http://www.grnet.gr

“eInfrastructure Development: from Regional to Intercontinental Collaborative Research”

Dr. Ognjen Prnjat
European and Regional eInfrastructure management
Greek Research and Technology Network

eResearch 2103 Conference, Cape Town

The case I am arguing

- Resource sharing (over different technologies) facilitates eResearch
- Regional collaboration as a vehicle for sustainable development

Outline

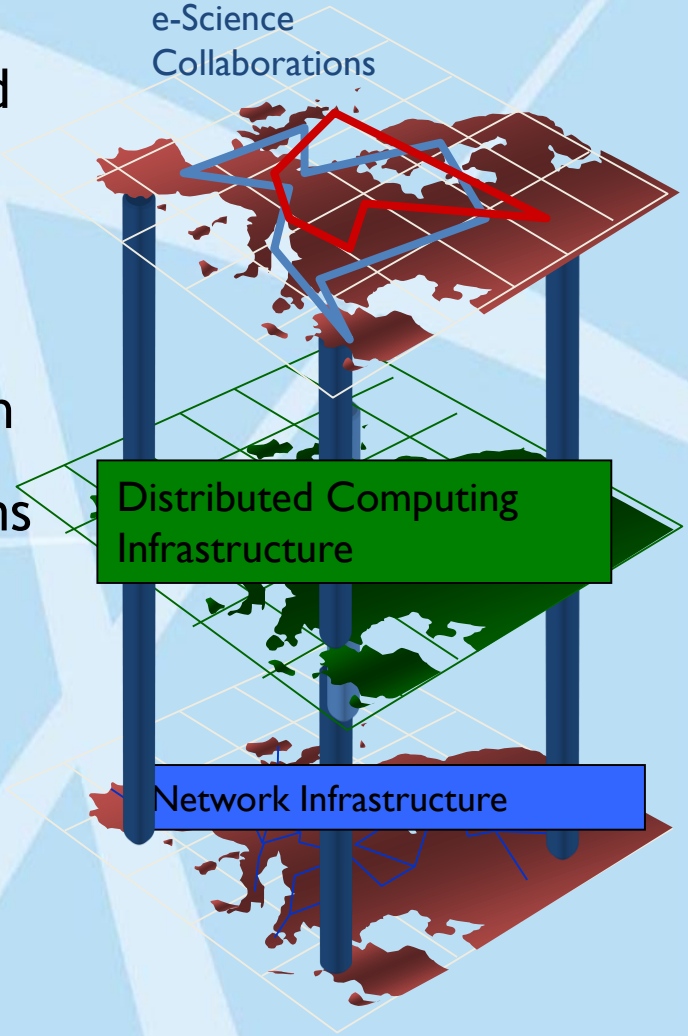
- State of the art in eInfrastructures in Europe
- Regional collaboration models in networking, Grid computing and HPC – a case for international collaboration in Africa
- CHAIN-REDS: worldwide collaborations
- CHAIN-REDS: Intercontinental Grids and HPC collaboration opportunities

Outline

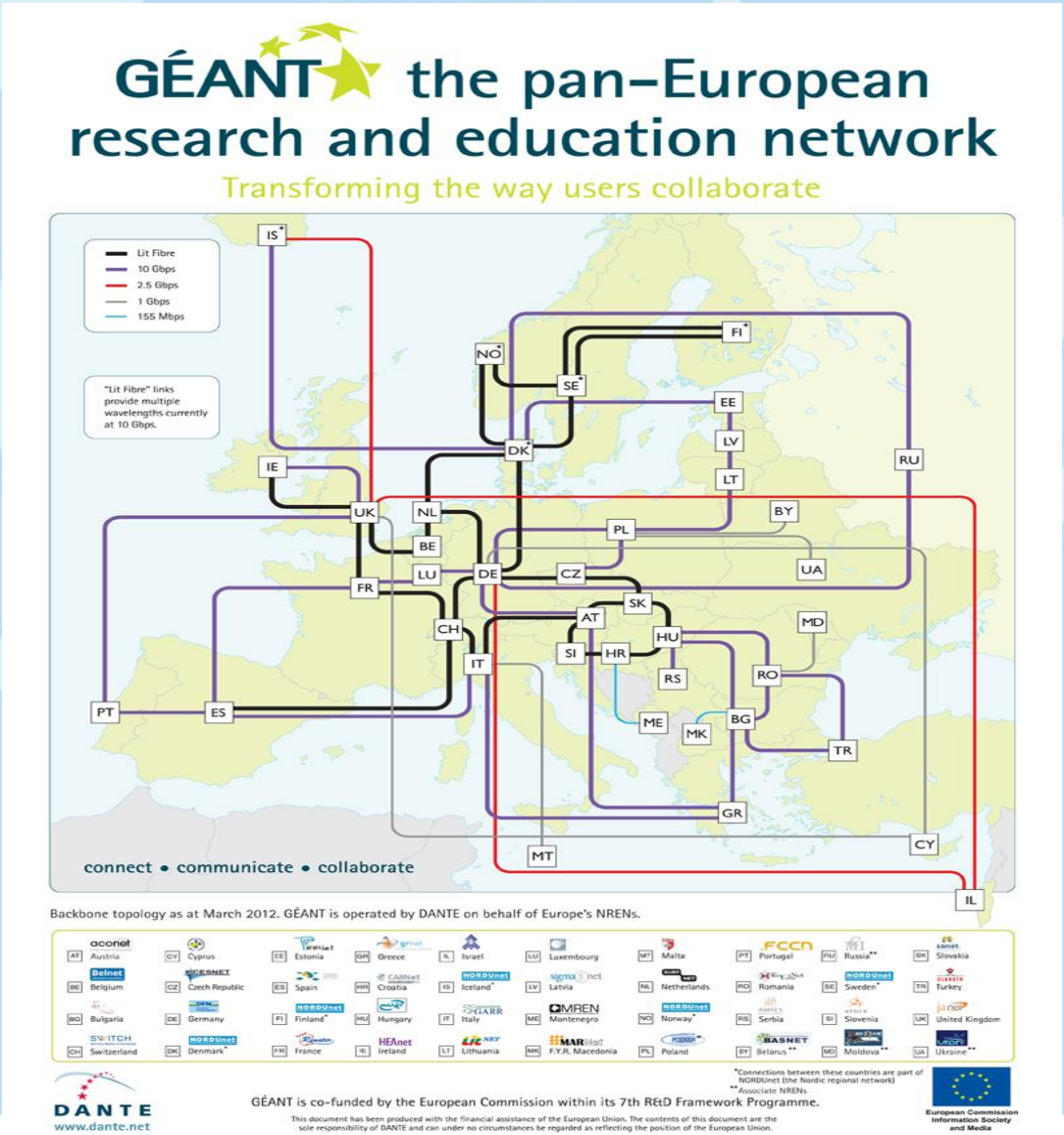
- **State of the art in eInfrastructures in Europe**
- Regional collaboration models in networking, Grid computing and HPC – a case for international collaboration in Africa
- CHAIN-REDS: worldwide collaborations
- CHAIN-REDS: Intercontinental Grids and HPC collaboration opportunities

