Applications and gateways for the LSGC: EGI's role as a helper and supporter

Steve Brewer, Chief Community Officer, EGI.eu

steve.brewer@egi.eu

e-infrastructure

Introduction

I am here primarily to listen, but first…

I have a story to tell…

1. Managing a sustainable infrastructure
2. Building sustainable user services
3. Cultivating sustainable communities

Where do you fit in?

presentation outline

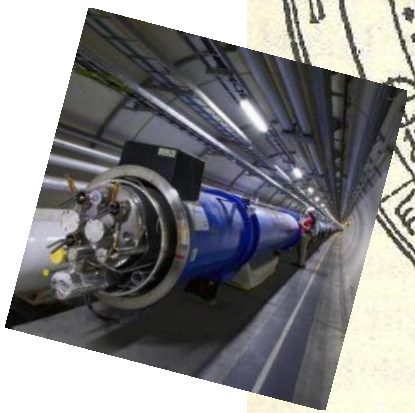
1. Managing a sustainable infrastructure
   - Background – Data deluge, Grid, infrastructure, EGI
   - Resource infrastructure, EGI-InSPIRE project
   - Virtuous Service Cycle, EGI supporting innovation
2. Building sustainable user services
   - User Support structure
   - Support services
   - Processing requirements, website
3. Cultivating sustainable communities
   - VRCs
   - Life Sciences Grid Community (LSGC)
   - What happens next for Life Sciences community

1) Managing a sustainable infrastructure

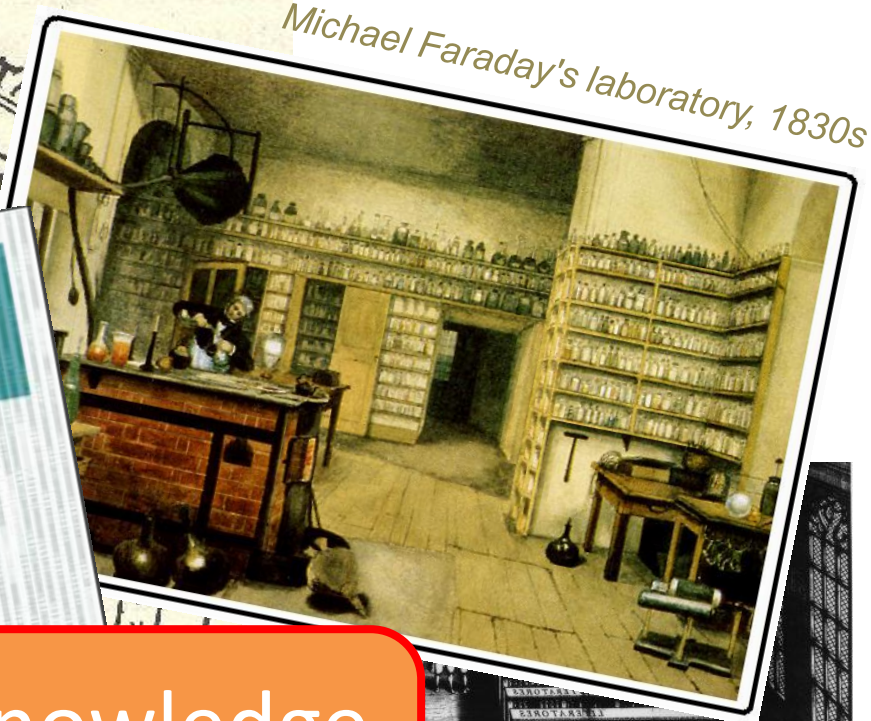
The Data Deluge

Barcelona Super Computer

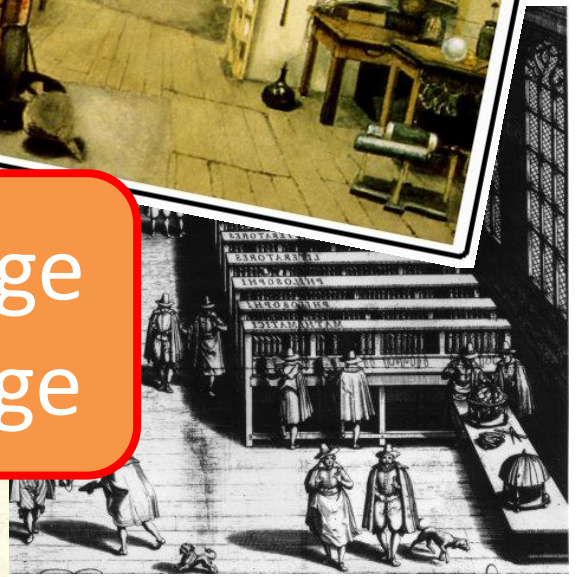
The LHC



The Geocentric universe, pre-1500



Michael Faraday's laboratory, 1830s



Leiden University chained library 1610

Extracting Knowledge from the Data Deluge

What is a Grid?

- A grid consists of distributed resources controlled by separate organisations that be systematically used securely by users external to that organisation
- Resources can include:
  - Commodity or HPC clusters
  - Disk or tape storage
  - Instruments
  - Data Archives or Digital Libraries

Infrastructure (Wikipedia)

