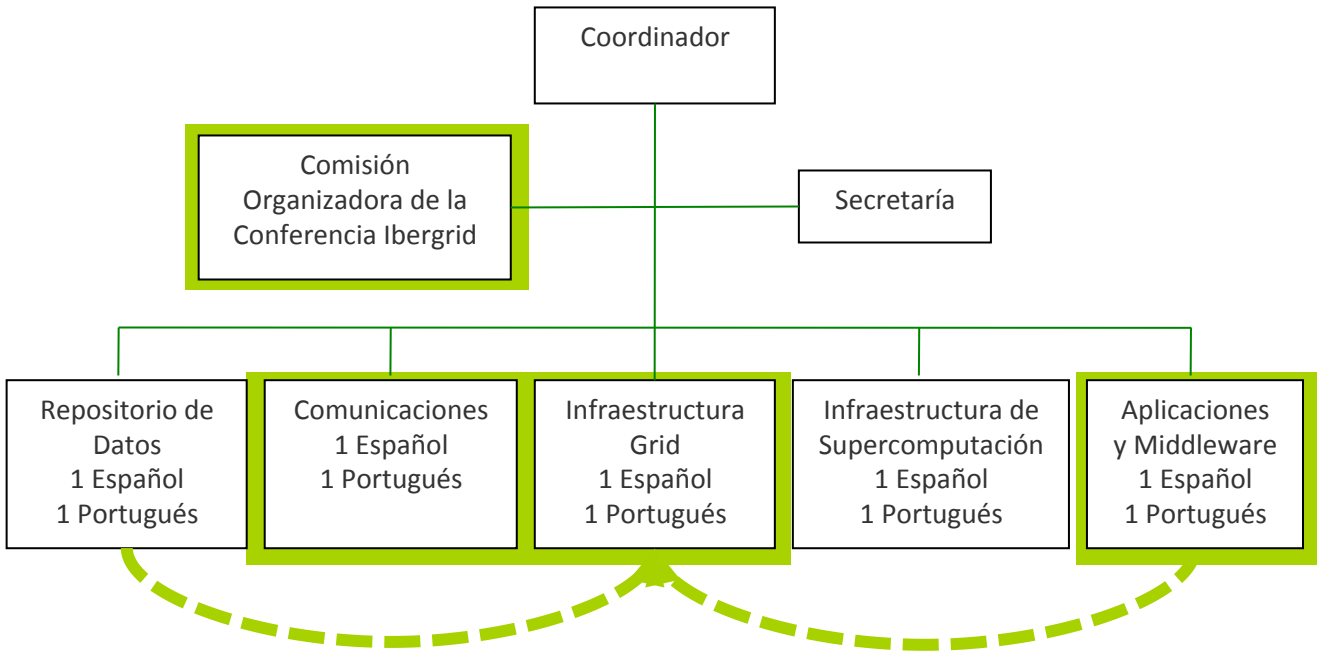
Many successful e-science projects with strong PT + ES partnership and also extending to Latin America

Timeline

- **8-Nov-2003 Scientific and Technology Cooperation Agreement between Portugal and Spain** (Figueira da Foz)
- **19-Nov-2005 Memorandum of Understanding** (Évora)**:** Ministério da Ciência, Tecnologia e Ensino Superior **+** Ministerio de Educación y Ciencia
- **25-Nov-2006 Memorandum of Understanding** (Badajoz)**:** Ministério da Ciência, Tecnologia e Ensino Superior **+** Ministerio de Educación y Ciencia
- **22-Jan-2009 Memorandum of Understanding** (Zamora)**:** Ministério da Ciência, Tecnologia e Ensino Superior **+** Ministerio de Educación y Ciencia ➔ **IBERGRID federated infrastructure**

IBERIAN common plan

Initial plan released in May 2007



- 1st phase
  - Operations Centre
  - Establish 4 VOs
  - Training
  - Sustainability study
- 2nd phase
  - Expansion to more Resource Centres
  - Expand VO coverage
- 3rd phase
  - Expand to supercomputing centres