Pan-EU e-Infrastructures vision

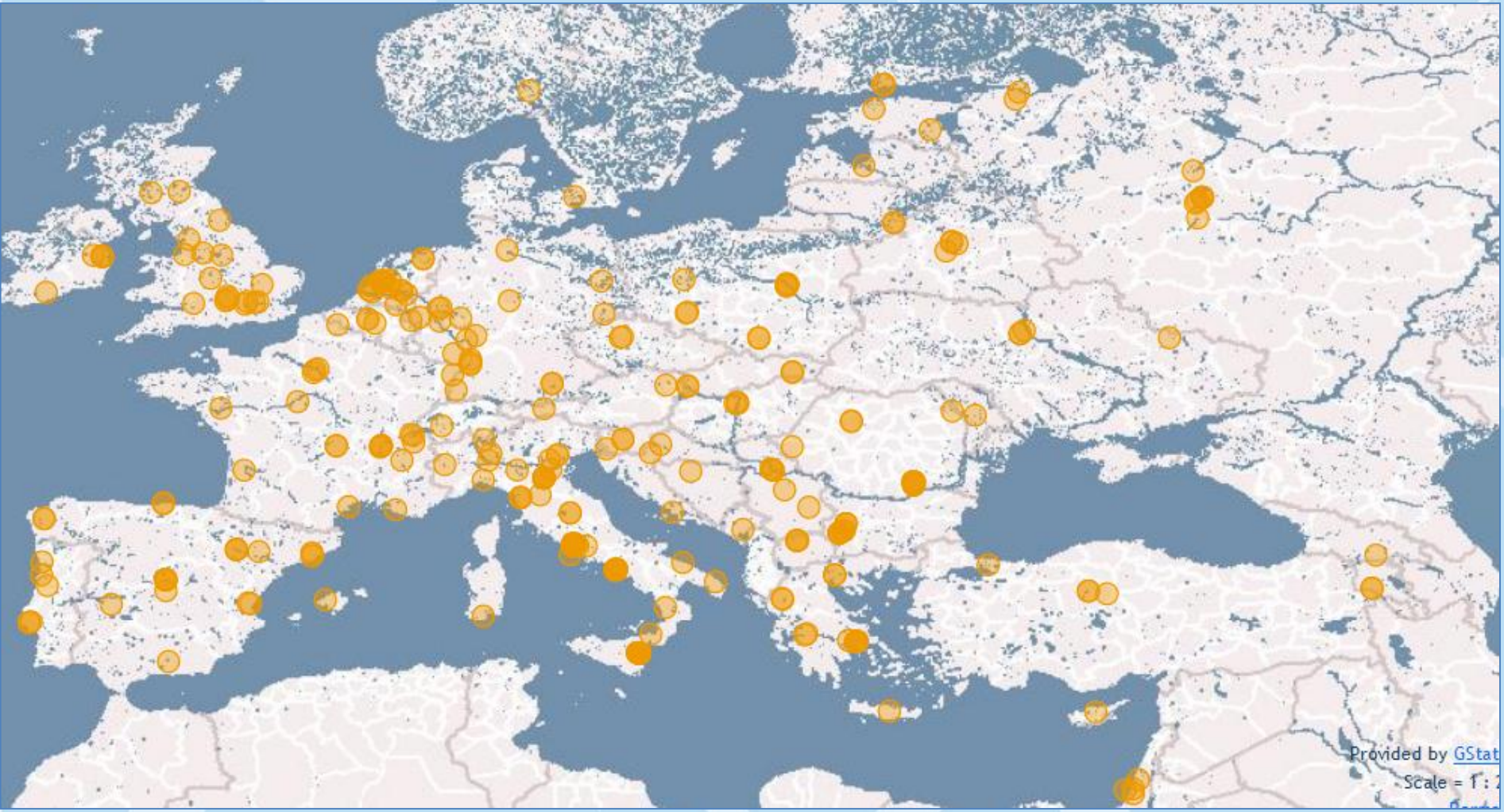
- The Research **Network infrastructure** provides fast interconnection and advanced services among Research and Education institutes of different countries'
- The Research **Distributed Computing Infrastructure** provides a distributed environment for sharing computing power, storage, instruments and databases through the appropriate software (middleware) in order to solve complex application problems
- This integrated networking & DCI environment is called **electronic infrastructure (eInfrastructure)** allowing new methods of global collaborative research - often referred to as **electronic science (eScience)**
- The creation of the eInfrastructure is a key objective of the **European Research Area**



Network: GEANT



Grid: European Grid Infrastructure

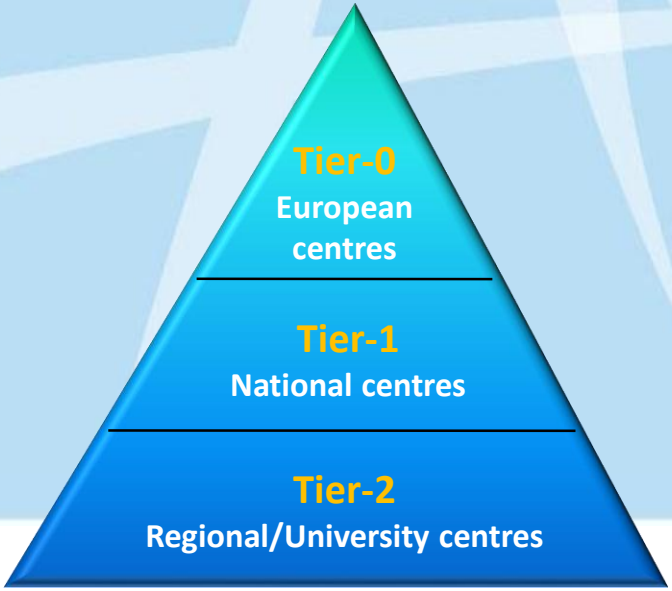


Grid: European Grid Infrastructure

Metrics		Value
Capacity	CPU cores (EGI and integrated resource providers) Disk/Tape (PB)	372,612 (315 resource production centres) 180/167
CPU wall clock time	Total normalized CPU wall clock time consumed – grid (Billion HEP-SPEC 06 hours)	15.5 (02-2012/02-2013) 31.4 (01-2010/02-2013)
Jobs	Job/year (Million)	528.4
	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%