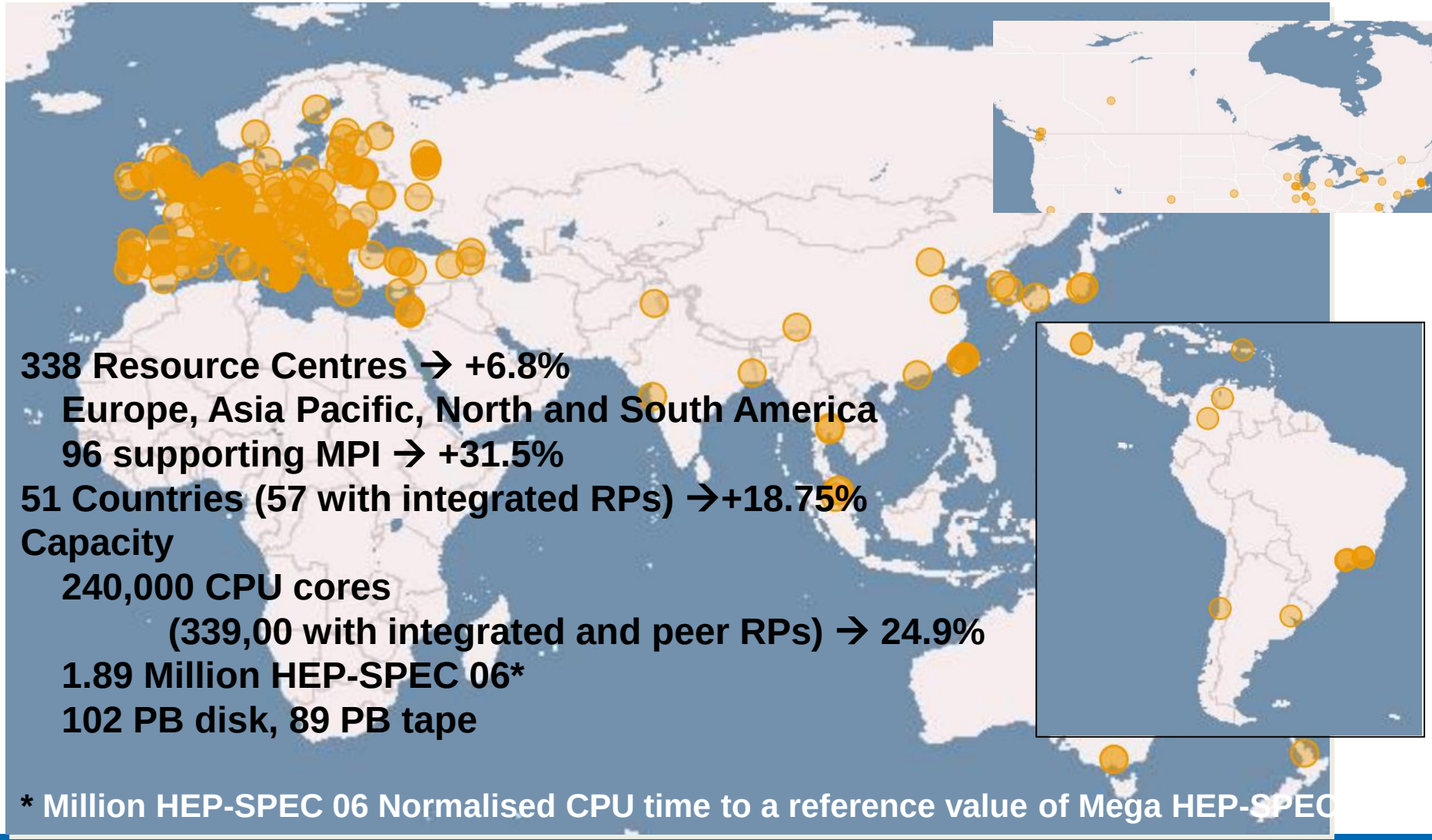
Infrastructure is the basic physical and organisational structures needed for the operation of a society or enterprise, or the services and facilities necessary for an economy to function

The Enterprise is the European Research Area

EGI provides a service infrastructure that exposes and helps coordinate a resource infrastructure

Resource Infrastructure

(Showing yearly increase)



338 Resource Centres → +6.8%

Europe, Asia Pacific, North and South America

96 supporting MPI → +31.5%

51 Countries (57 with integrated RPs) →+18.75%

Capacity

240,000 CPU cores

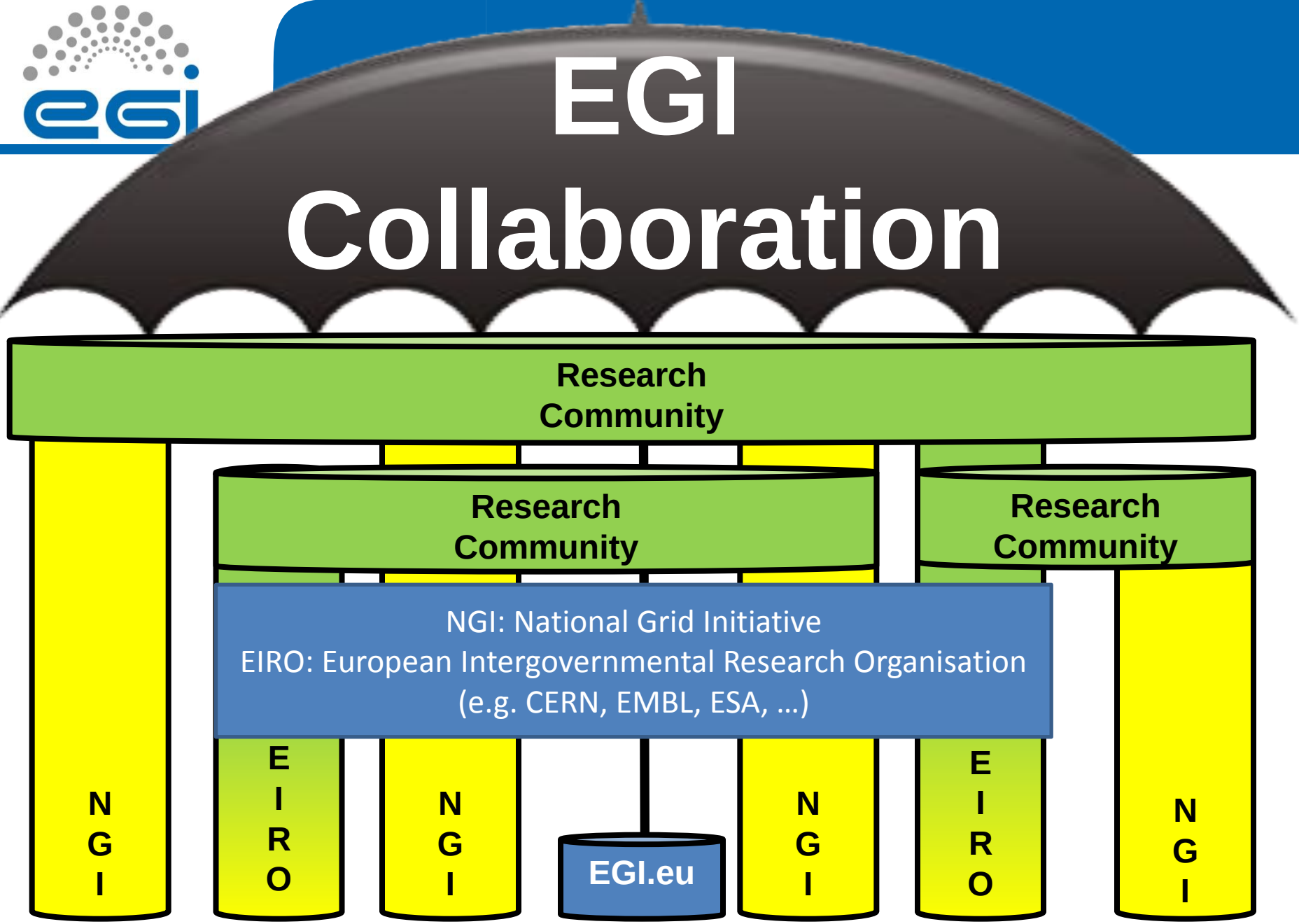
(339,00 with integrated and peer RPs) → 24.9%

