Status and future of the European Grid Infrastructure after 10 years of production activities

2004-2013

Tiziana Ferrari/EGI.eu
EGI Chief Operations Officer

e-infrastructure

… a short journey through the history of international e-Infrastructures and their challenges


Efficiency, quality of service, cost effectiveness
Collection and exchange of information
Global and integrated
Collaboration and Communication
Research
e-Infrastructure

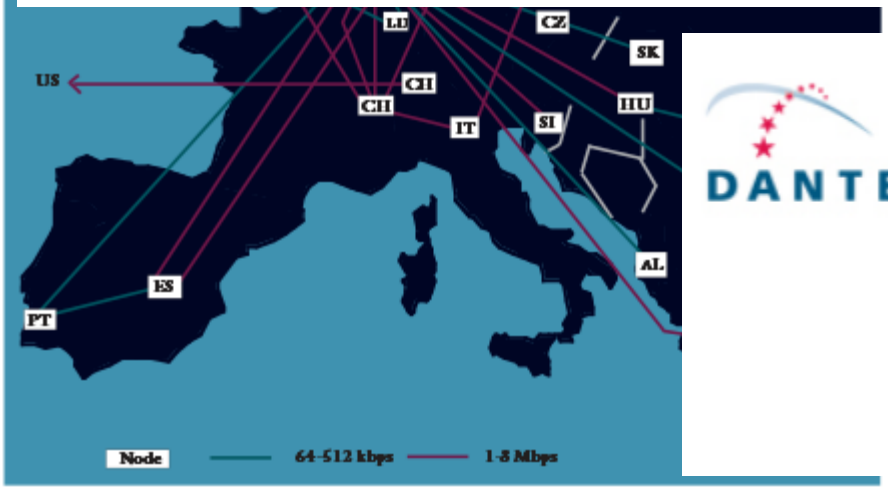
Capacity building

February 1994 - News from the Intercontinental "Front"

A contract for the provision of a 64 kbps line between EuropaNET and KREONET, the Korean research network, has been signed between DANTE and the CEC (DG-XIII). The line will be the first direct link between EU countries and an R&D network in the Pacific Rim. The aim is to stimulate the - so far - relatively weak co-operation between European and Korean researchers.

Intercontinental connectivity

45 Mbps between Europe and the US (shared with BT's commercial customers).
2 Mbps between Europe and Japan (NACSIS).



6 March | 2013

DANTE pushing towards world's first 100G transatlantic network for research and education

DANTE, representing the pan-European GÉANT network, and the America Connects to Europe project (ACE), managed by Indiana University, have launched a Prior Information Notice (PIN) to progress the implementation of the first ever 100G (gigabits per second) transatlantic links for the research and education community

When EGI was born. DataGrid

The DataGrid Project

DataGrid is a project funded by European Union. The objective is to build the next generation computing infrastructure providing intensive computation and analysis of shared large-scale databases, from hundreds of TeraBytes to PetaBytes, across widely distributed scientific communities.

January 2004:

Total number of jobs run by REGION and DATE (Excluded dteam and ops VOs)						
REGION	Jan 04	Feb 04	Mar 04	Apr 04	Total	%
CERN	13	17	2,117	1,198	3,345	9.03%
NGI_FRANCE	11	0	0	0	11	0.03%
NGI_IBERGRID	0	0	0	180	180	0.49%
NGI_IE	0	2	21	42	65	0.18%
NGI_NL	1,494	16,836	13,591	1,509	33,430	90.25%
Russia	11	0	0	0	11	0.03%
Total	1,529	16,855	15,729	2,929	37,042	
Percentage	4.13%	45.50%	42.46%	7.91%		

Deployment Status

Enabling Grids for E-science in Europe

Core Sites already integrated

Site	CPU	
CERN (CH)	324	
FZK (D)	144	
PIC (E)	160	
FNAL (US)	4	
CNAF(I)	715	
Nikhef (NL)	250	
Taipei (AP)	98	
RAL (UK)	146	
Total	**1841**	

With the other sites (currently running LCG-1), the expected capacity will exceed the previsions foreseen for 2004:

around 4000 CPUs at about 30 sites

EGI Capacity building 2004-2013

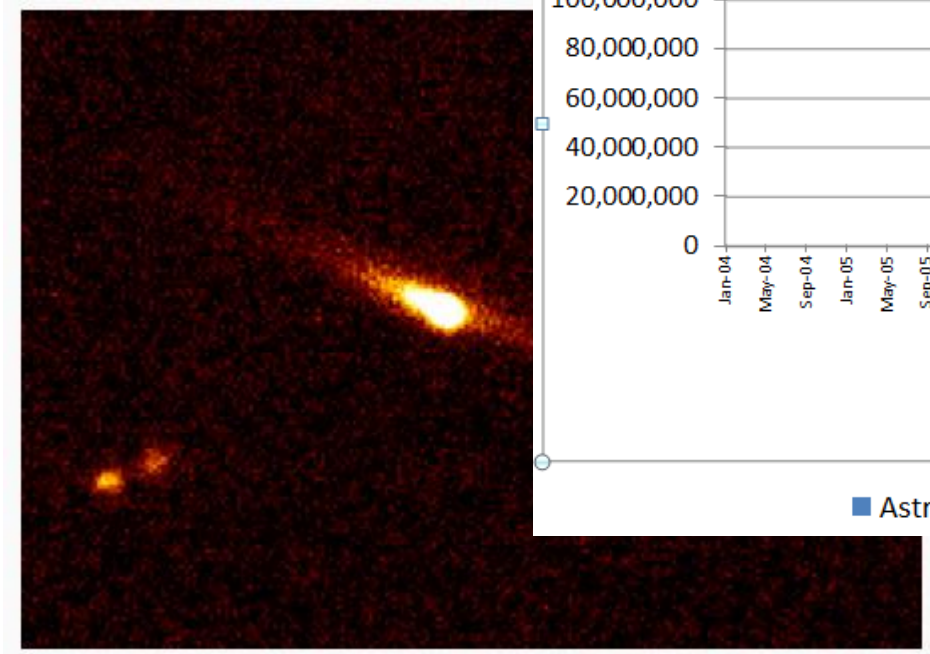
2004

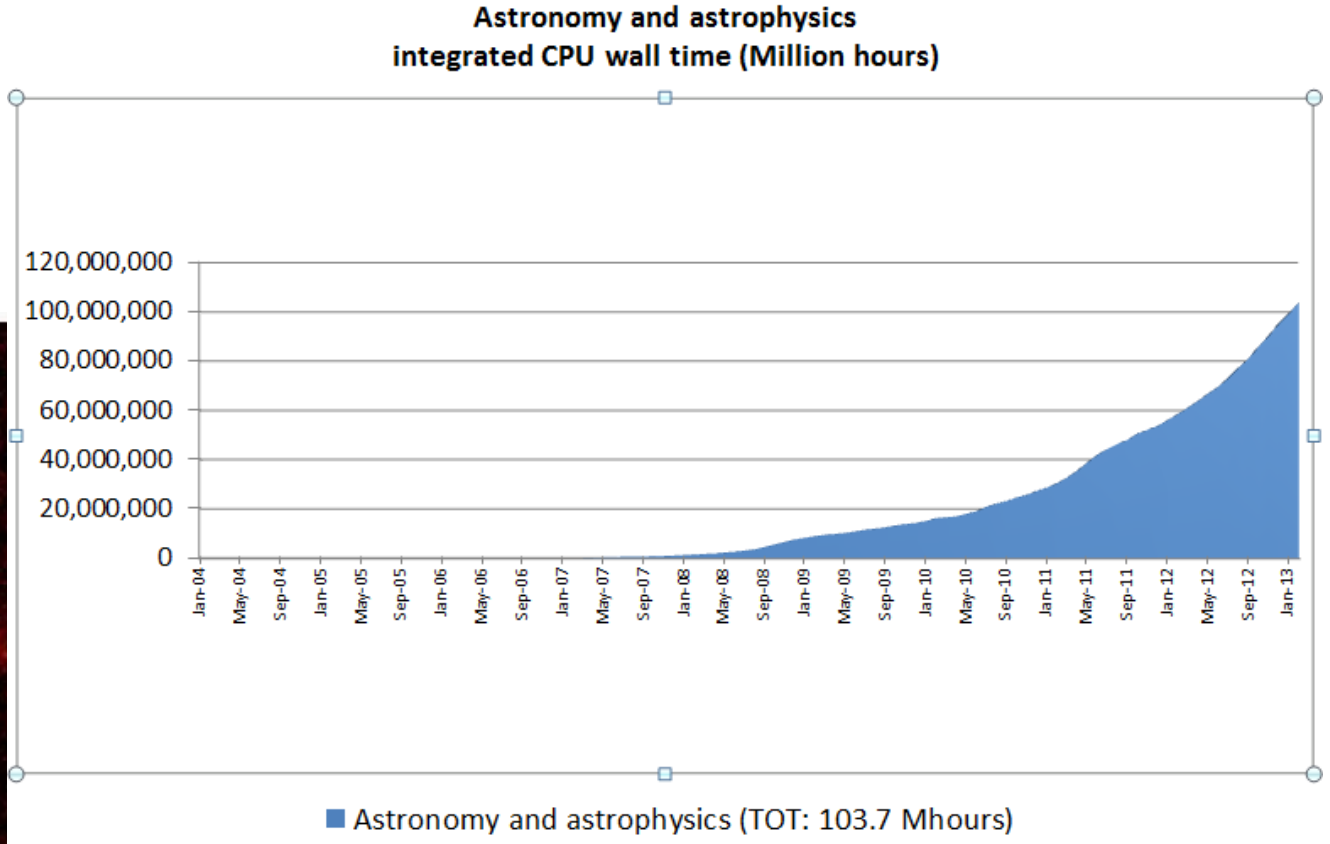
Sites 2004	
CERN	1
NGI_NL	2
Sweden	2
NGI_BG	1
Canada	1
NGI_FR	1
NGI_DE	3
Russia	1
NGI_UK	1
NGI_GR	1
NGI_FI	1
Taiwan	1

2013

Metrics March 2013		Value
Capacity	CPU cores (EGI and integrated resource providers) Disk/Tape (PB) Countries	372,612 (315 resource centres) 180/167 56
Jobs	Average Job/day (Million)	1.67 (2.25 including local computation)
% of total norm. CPU wall time consumed	High-Energy Physics	93.78%
	Astronomy and Astrophysics	2.78 %
	Life Sciences	1.31%
	Remaining disciplines	2.13%
CPU wall time	Integrated, Billion hours Jan 2004-Mar 2013	4.8 **36.8 (normalized HEP-SPC06)**