High-Performance Computing: PRACE

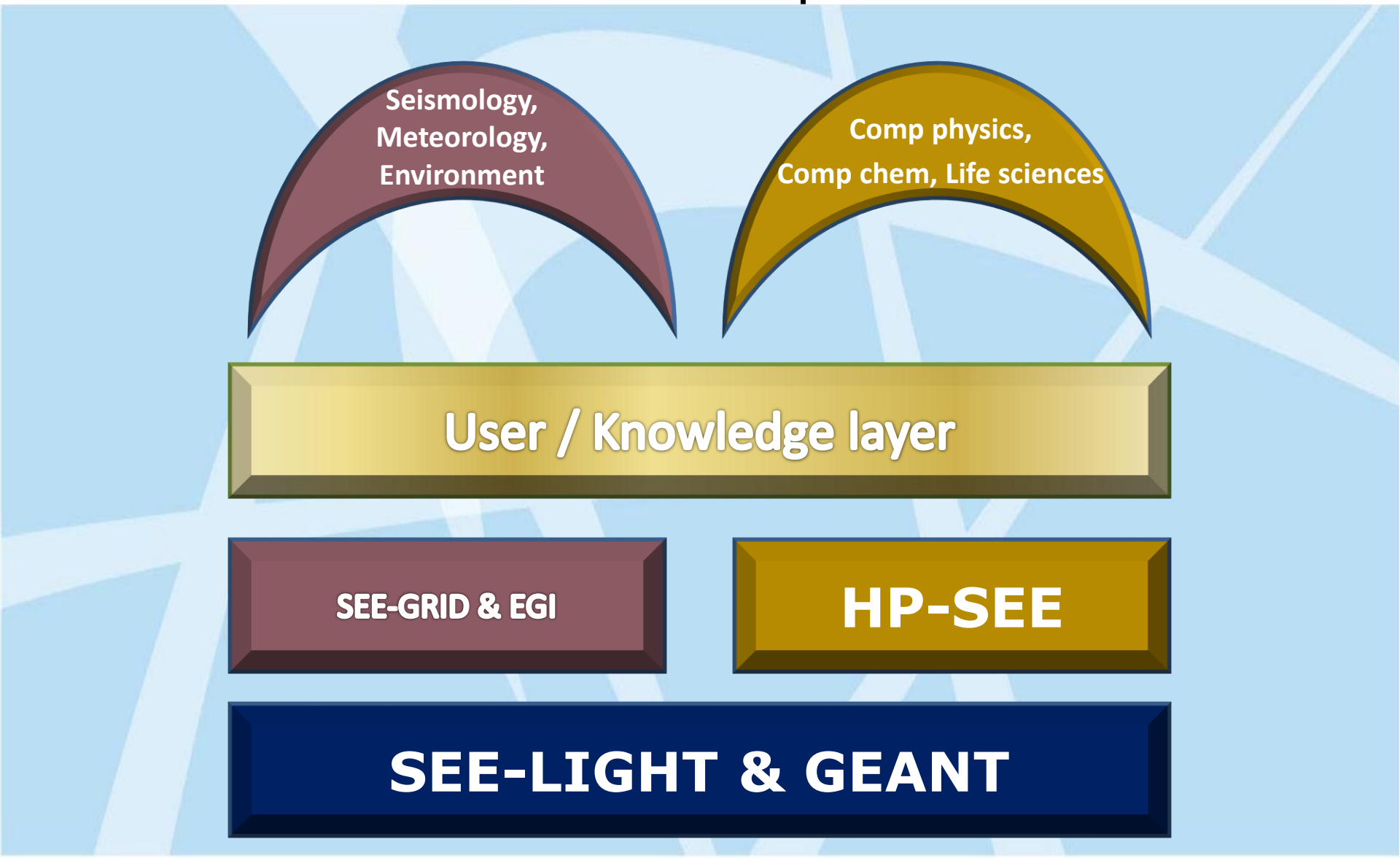
- 24 countries
- 5 Tier0 systems currently deployed
- ~9 Pflops total power (Tier-0)
- 15+ Tier 1 systems
- PRACE PP + phases of implementation projects, total EC investment ~ 70MEuro
- Access by European peer review



Outline

- State of the art in eInfrastructures in Europe
- **Regional collaboration models in networking, Grid computing and HPC - a case for international collaboration in Africa**
- CHAIN-REDS: overview
- CHAIN-REDS: Intercontinental Grids and HPC collaboration opportunities

The model - eInfrastructure for South-East Europe



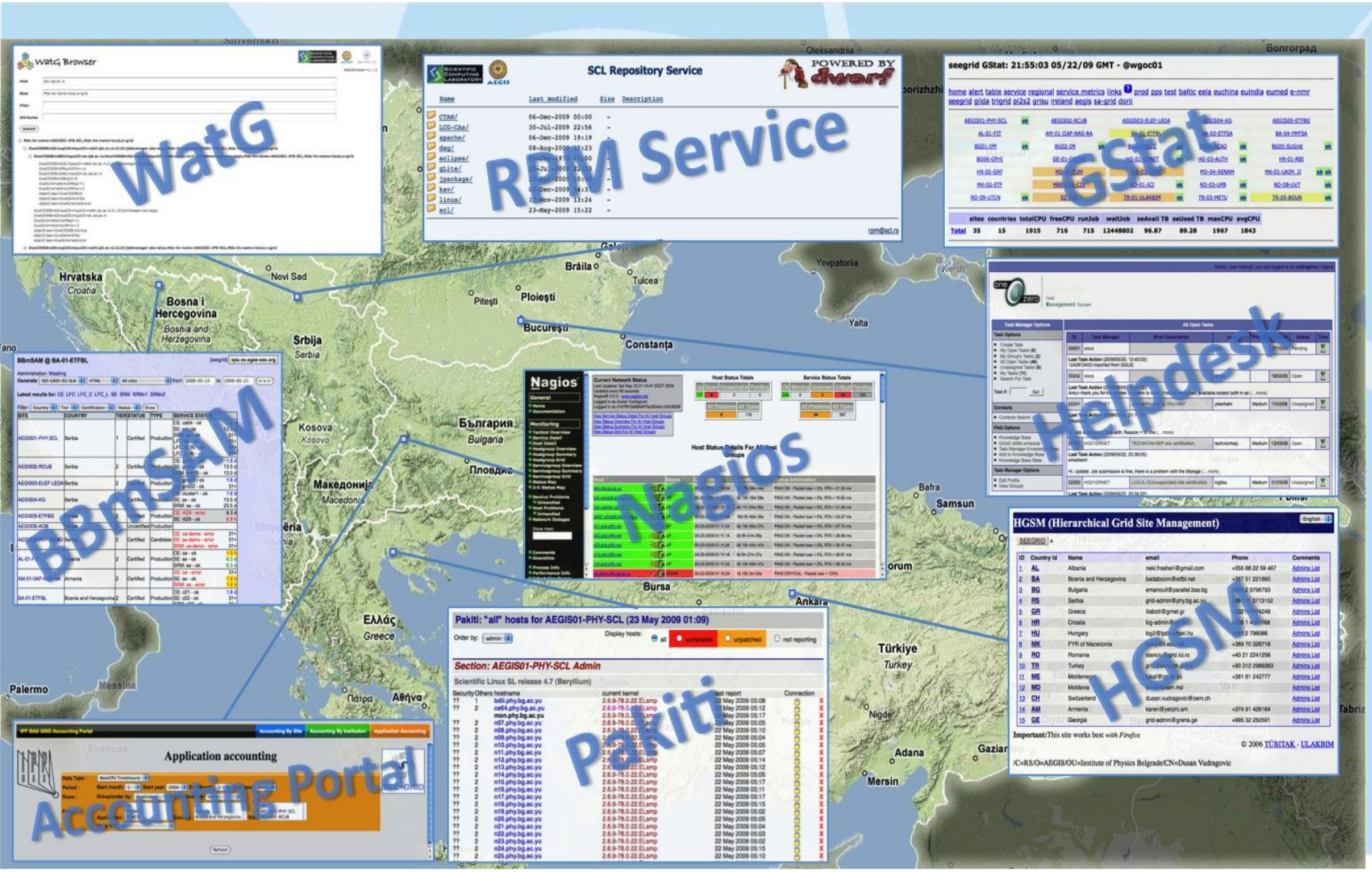
Network: SEELIGHT

- SEEREN projects set up regional NREN connectivity and GEANT links
- SEE-LIGHT: South-East European Lambda Network Facility for R&E
- Deployment of an advanced regional network infrastructure, fibres and equipment
- Under the Hellenic Plan for the Economic Reconstruction of the Balkans – HiPERB (80-20)
- Serbia implementation stage, Bulgaria tender stage, Romania on own funds, FYR of Macedonia ongoing
- SEENet: a management body for SEELIGHT