1.89 Million HEP-SPEC 06*

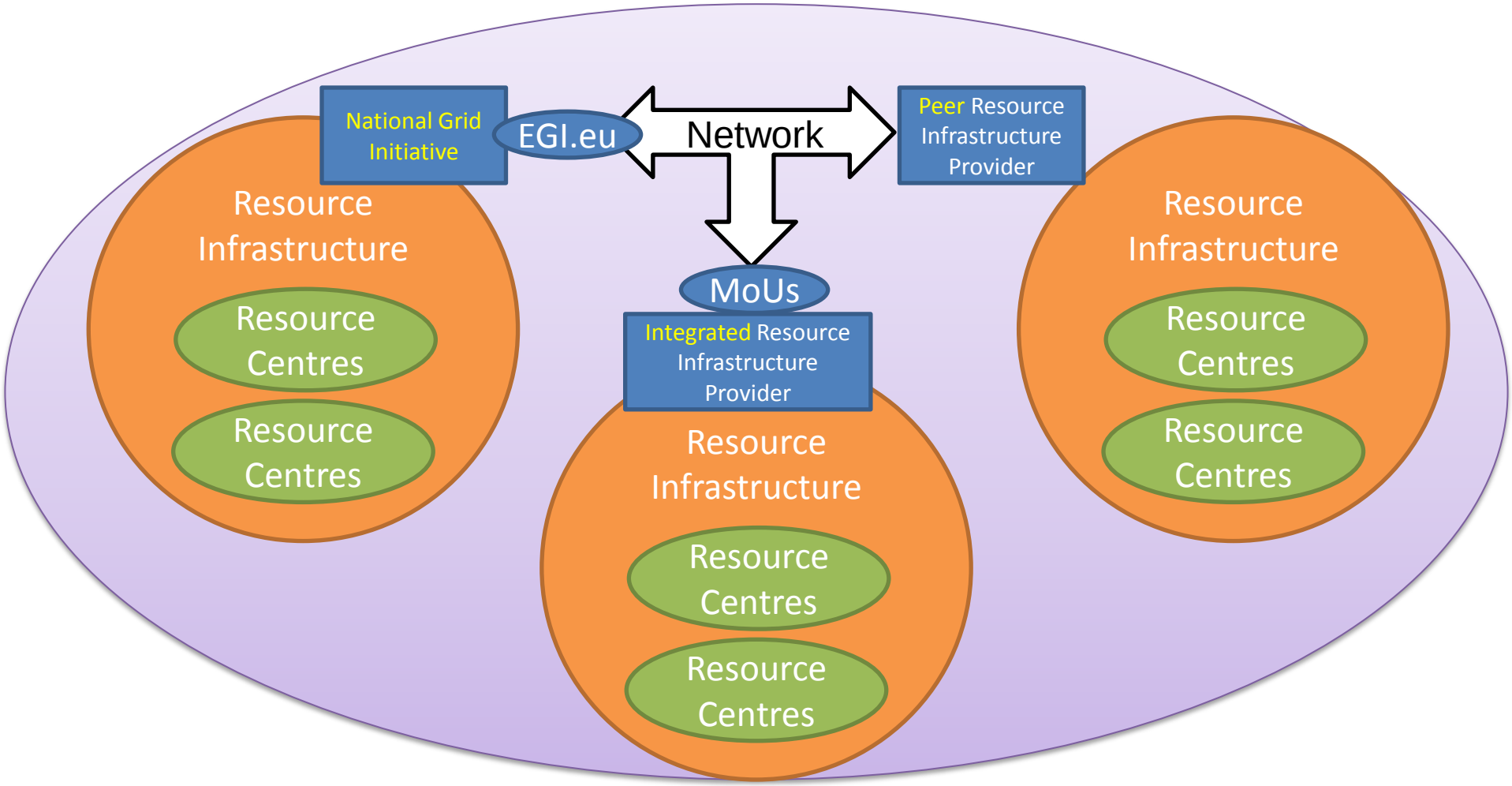
102 PB disk, 89 PB tape

*** Million HEP-SPEC 06 Normalised CPU time to a reference value of Mega HEP-SPEC**

EGI Collaboration



EGI Resource Infrastructure



Integrated: infrastructure operated by a non-EGI-InSPIRE partner but relying on EGI operational services (MoU)
Peer: infrastructure accessible to EGI users, but relying on own operational services

EGI.eu

- Coordination for European Grid resources
  - Established February 8th 2010
  - Central policy & services needed to run a grid
  - Sustainable small coordinating organisation
- Based in Amsterdam
  - Coordinating core (~20 people) in Amsterdam
  - Technical services from partners (~20 people)

EGI and EGI.eu supported by EGI-InSPIRE project

EGI-InSPIRE Project

Integrated Sustainable Pan-European Infrastructure for Researchers in Europe

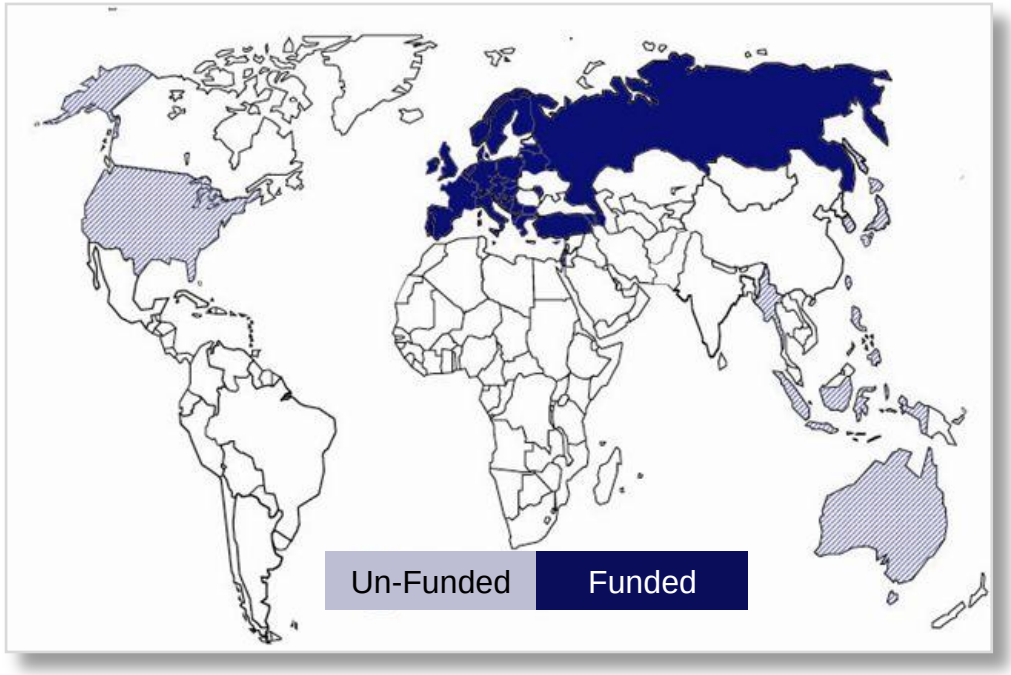
A 4 year project with €25M EC contribution

- Project cost €72M
- Total Effort ~€330M
- Effort: 9261PMs

Project Partners (50)