- In green the areas that become active and are operational

IBERGRID

Distributed computing infrastructure

GOBIERNO DE ESPAÑA
MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

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

ES-NGI

IFCA
Instituto de Física de Cantabria
Tech coordination

INCD

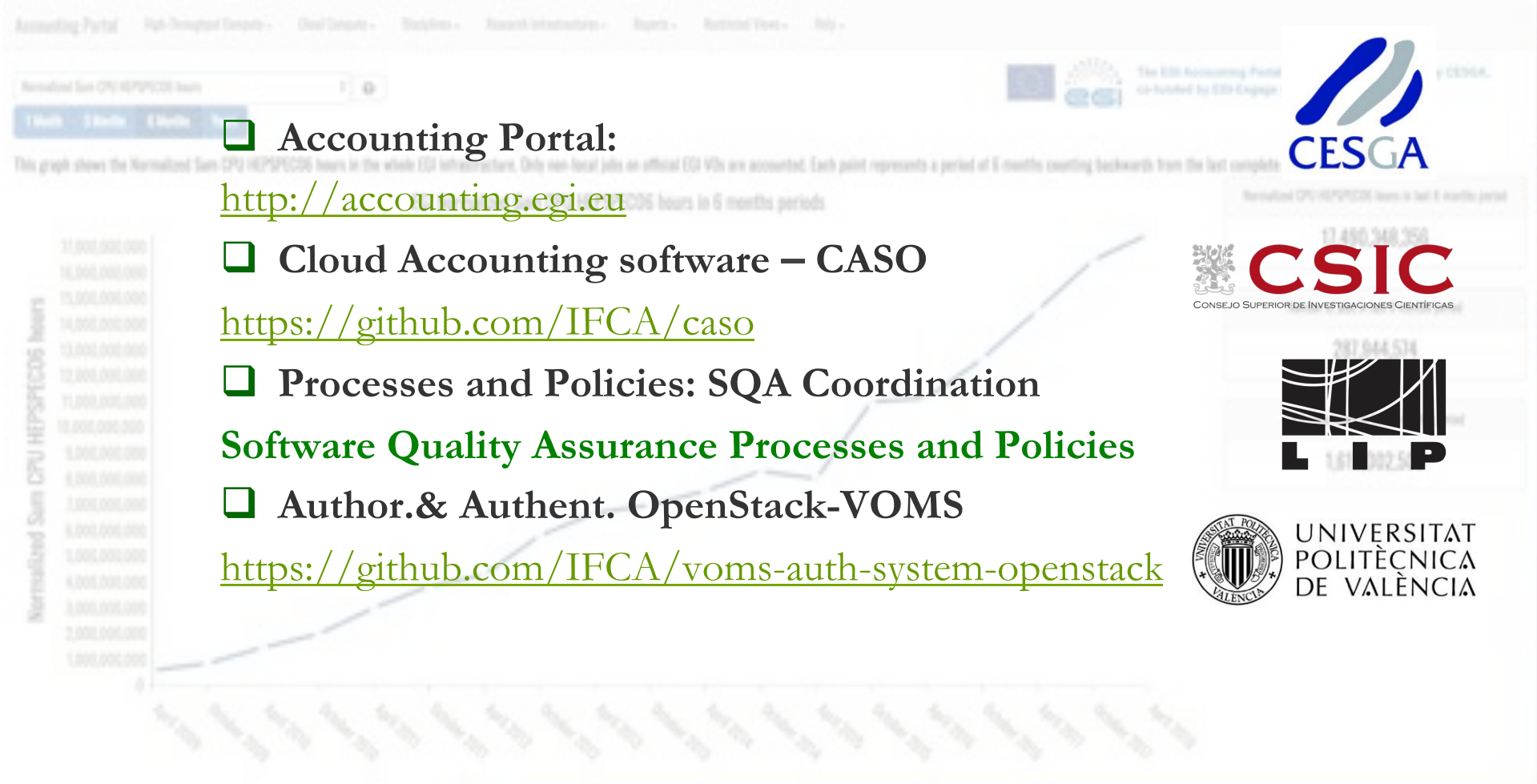
LIP
Tech coordination

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 2018 in EOSC-hub
  - Configuration and Change Management, Release and Deployment Management
  - EGI accounting portal
  - Processing and Orchestration
  - Common service requirements
  - Stakeholder Engagement Programme
- Global tasks:
  - Coordination of European wide activities