Grid: the SEE-GRID project series

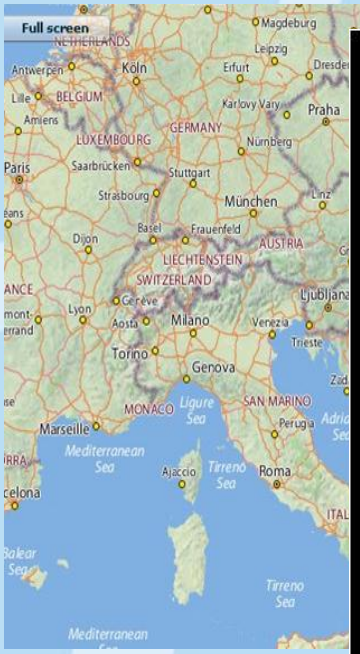
- 55 sites in production
- Dedicated CPUs: 3k, 6k total
- Dedicated storage: 745 TB

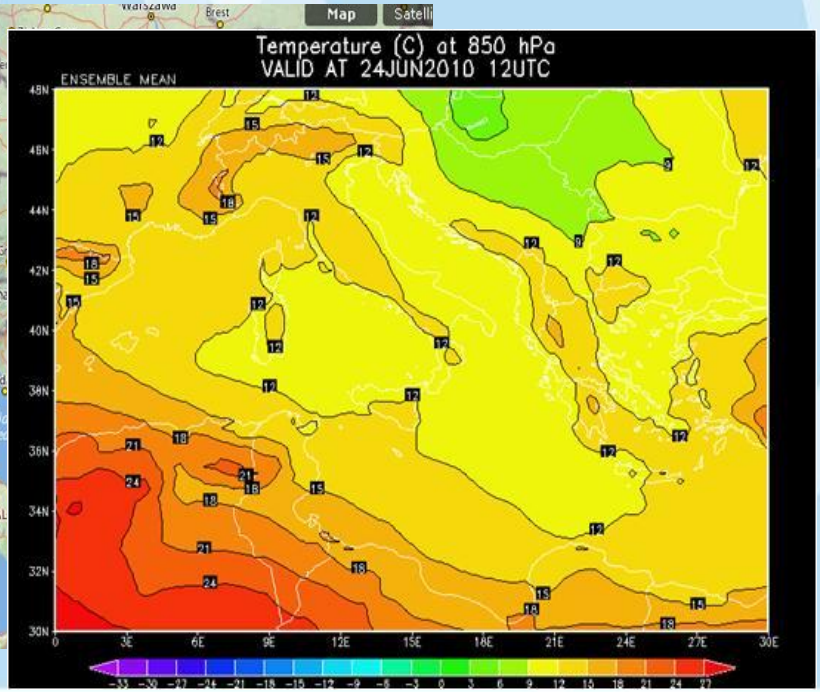
Grid: Distributed management

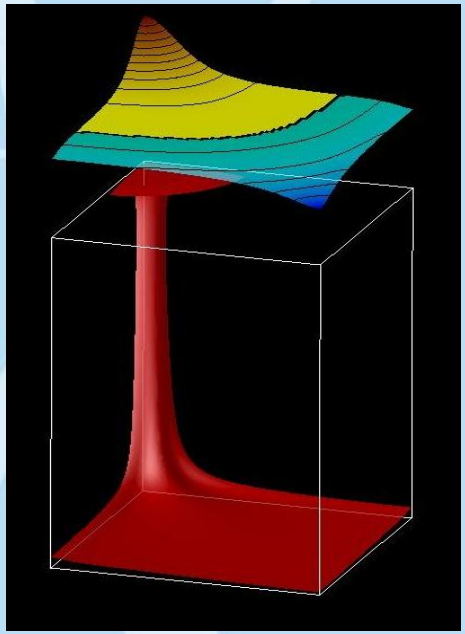


Grid: User communities support

- 22 initial applications fro various domains
- Further focus: seismology (6 major applications), meteorology (2) and environmental protection (8)
- Cross-border user communities and beneficiaries
- Clear and efficient procedures for support