EGI.eu, 38 NGIs, 2 EIROs

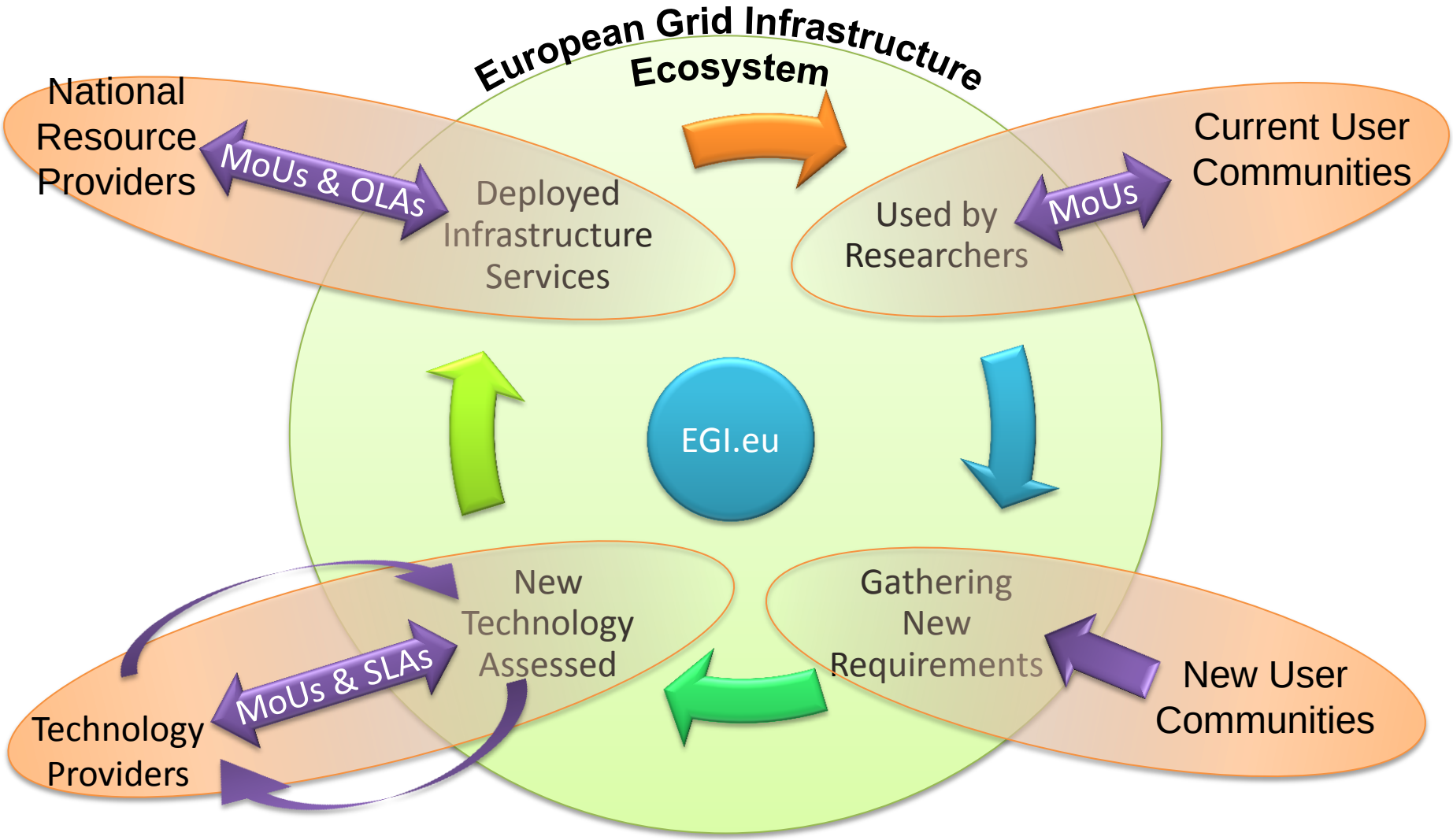
Asia Pacific (9 partners)



EGI-Inspire project objectives

- A sustainable production infrastructure
  - Resource providers in Europe and worldwide
  - With new technologies as they mature
- Support structured international research
  - Sustain current domain specific services
  - Attract new user communities (e.g. ESFRI)

A Virtuous Service Cycle



EGI supports Innovation

- Deploy Technology Innovation
  - Distributed Computing continues to evolve
    - To include: Grids, Desktops, Virtualisation, Clouds, …
- Enable Software Innovation
  - Provide reliable persistent technology platform
    - Tools built on gLite/UNICORE/ARC/Globus
- Support Research Innovation
  - Infrastructure for data driven research
    - Support for international research (e.g. ESFRI)

Key Points

- EGI is what you make of it!
  - Many national resources already engaged
  - Easy to federate in new resources
  - You retain control as to which VOs can use it
- Mature services to expose resources
  - Ideal for large-scale data analysis
- Platform for collaborative sharing
  - Across compute, data, ….

2) Building sustainable user services

www.egi.eu/usersupport

User Support & Services

- Support User Communities:
  - Researchers in International Collaborations (VRCs)
  - National Research Collaborations through the NGI
  - Scale up from the single VO to a community
- Provide core services to support users:
  - Manage VOs, AppDB, Training Services, Requirements
- Support teams:
  - EGI.eu User Community Support Team
  - NGI User Support Teams
  - NGI Operations Teams
  - Experts within user communities or projects

Support coordination: cultivating sustainable communities

Success stories, user requirements, application details and training information

EGI support services and dissemination activity

User communities consist of:

- Scientific end users
- Software and technology experts
- System administrators
- Community representatives

National and specialist support units deliver services, tools & training

Sustainable EGI Virtual Research Communities (VRCs)

Lessons learnt

Sustainability of user communities is the key to achieve a sustainable e-Infrastructure.

Coordinating support: developing and deploying sustainable services

Feedback through VRCs in the User Community Board

Discover

Where is the community?

- VRCs
- Mailing lists
- Workshops
- Forums
- Blogs
- Projects
- Sharing stories
- Collaborating

Design

User Services Advisory Group to drive detailed design

How can I contribute?

- Applications
- Data collections
- Requirements
- Proposals
- Projects
- Success stories

Deliver

Integrated Services

- Human
- Technical
- Infrastructure

How do I use these resources?