IBERGRID role to Federate in the EOSC-hub Catalogue

https://ibergrid.eu/federating-clouds/



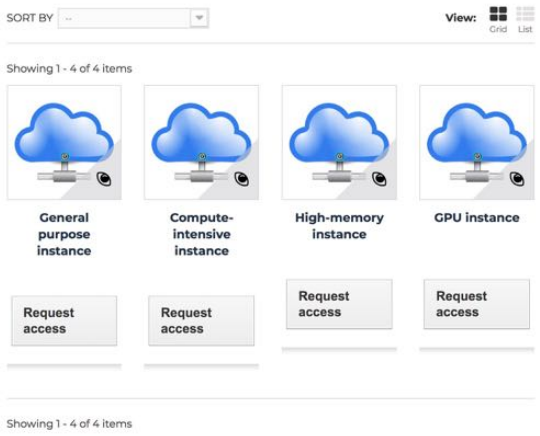
- **Accounting Portal:**
  http://accounting.egi.eu
- **Cloud Accounting software – CASO**
  https://github.com/IFCA/caso
- **Processes and Policies: SQA Coordination**

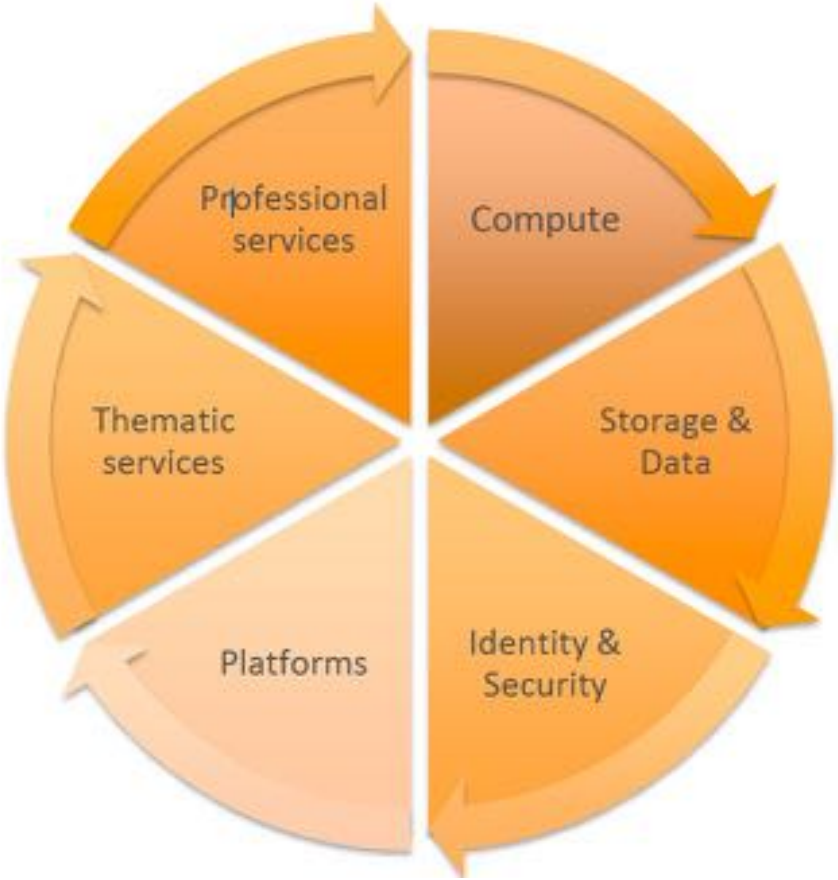
Software Quality Assurance Processes and Policies

- **Author.& Authent. OpenStack-VOMS**
  https://github.com/IFCA/voms-auth-system-openstack

EOSC Marketplace

- **Open catalogue** of services
- **Integrated** discovery, order & access management
- Integrated with **Authentication**
  (AARC blueprint conformant)





https://marketplace.eosc-hub.eu

IBERGRID developments in the EOSC-HUB Marketplace

- **CONTAINER SERVICES**
  - **UDocker** (LIP)