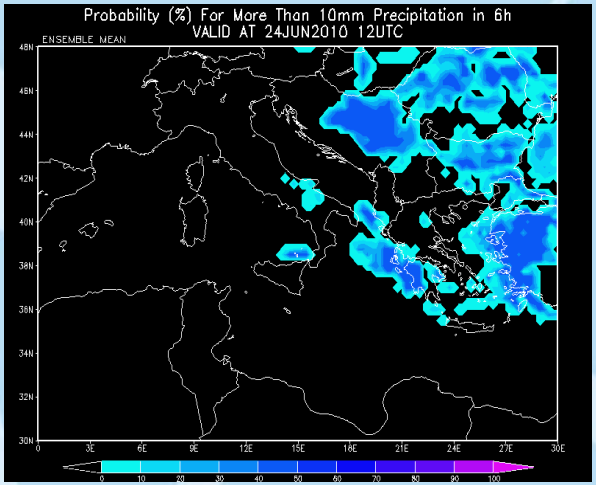


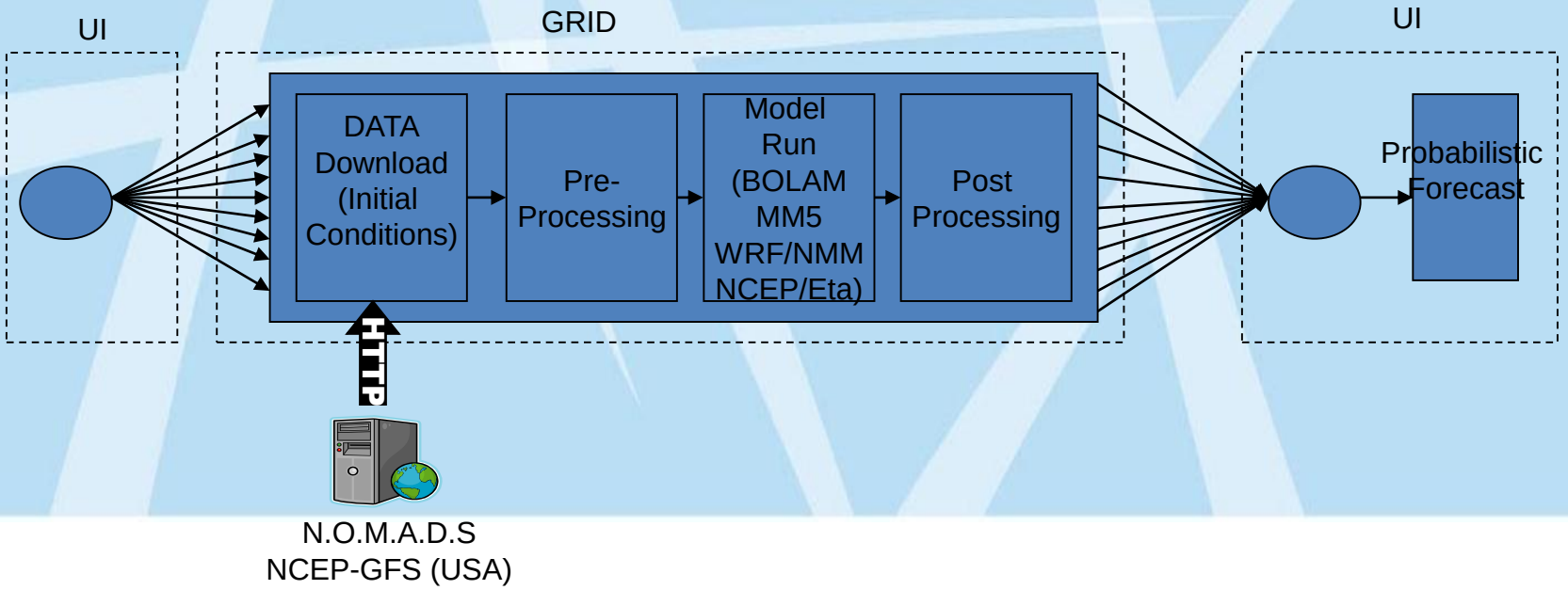


Grid: Meteorology

Regional Multi-model, Multi-analysis Ensemble Prediction System

- BOLAM, MM5, NCEP/Eta, NCEP/WRF-NMM
- SEE-wide scale detailed forecasts
- Coordinate, collect and analyze outputs of all models to generate of probabilistic forecasts
- Complex and CPU-intensive





Grid: the SEE-GRID series

- Regional infrastructure and operations built through 3 projects
- User community buy-in secured
- National structuring via NGIs
- All countries in European Grid Initiative
- Key to success: distributing operations and supporting cross-border communities; joint lobbying strategies

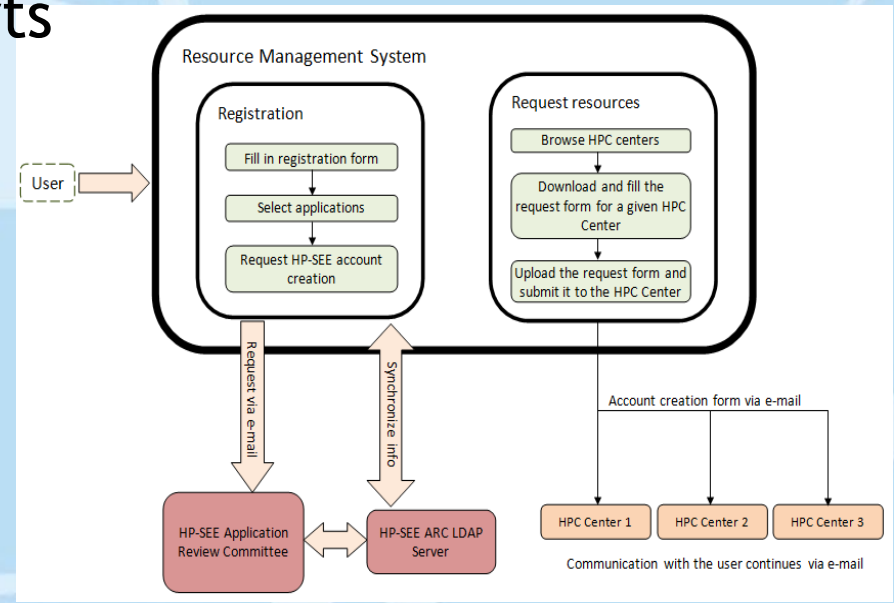
HPC: HP-SEE project

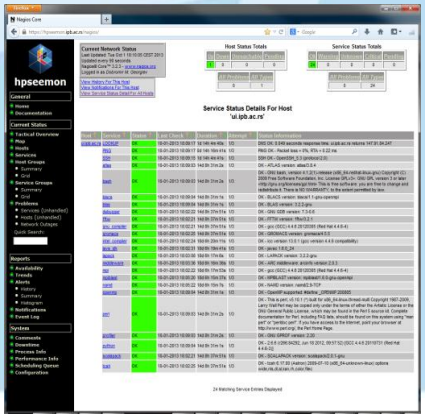
- Total peak performance in double precision >150 Tflops CPUs; the single precision performance by GPUs is >150 Tflops.
- Heterogeneous infrastructure - 2 Blue Gene/P SCs, several HPC clusters with advanced interconnects. Advantage to users
- Substantial number of libraries, toolkits and application software were deployed, tested and optimised

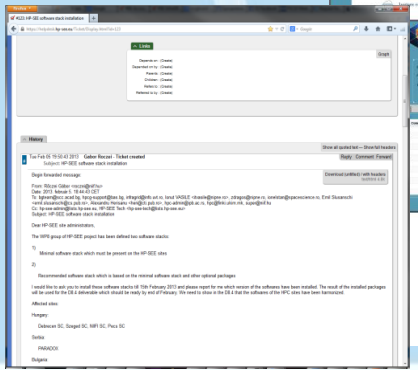
HPC: Distributed management

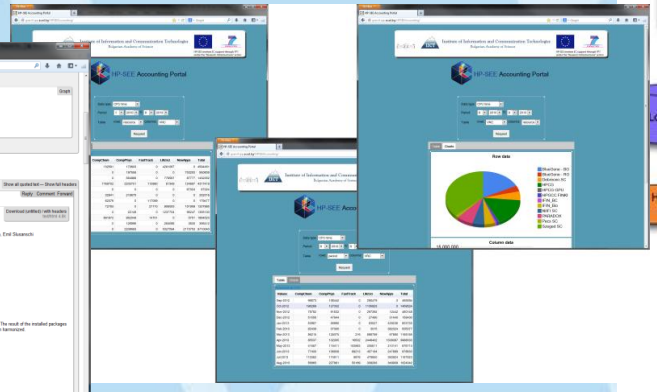
Distributed set of services supports infrastructure operations

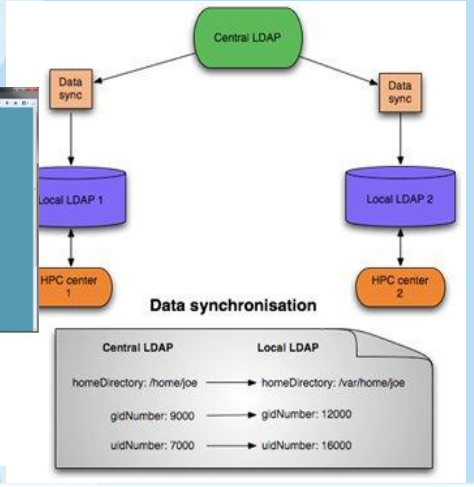
- AAA framework
  - Resource Management System
  - ARC-LDAP service
  - Accounting System
- Helpdesk (Request Tracker based)
- Monitoring (Nagios + local tools)