- Attend training courses
- Utilise training material
- Access data
- Run applications on the grid

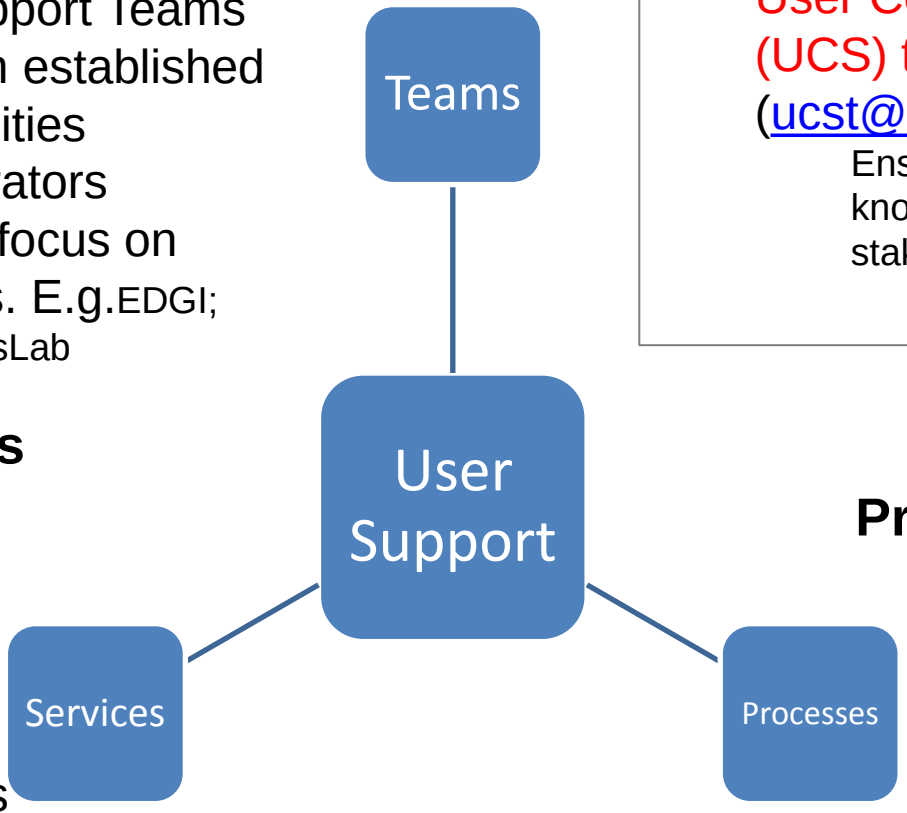
User Support structure

Support teams

- NGI User Support Teams
- Experts within established user communities
- Site administrators
- Projects with focus on specific areas. E.g.EDGI; MAPPER; StratusLab

Technical services

- Application Database
- Training Marketplace
- VO services
- Requirements Tracker



User Support Coordination

User Community Support (UCS) team of EGI.eu (ucst@egi.eu)

Ensures the flow of information , knowledge and solutions among stakeholders

Processes

- User Community Board
- VRC accreditation
- Requirements processing
- Training Working Group

NGI User Support Teams

NGI User Support Teams provide:

- Consultancy & training
- Access to grid services
- Application porting
- Developing and operating software services
- Collecting feedback
- Documentation & helpdesk

NGIs receive:

- Web & wiki resources
- Requirements process
- Applications and services
- VRC coordination support

Support teams are available across 50 partners – they are all different… as are their support skills.

Support services

Provided by a few NGIs **for the whole community**

- To facilitate the work of NGI support teams
- To improve communication among NGIs, among users and with users

Tools and services:

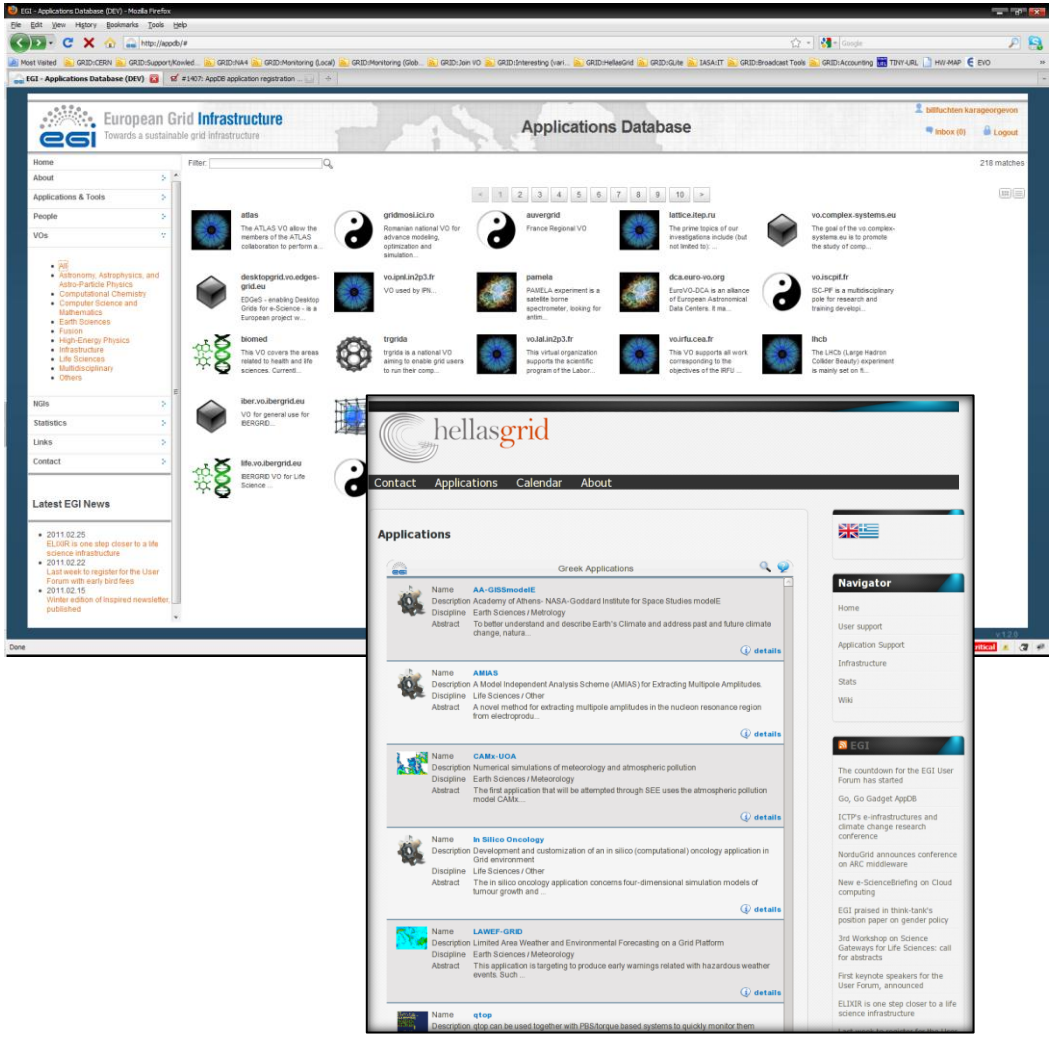
1. EGI Application Database (GRNET)
   - Domain specific application – for scientific end users
   - Reusable toolkits and services – for NGI support teams
2. Services for VOs (CSIC; LIP)
   - Tools, manuals, assistance for VO setup and management
3. Training Marketplace (STFC, UK)
   - Repository of events and materials
   - Place to request and provide NGI training services
4. Requirement Tracker (EGI.eu with CESNET)
   - To capture needs from the whole community
   - To offer solutions from NGIs