- **COMPUTING SERVICES**
  - OpenStack Preemptible Instances - **OPIE** (CSIC)
- **PaaS SERVICES**
  - Infrastructure Manager - **IM** (UPVLC)
- **THEMATIC SERVICES**
  - LIFEWATCH (Lifewatch ERIC)
  - OPENCOASTS (LNEC)

IBERGRID

IBERGRID towards the EOSC

INNOVATION IN THE EOSC ECOSYSTEM

Innovation Potential:
"*Engaging Human Talent*"

- EOSC Infrastructure: in general a *hybrid landscape,* based on
  - Public and Private Clouds
  - Commercial or Publicly Funded Infrastructures
- Scientific Communities usually **pilot the deployment of innovative services**: advanced requirements.
- Software developers are a fundamental part in this ecosystem: **recognition**
- EOSC should implement policies to harness the potential of European developers in academia and industry.

Breakthrough ideas leading to innovation need to be awarded with the proper Recognition

IBERGRID towards the EOSC

IBERGRID INNOVATION POTENTIAL IN SOFTWARE DEVELOPMENT

Example: *Preemptible* (or *Spot*) cloud instances or, how to make extra money with idle CPU cycles?

Google Cloud launches preemptible GPUs with a 50% discount

Google Cloud

Short Lived, Low Cost VMs

Preemptible VMs are highly affordable, short-lived compute instances suitable for batch jobs and fault-tolerant workloads. Preemptible VMs offer the same machine types and options as regular compute instances and last for up to 24 hours. If your applications are fault-tolerant and can withstand possible instance preemptions, then preemptible instances can reduce your Google Compute Engine costs significantly.

amazon web services™ EC2

The only difference between On-Demand instances and Spot Instances is that Spot instances can be interrupted by EC2 with two minutes of notification when EC2 needs the capacity back. You can use EC2 Spot for various fault-tolerant and flexible applications, such as test & development
transcoding, and to run an
(HPC) workloads. EC2 Spo
Auto Scaling, Elastic Conta
Batch, providing you freed
running on Spot instances.

TM
closed source

Example: *Preemptible* (or *Spot*)cloud instances

Independent, State of the Art Research, delivering Innovation

INDIGO - DataCloud

OpenStack Preemptible Instances – OPIE
https://github.com/indigo-dc/opie

Alvaro, Premio a la mejor
Tesis en "Ciencias Experimentales"

open source™

EOSC-hub

IBERGRID

Incentives for Software Developers: *"Engaging Human Talent"*

- *"EOSC-Ready"* **as a Branding for software** products, will harness the potential of the European developers, both in academia and industry.

What does it mean "EOSC-Ready" **? Trustable**

- **Open Source** → contributed upstream (potentially sustainable)
- Based on **Open Standards**
- Adheres to code style quality standards → **Software Quality**
- It **respects the EC directives** on cloud security (**NIS**), personal data protection (**GDPR**) and Free Flow of non-personal Data (**FFD**).

Software activity history as Iberian Federation

INDIGO - DataCloud

2014-2017

2010-2013

CSIC coordinated **2005-2007**
7 countries / 12 partners
(**5 Iberian**)

2009…

EGI.eu is funded in Feb. 2009.
Software Quality task EGI-wide

2006-2010
Many…

2002-2005
11 countries / 21 partners
(**5 Iberian**)

2005-2010
"Southwest Federation"
50 partners/8-10 Iberian

@ Jesús Marco & Jorge Gomes

IBERGRID

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 are 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.

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

IBERGRID

Hybrid DataCloud

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777435.

- **Evolve up to production level intensive computing services** exploiting specialized hardware: **Hardware Accelerators**: **GPUs** and **low-latency** interconnects
- **Integrate** intensive computing **Services under a Hybrid Cloud** approach ensuring interoperability expanding over multiple IaaS using high level networking technologies
- **Offer a DevOps approach for application development**

CSIC LIP INFN PSNC KIT

UNIVERSITAT POLITÈCNICA DE VALÈNCIA cesnet Ústav Informatiky Atos HelmholtzZentrum münchen