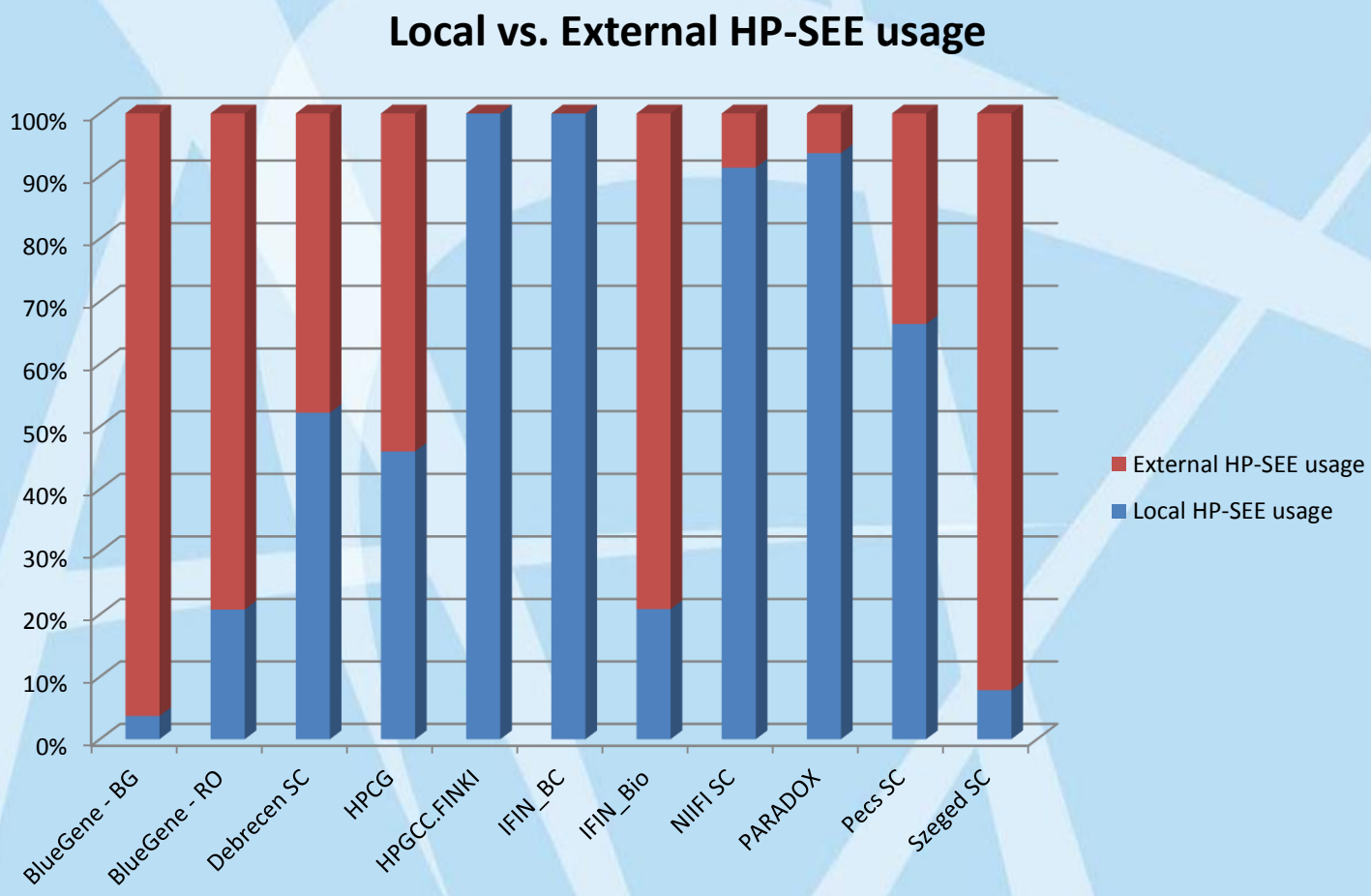








HPC: Trans-national usage



HPC: User communities

- Number of registered users- 200
- Total number of scientific publications >200

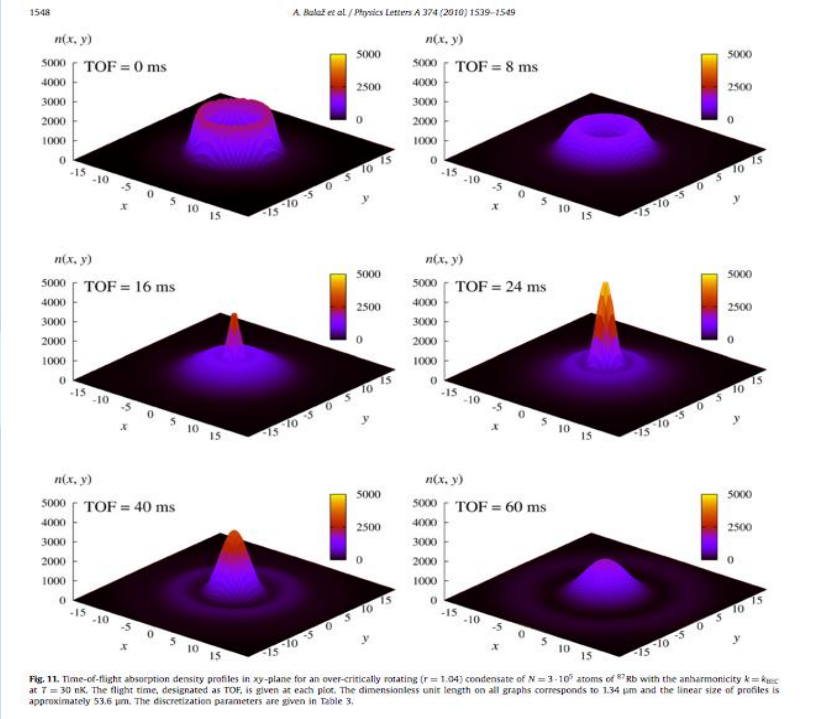
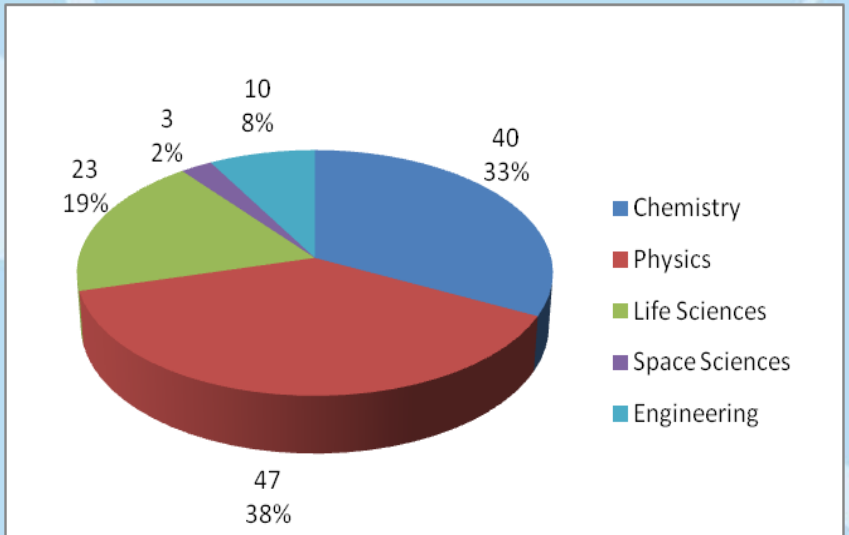
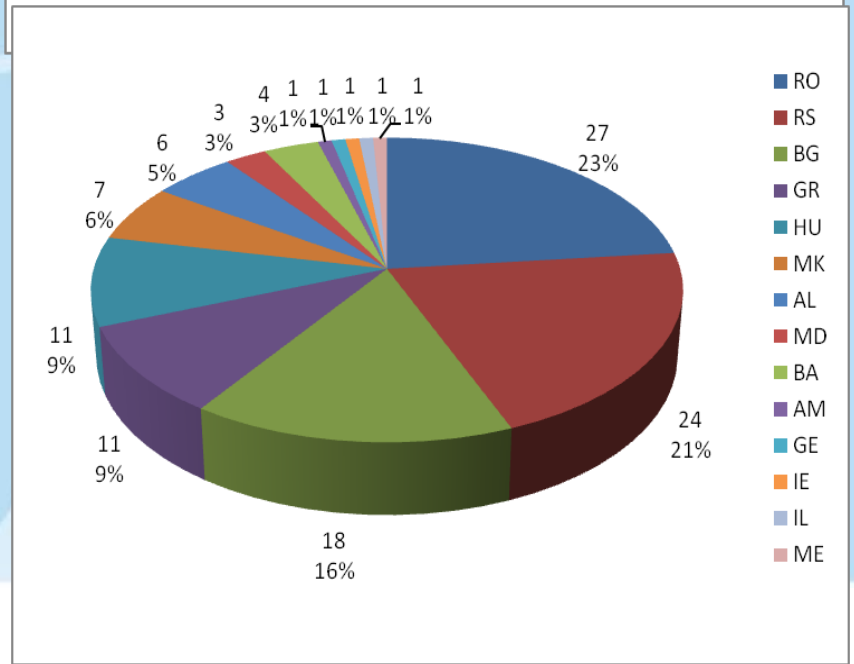