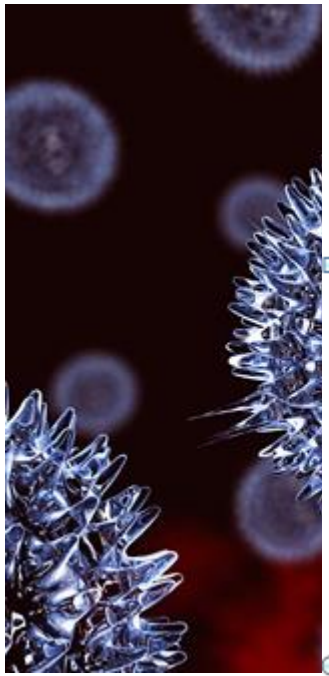
Supporting research
Astronomy and Astrophysics

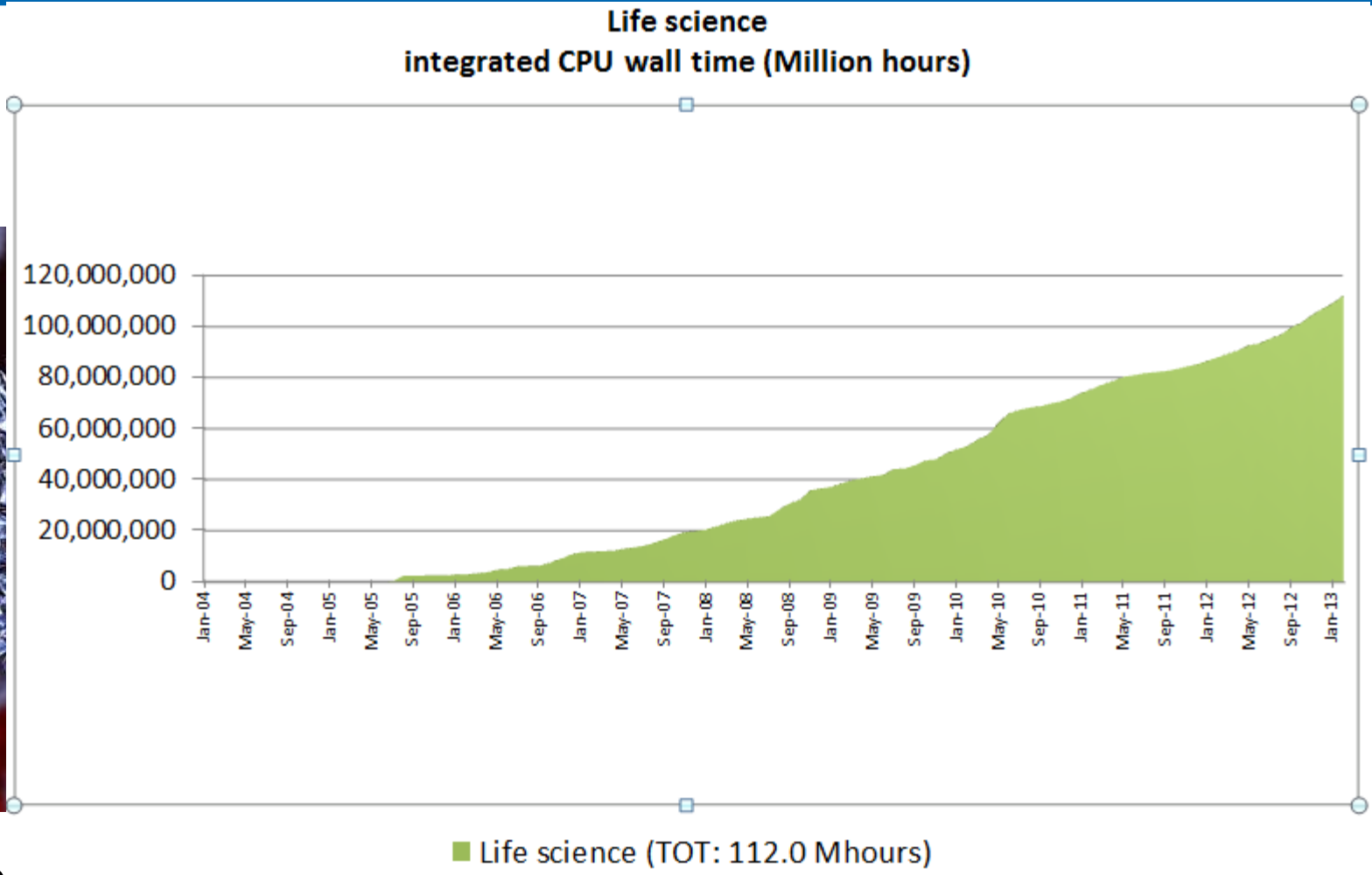




Origin of main-belt comets
http://www.egi.eu/case-studies/main-belt_comets.html

Supporting research
Life Science

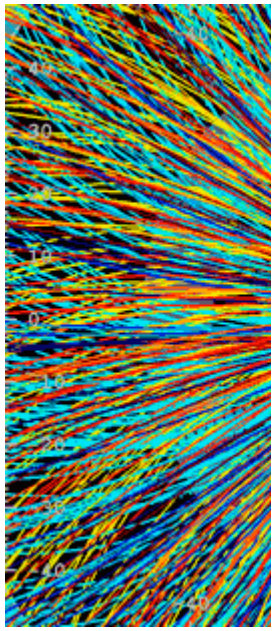