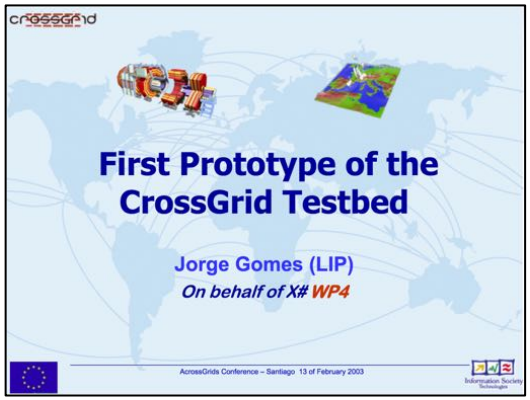
IBERGRID

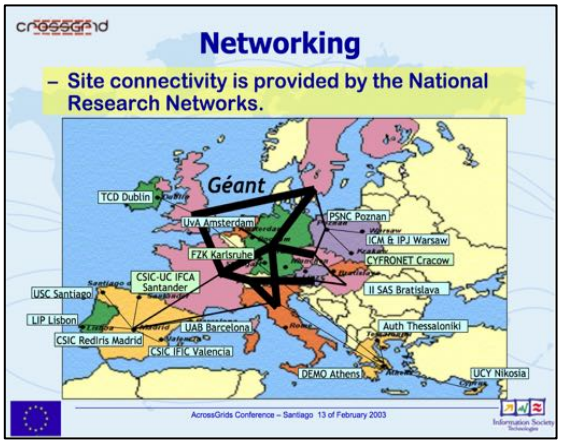
… OUR ROOTS GO DEEP

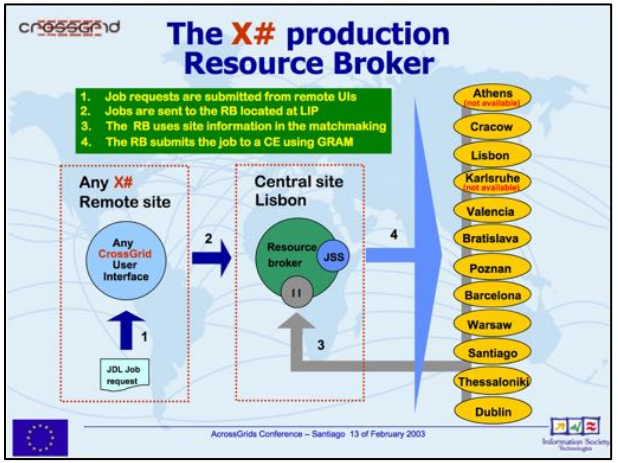
In retrospective: CrossGrid 2002-2004:

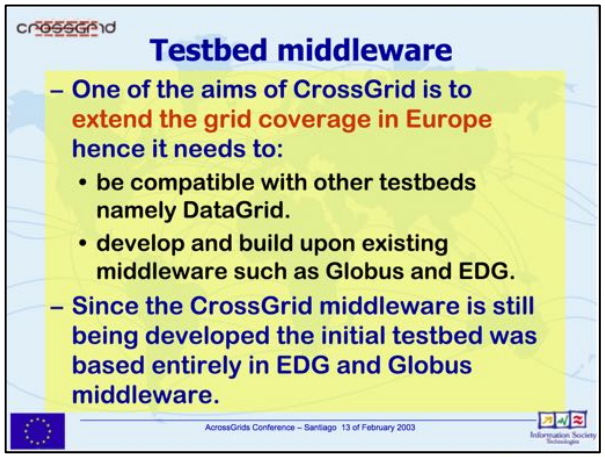
AcrossGrids Conference, **February 2003, Santiago de Compostela**