Fig. 11. Time-of-flight absorption density profiles in xy-plane for an over-critically rotating ($\tau = 1.04$) condensate of $N = 3 \cdot 10^5$ atoms of ^{87}Rb with the anharmonicity $k = k_{\mathrm{REC}}$ at $T = 30$ nK. The flight time, designated as TOF, is given at each plot. The dimensionless unit length on all graphs corresponds to 1.34 μm and the linear size of profiles is approximately 53.6 μm. The discretization parameters are given in Table 3.





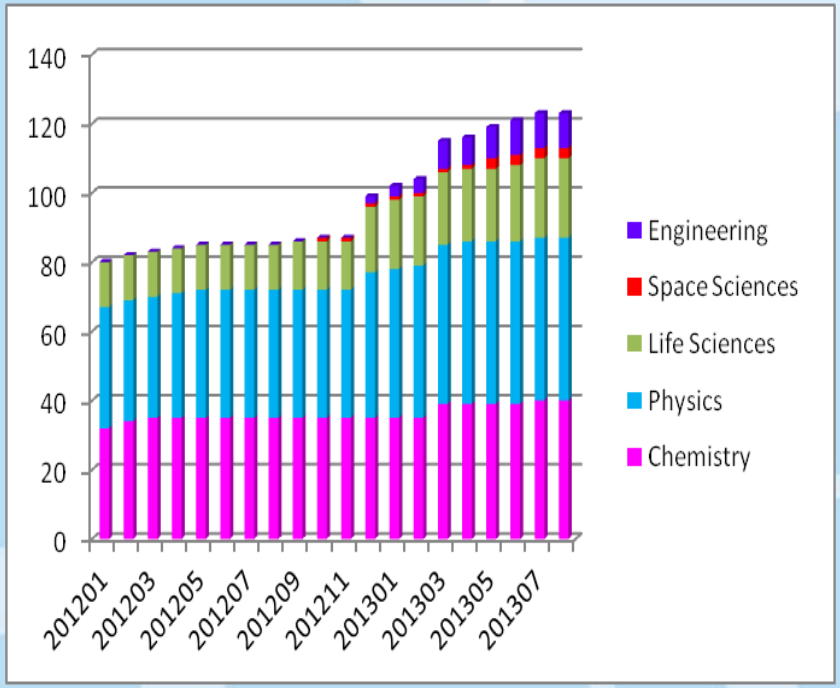
HPC: Flexible access mechanisms

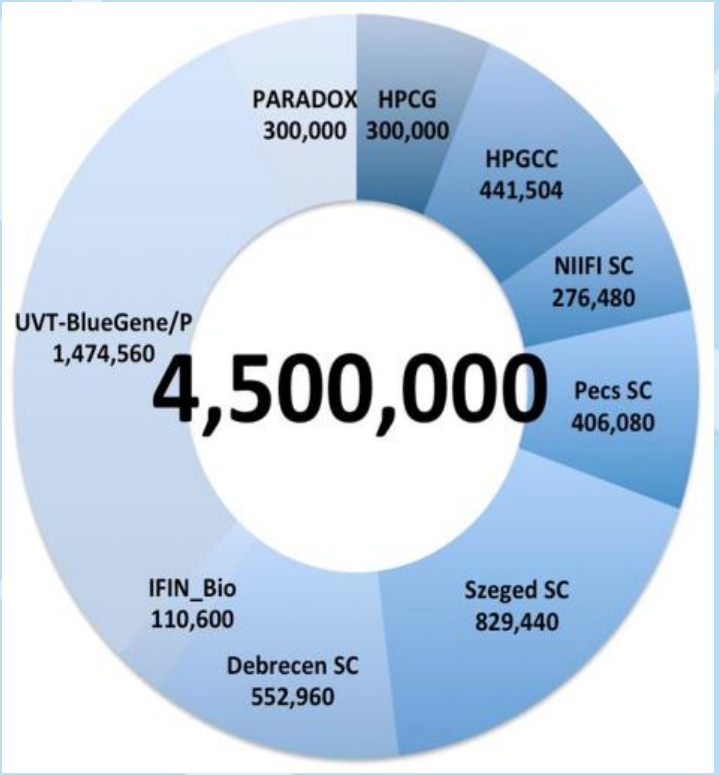
- Pilot call for access to resources
  - Based on regional resource sharing MoU
  - Resources offered: 4.6M Core hours, 1.8 M GPU hours
  - Allocations for 1 year – starting December 2012
  - Peer review based
  - Access to the resources from all countries of the region
  - Access to Mediterranean region also
- Fast track access mechanism
  - Limited resources provided
  - 2 Month allocation period
  - Suitable for: New user communities – Non experienced users; 9 new applications