EGI Application Database

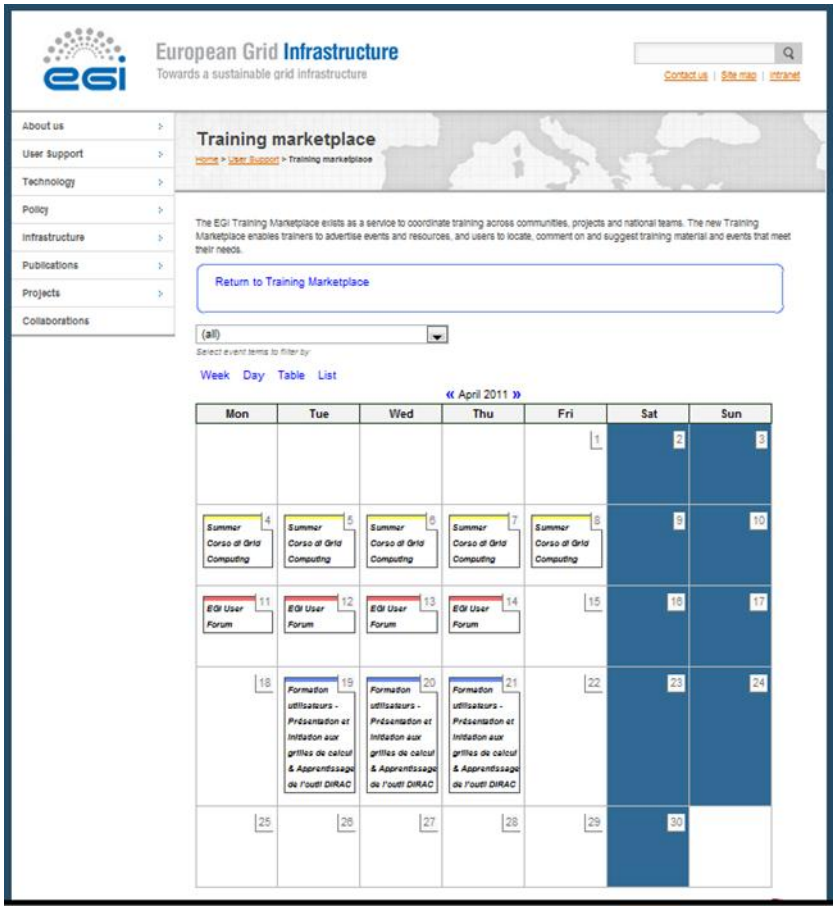
Benefits of AppDB:

- gives recognition to reusable applications
- gives recognition to application developers

- How to get involved?
  - Register applications
  - Reuse applications
  - Integrate AppDB through its gadget into any Webpage!



Training marketplace

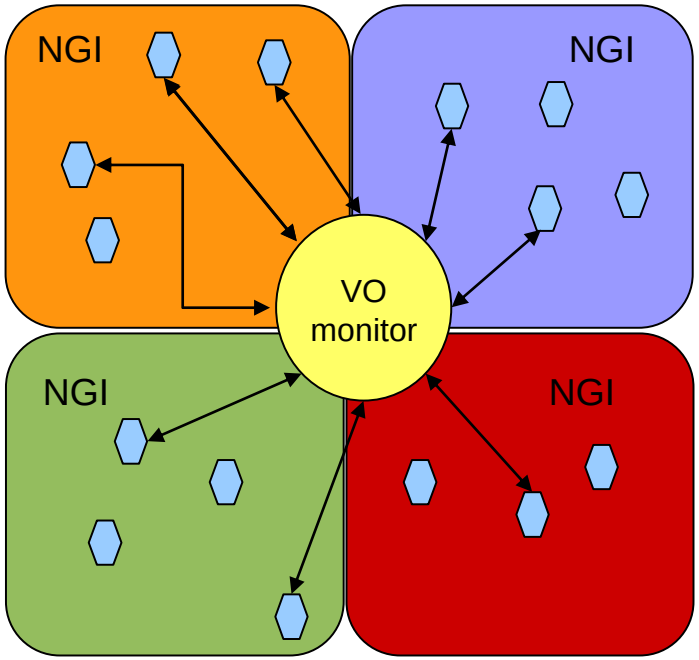


- Integration of:
  - training event calendar
  - Digital library (of training materials)
  - Service offering & requesting form
- How to get involved?
  - Register training events
  - Use, reuse and share training materials
  - Offer training-related services (infrastructure, VO, CA, ...)
  - Feed training-requirements
  - Delegate representative to EGI Training Working Group

You can offer and request ANY type of training-related service, tool, expertise through the new Training Marketplace

Services for VOs

- Activities:
  - Consultancy and helpdesk for VO managers
  - Evaluation of VO management, monitor and accounting tools
  - Provision of VO support software for VRCs