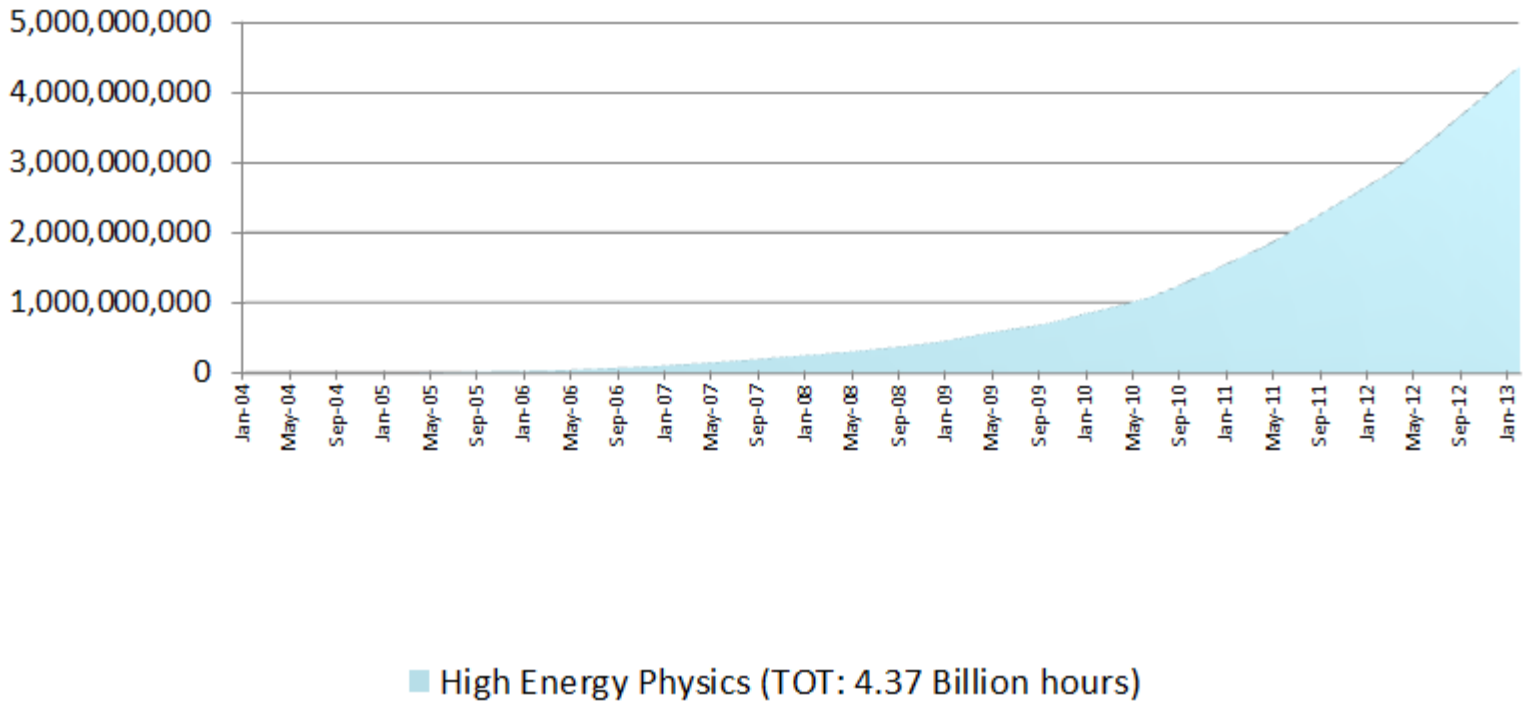


Hunting fo. viruses

http://www.egi.eu/cms/case-studies/hunting_for_viruses.html

Supporting research
High Energy Physics

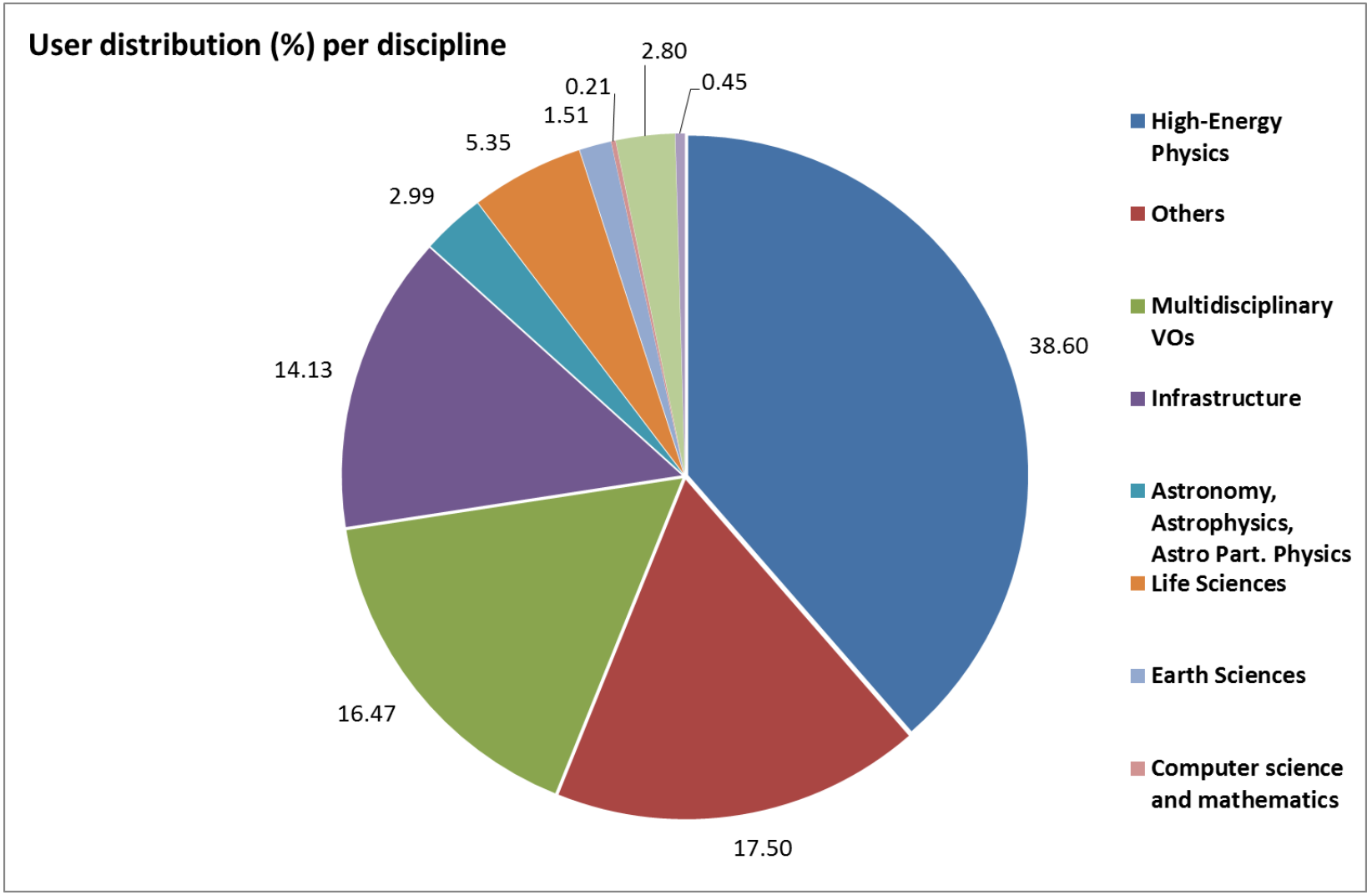




Large Hadron Collider

http://www.egi.eu/cms/case-studies/WLCG.html