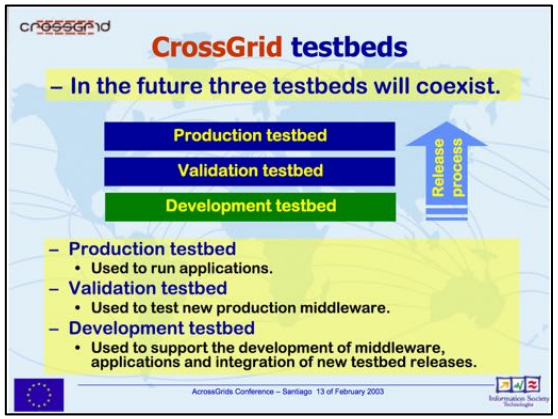
http://slideplayer.com/slide/4798848

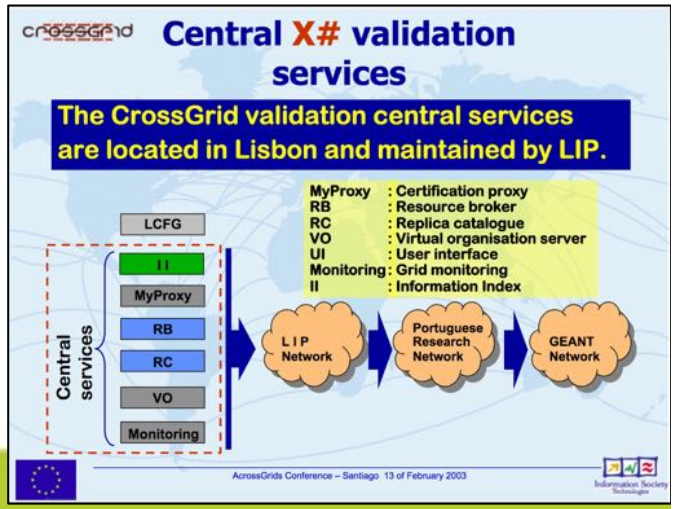










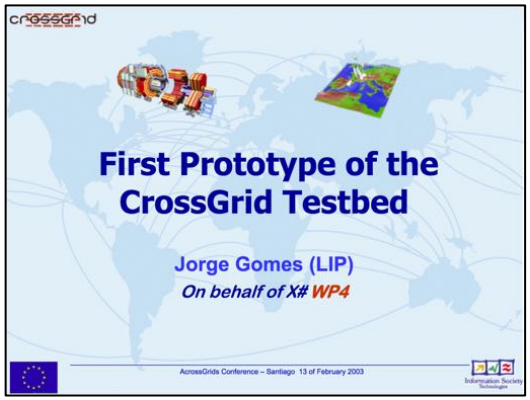


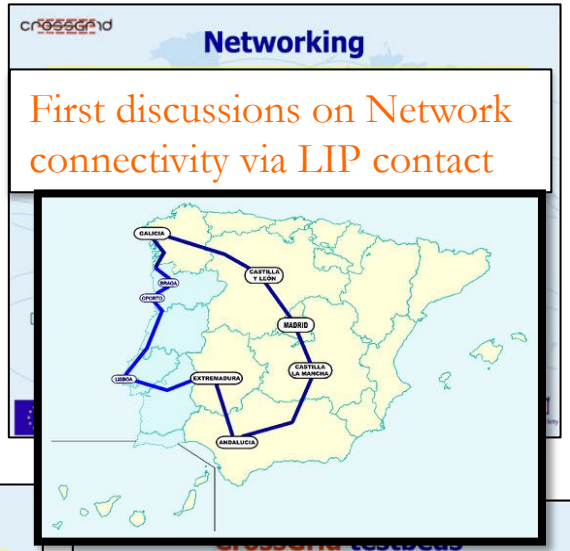
IBERGRID

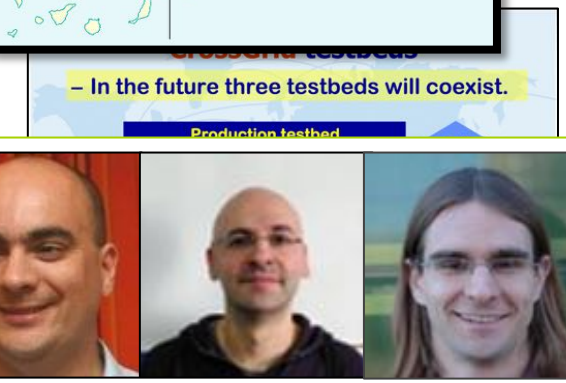
In retrospective: CrossGrid 2002-2004:

AcrossGrids Conference, **February 2003, Santiago de Compostela**

http://slideplayer.com/slide/4798848







Pablo, Alvaro, Mario, Joao, Miguel

EOSC-hub

Software Verification Team