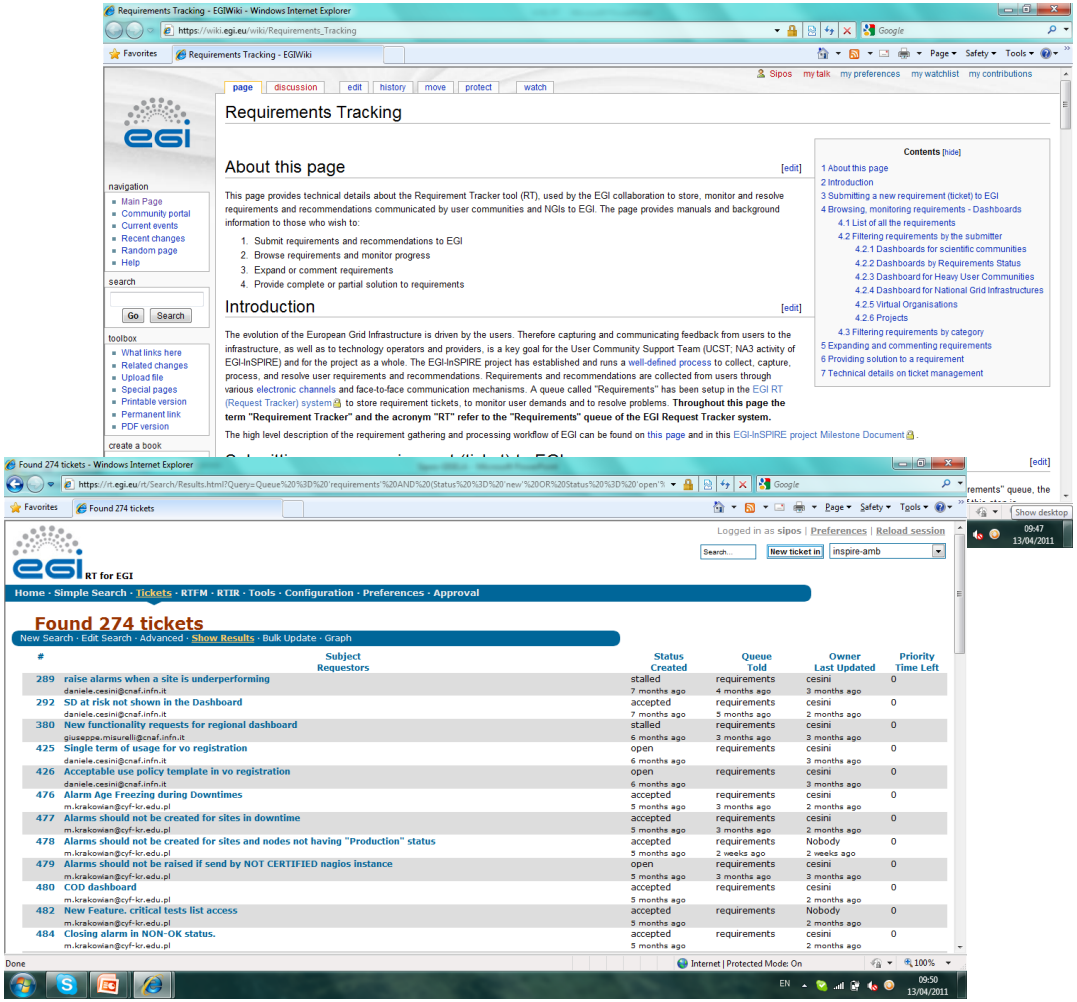


- VO-specific monitoring:
  - Monitor only those sites that support you
  - Create and plug-in VO-specific probes

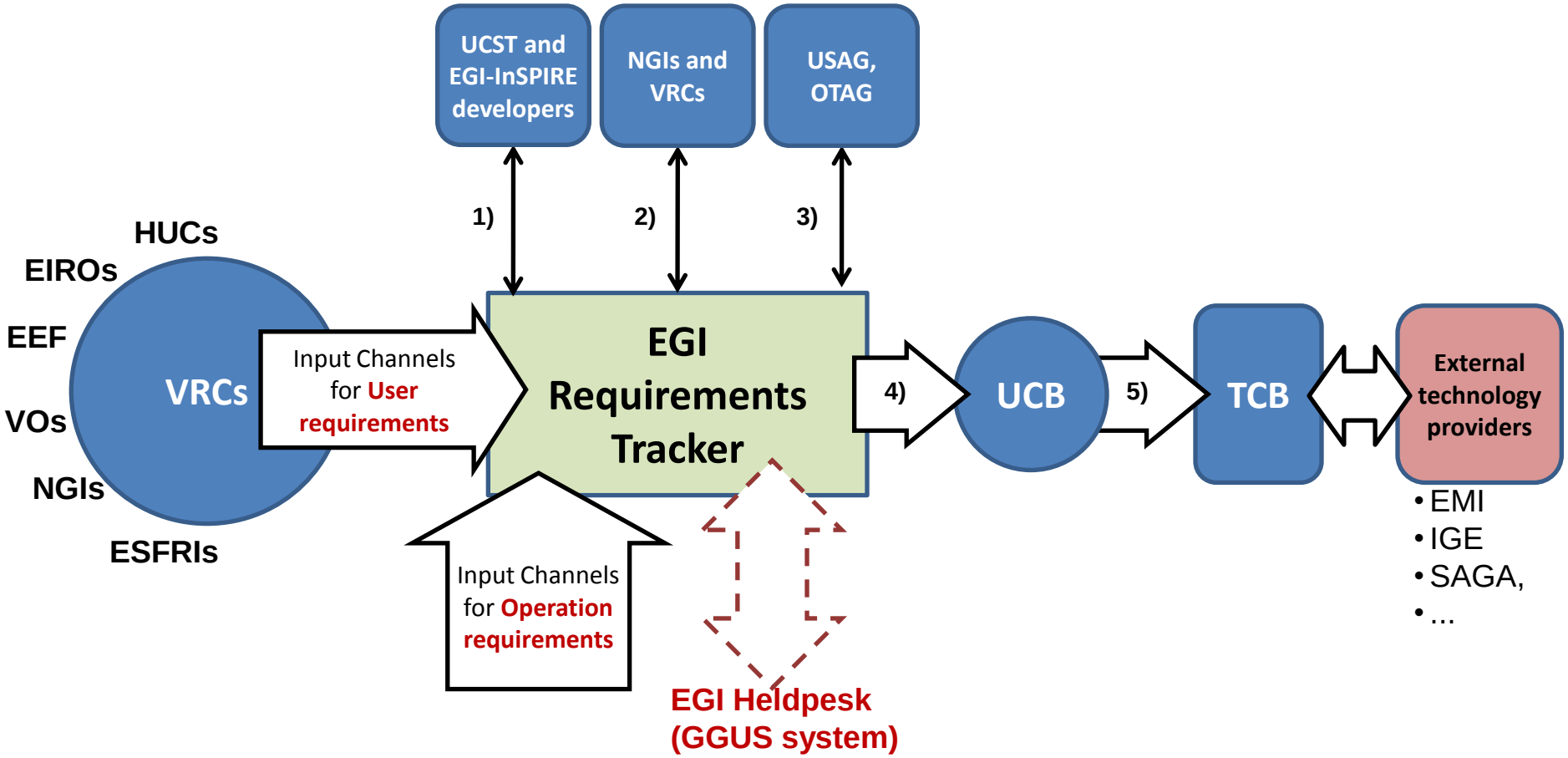
- How to get involved?
  - Request support
    - vo-services@mailman.egi.eu
  - Prepare and share reviews of VO tools
  - Offer local solutions for VOs through the group

Requirements Tracker (RT)