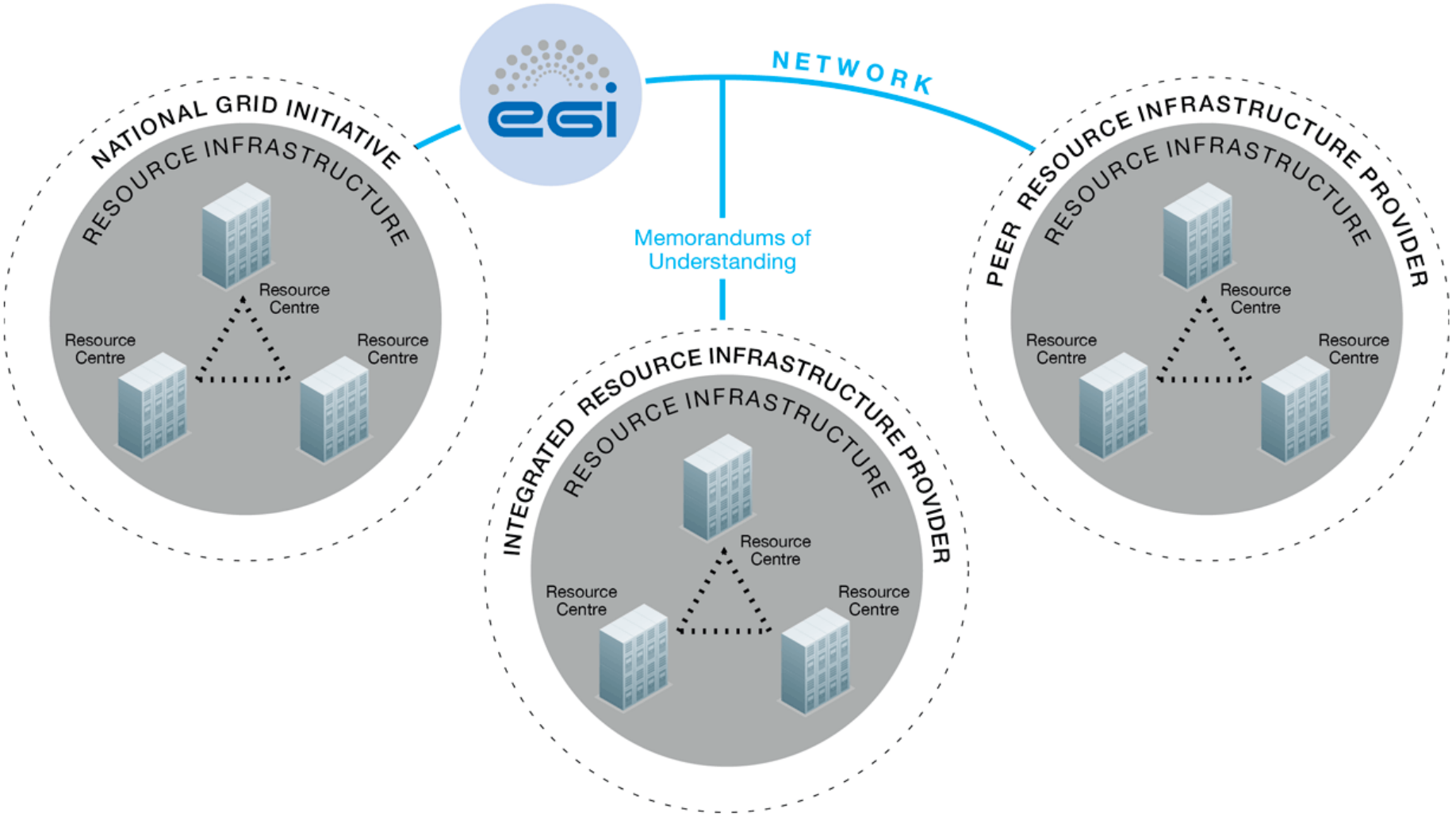
… and many other sciences



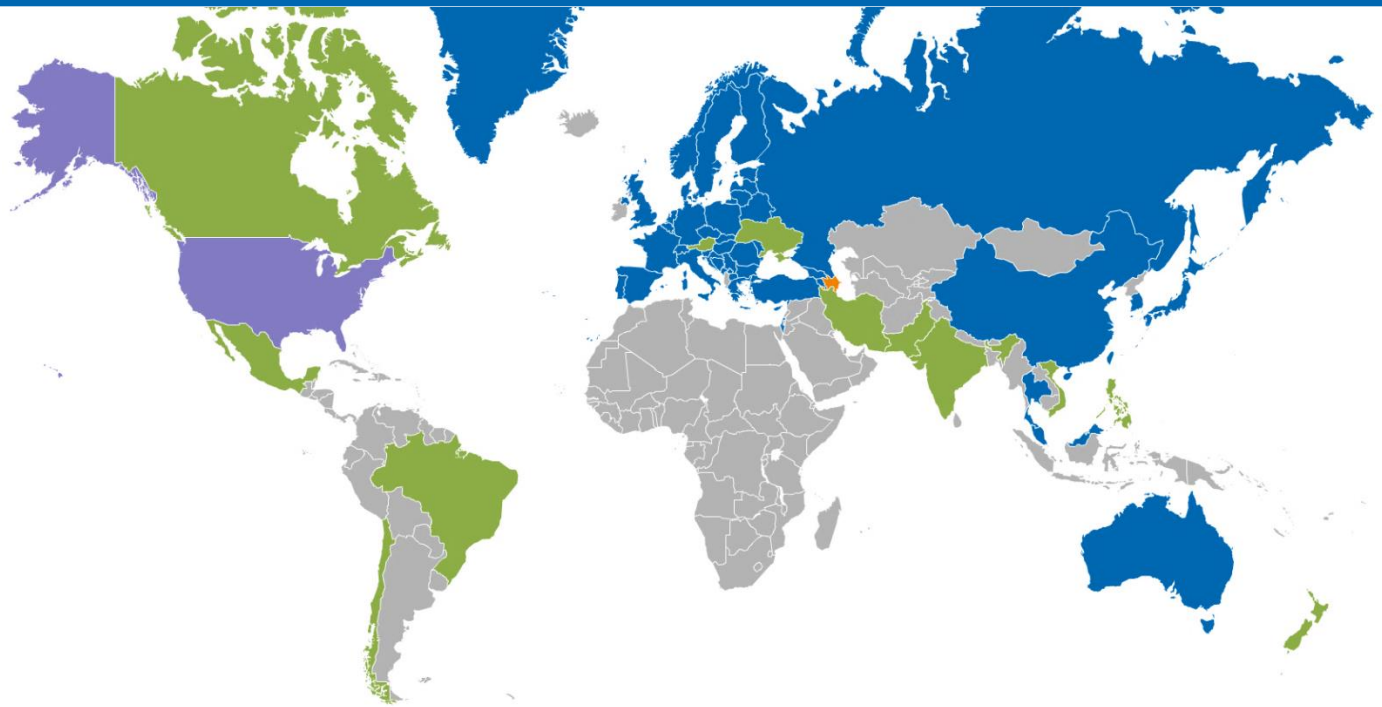
Global and integrated 1/3

- Integrated infrastructure providers → sharing policies, procedures, tools, QoS agreements and part of the same operations structure
  - Members of the EGI collaboration (EGI Council/EGI-InSPIRE)
  - External providers
    - Austria and Ukraine, Latin America, Canada, Asia Pacific
- Peer infrastructure providers → own operations tools and procedures, compatible policies, loose operations collaboration with EGI
  - Open Science Grid

Global and integrated 2/3



Global and integrated 3/3



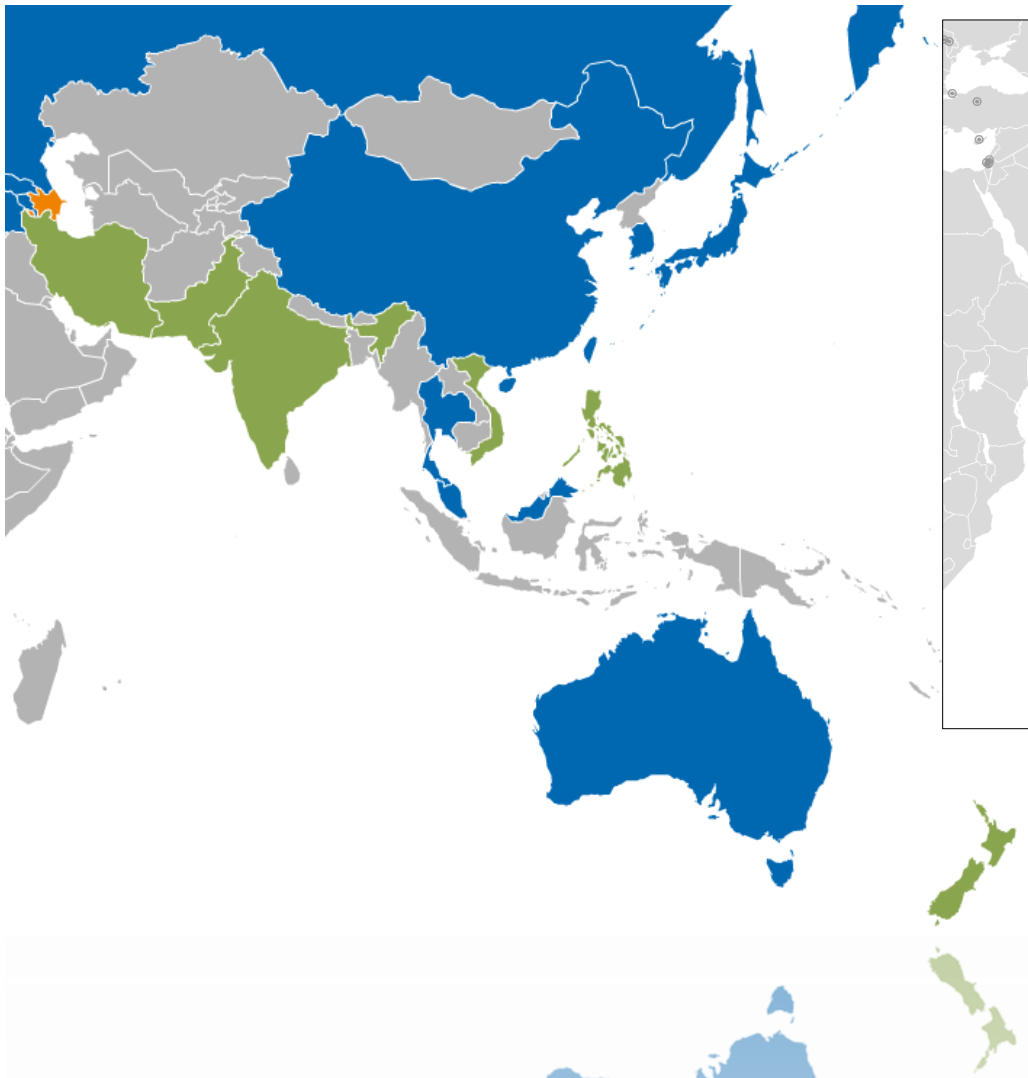
Integrated EGI-InSPIRE Partners and EGI Council Members

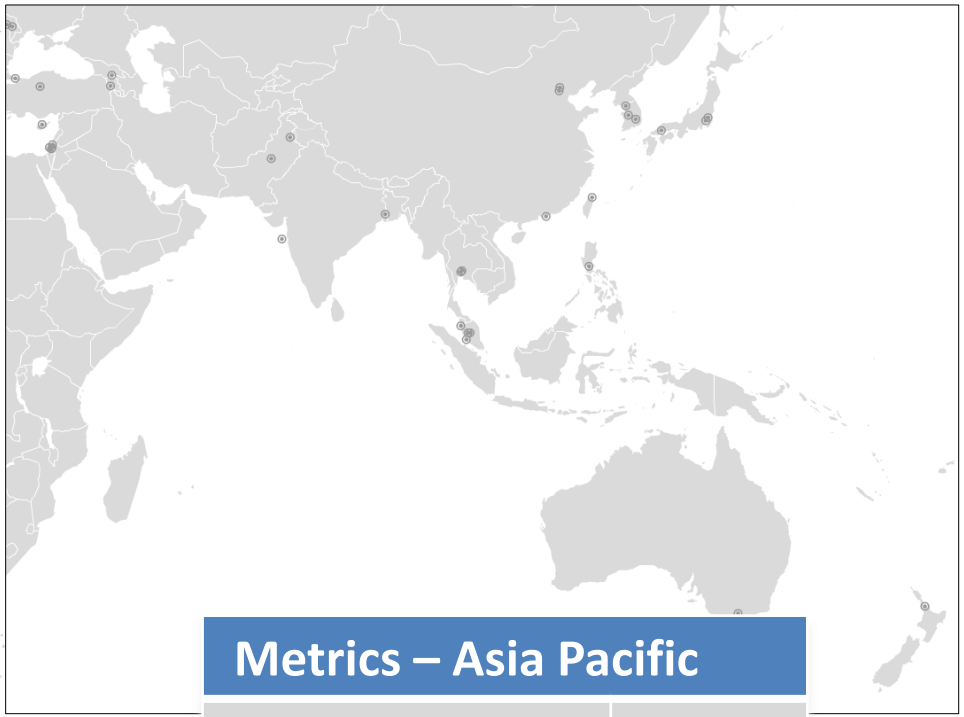
Internal/External RPs being integrated

External RP

Peer RP

Asia Pacific region





Metrics – Asia Pacific	
Countries	13
Sites	32
Integrated CPU wall time (Million hours), 2004-2013	143.0