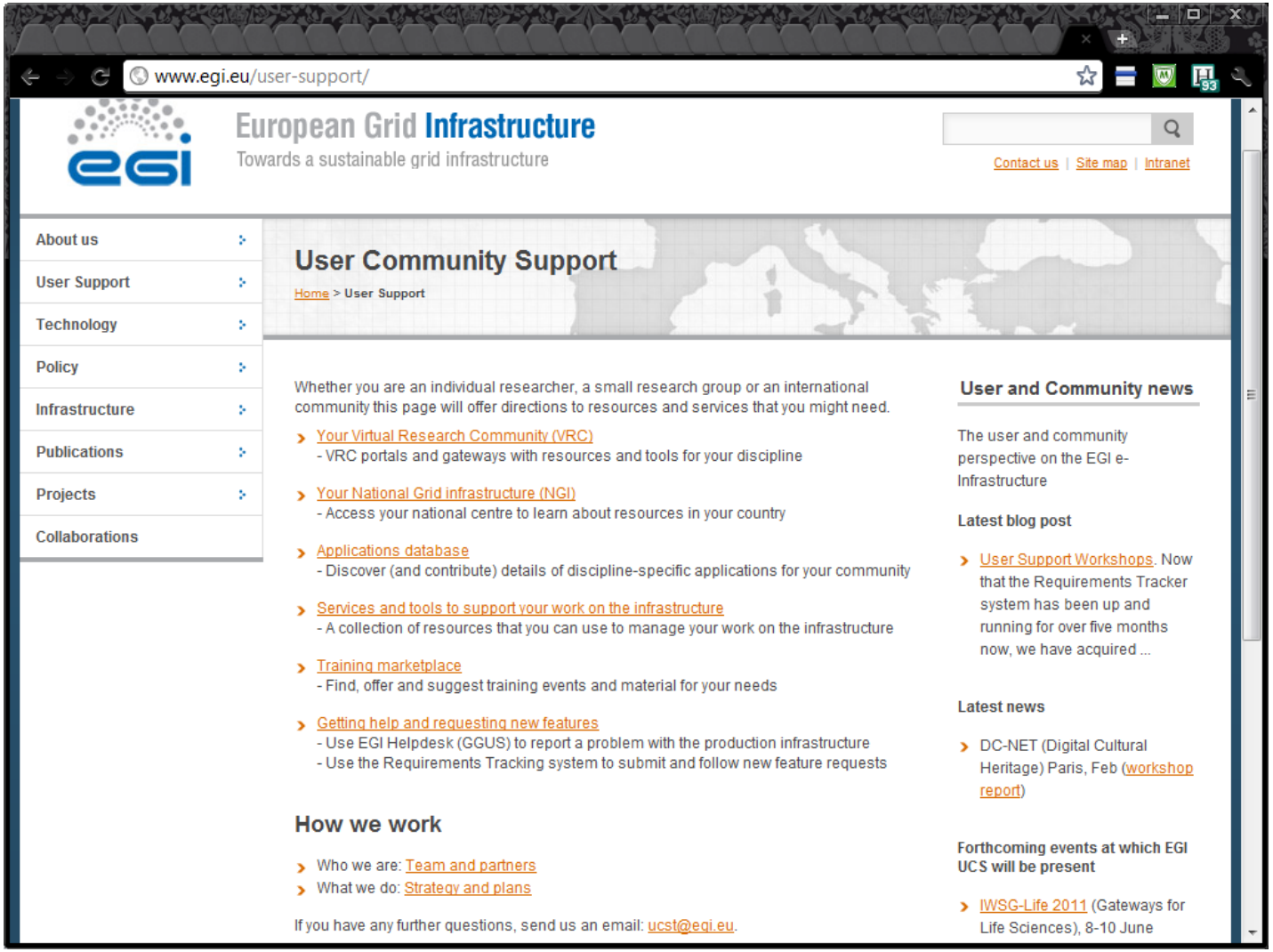
- Permanent storage for EGI requirements
  - User
  - Operational
- Complete transparency
  - For user communities
  - For NGIs
  - For technology providers
- Requirement submission interface
- Query interfaces
- Manuals available



Processing EGI requirements, implementing solutions



www.egi.eu/user-support



3) Supporting sustainable communities

Virtual Research Community

- What is a VRC?
  - Community of researchers
  - Shared technology or research interests
  - Technical implementation: Group of VOs
- Why be a VRC?
  - Effective coordination & integration with EGI
  - Build support structures within your community
- Constraints to being a VRC
  - Have supporting resources, i.e., supporting NGIs
  - Involve significant VOs, i.e., represent users

Life Sciences Grid Community

LSGC has just signed an MoU with EGI

- Application developers will be major element
- Portal and gateway developers too
- Developers are the bridge between research and the infrastructure
- Event such as IWSG-Life2011 and the HealthGrid conference bring such people together

MoU – Memorandum of Understanding

LSGC VRC

The Life Sciences community:

- Clearly well established, however…
- Many new requirements to be satisfied
- Especially to increase researcher numbers

How will this be achieved?

- Use of the existing EGI support services
- Participation in UCB
- Workshops, forums and road-shows

Life-Science Grid Community

- **Timeline**
  - June 2010: open workshop at the HealthGrid conference
  - Agreed on a statement of goals and missions
  - Monthly phone meetings
- **Participating user groups**
  - VOs: biomed (catch-all), vlemed, lsgrid
  - German life-science user community
  - Lifewatch
- **Supporting NGIs**
  - Dutch, French, Italian, Spanish and Swiss
- **Hosting body**
  - HealthGrid, a non-profit association

Purposes and goals (1)

- **Represent the LS grid users, in particular to**
  - Negotiate resources
  - Promote their requirements
  - Liaise with the European Grid Infrastructure

- **Coordinate actions**
  - Serve as a contact point for new users
  - Share expertise in the community
  - Avoid replication of efforts
  - Define common requirements
  - Encourage sharing of resources, data and tools

Purposes and goals (2)

- **Provide technical services**
  - Operate and support common VOs
  - Operate shared services
  - Provide targeted user support and application porting
- **Training and induction**
  - Organize community-specific training events
  - Smooth the learning curve, lower the start-up cost
- **Dissemination**
  - Transfer knowledge among VRC partners
  - Advertise actions
  - Liaise with other groups of interest

Applications Database

< 1 2 3 4 >

2d-ANACONDA
2d-ANAlysis of COpy Number DAta

ALMOST
almost - all atom molecular simulation toolkit

AMBER
The AMBER molecular dynamics workpackage

AMCBI
AMC e-Bioscience Infrastructure

AMIAS
A Model Independent Analysis Scheme (AMIAS) for Extracting Multipole Amplitudes.

AquaMaps
AquaMaps

ARCEM
Grid Computing Technology and the Recurrence Quantification Analysis to Predict ...

AutoDock (EUAsiaGrid)
Biomedicine / Molecular Biology - using Molecular dynamics

BiG (GISELA)
Blast in Grids

BioMD (EPIKH)
Life Sciences / Drug discovery - using Molecular dynamics

BLAST
Basic Local Alignment Search Tool

BLAST
To draw a comparison by BLAST between all sequences of prokaryotes, fungi, plant...

BLAST_in_JST
Large Scale Proteome Comparison

BM Portal
Bio Medical Portal

BTR
BIOTREE

CALnet
Combinatorial Association Logic Networks Inference

CardioGrid Portal (GISELA)
Laboratory for Complex Systems (LSC) Cardiology Portal

CD-HIT-Grid
Protein clustering on the Grid with CD-HIT

ClustalW
ClustalW

CLUSTALW
ClustalW

CLUSTER
Chaotic map clustering

CNS1.2
Crystallograph and NMR System - structure

CNS1.2-para
Crystallography and NMR system - structure

CS-Rosetta
NMR chemical shift-based structure prediction

More information

- **Contact**
  - Wiki page: http://wiki.healthgrid.org/LSVRC:Index
  - Mailing list: lsvrc@healthgrid.org

- **How to benefit from the LSGC?**
  - Established VOs
    - Discuss technical issues with other VOs
    - Submit requirements to the EGI
    - Advertise your services ; use others'
  - New communities
    - Join an existing VO and benefit from its expertise
    - Get information on related applications and porting efforts
    - Get application porting support
    - Browse training material and events

What happens next?

IWSG-Life2011

- Capture latest developments and needs

Life Sciences Grid Community

- Define new requirements; update AppDB
- Take stock of existing LS applications

EGI

- Process LSGC requirements; plan user workshops

Thanks for your time!

steve.brewer@egi.eu