Integration: EGI service platforms

- Core platform
  - to enable the federation of resources
- Community platform
  - access computing resources
  - access and manage data and storage
- Cloud platform
  - access cloud resources

THE platform - 2005

SA1: grid operations, mgmt & support

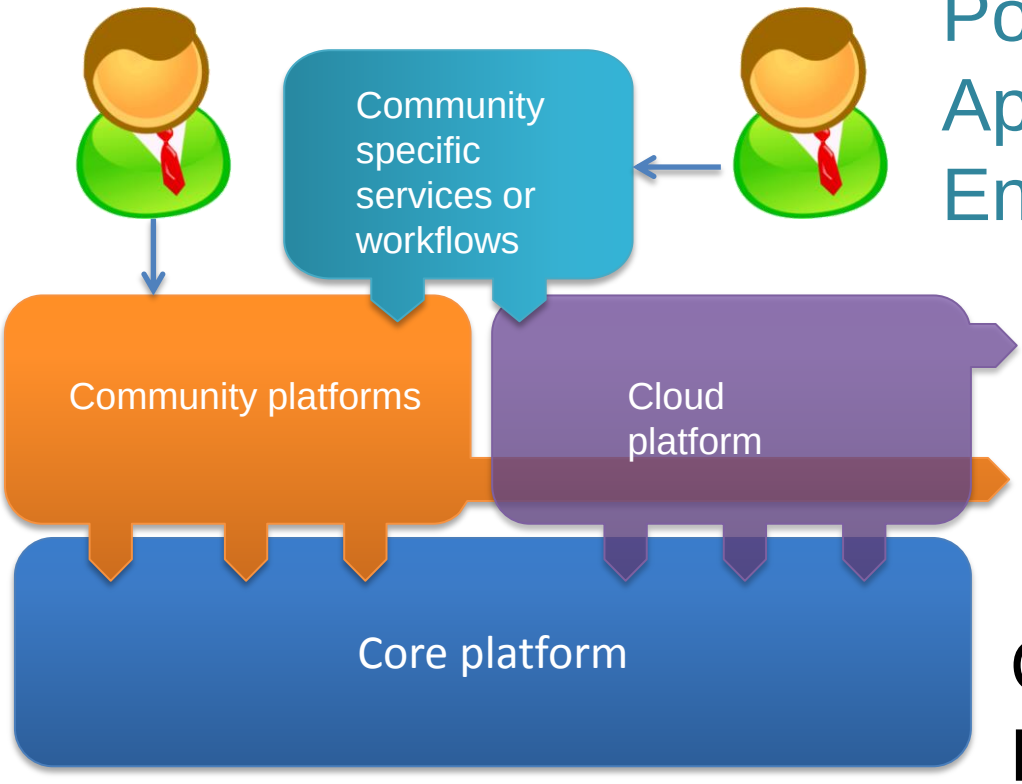
Enabling Grids for E-sciencE

- **Scale of the production service**
  - ***Den Haag***: ~8k CPUs/80 sites ***Athens***: ~14k CPUs/130 sites
    This greatly exceeds the no. sites planned for the end of EGEE
  - Continuous improvements to LCG-2 middleware
    LCG-2.4 recently released

- **Set-up of CIC/ROCs**
  - Rotating operational responsibility now in place
- **Inter-operability increasing with other grids (Grid3/OSG)**
- **gLite certification now a priority**
  - Setting-up pre-production service for end-users

EGI platforms - 2013

ARC,
dCache
gLite
Globus
QCG
Unicore



Portals, gateways,
Applications, wflow
Engines

Compute, Storage,
Data archives,
Desktop Grids,
HPC

Monitoring, Accounting,
EGI Helpdesk
Service registry and discovery

Last step in our journey…

Governance
- EGI.eu
- NGIs/EIROs

Sustainability
- Operations
- Innovation

EGI as vehicle for Innovation

- Deploying Technology Innovation
  - Distributed Computing continues to evolve
    - To include: Grids, Desktops, Virtualisation, Clouds, …
- Enabling Software Innovation
  - Provide reliable persistent community platform
- Supporting Research Innovation in Europe
  - Infrastructure for data driven research
    - Support for international research (e.g. ESFRI)

Service Evolution

- ‘Old’ Model
  - Isolated Technology Platforms
  - High Throughput Computing
- ‘Current’ Model
  - Integrated Technology Platforms
  - HPC, HTC, Data
- ‘Future’ Model
  - Federated Cloud Platform
  - Community Platforms

Towards Horizon2020

- EC projects to support innovation
  - E-Infrastructure integration
  - Virtual research environments
  - Collaboration and human capital
- Community funding
  - EGI.eu Core platform
  - NGI Core platform, operations
- Exploring
  - Resource allocation on-demand
  - Pay per use models