HPC: Pilot call

HPC resources committed ->

HPC resources used ->



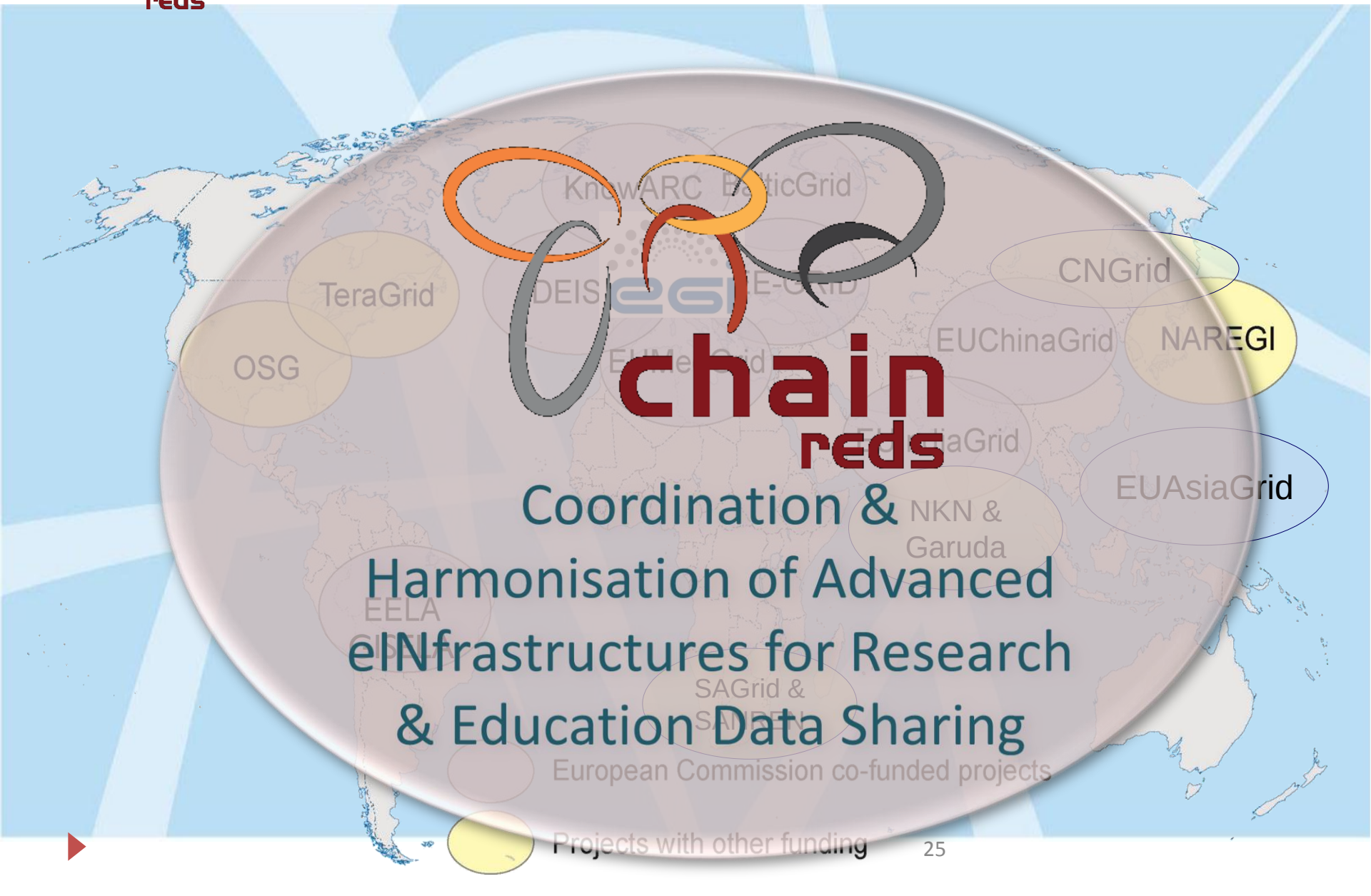


Work-flow and Timelines ->

Outline

- State of the art in eInfrastructures in Europe
- Regional collaboration models in networking, Grid computing and HPC - a case for international collaboration in Africa
- **CHAIN-REDS: worldwide collaborations**
- **CHAIN-REDS: Intercontinental Grids and HPC collaboration opportunities**

Regional Grid infrastructures



chain reds

Coordination & Harmonisation of Advanced eINfrastructures for Research & Education Data Sharing

CHAIN-REDS: Status

- Co-ordination & Harmonisation of Advanced e-Infrastructures for Research and Education Data Sharing
- Research Infrastructures – Support Action
- Grant Agreement n. 306819
- Total Costs of € 2.3 M
- EC contribution: € 1.52 M
- Start date: 1 December 2012
- Duration: 30 Months

CHAIN-REDS: Partners and roles

- INFN (IT) – Coordinator
- CIEMAT (ES) – WP4 Leader
- GRNET (GR) – WP3 Leader
- CESNET (CZ) – WP5 Leader
- UBUNTUNET (MW) – Africa
- CLARA (UR) – Latin America
- IHEP (CN) – China
- ASREN (DE) – Arab States
- SIGMA-ORIONIS (FR) – WP2 Leader
- C-DAC (IN) - India

CHAIN-REDS: Objectives

- Extend and consolidate international cooperation of Europe with other regions of the world in the domain of e-Infrastructures for Research and Education
- Promote, coordinate and support the effort of a critical mass of non-European e-Infrastructures for R&E to collaborate with Europe **addressing interoperability and interoperation of Grids and other DCIs**
- Study the opportunities of data sharing across different e-Infrastructures and continents widening the scope of the existing CHAIN Knowledge Base to Data Infrastructures and Cloud implementations
- Promote trust building towards open Scientific Data infrastructures across the world regions, including organisational, operational and technical aspects
- Demonstrate the relevance of intercontinental cooperation in several scientific data fields addressing existing and emerging VRCs and propose pragmatic approaches that could impact the everyday work of the single researcher, even if not structured in a VRC
- Provide guidance and recommendations for roadmaps for long-term global collaboration in e-Infrastructures & harmonization of existing polices

CHAIN-REDS: Action lines

DCI

Distributed Computing Infrastructure

- Provide ongoing support of the DCI road-map for intercontinental DCI collaboration, specified within the CHAIN project

ROC

Regional Operation Centres

- Support stability of existing and emerging Regional Operation Centres. Cooperate with other projects & initiatives (e.g. AfricaConnect, TEIN3) to support the development of eInfrastructures and key VRCs in Africa, Asia, Latin America and the Middle-east

Cloud

Clouds for Research and Education

- Support for coordination of Cloud developments for Research & Education with other regions (e.g. China, India, Latin America)

CHAIN-REDS: Action lines

Data

Infrastructures and Repositories

- Extend the CHAIN Knowledge Base with information on Data Infrastructures: collecting issues, best practices and identifying data repositories of direct interest for VRCs

  Support the study of data infrastructures for a target subset of VRCs (e.g. Agricolture, Climate Change, Health, Biomedicine, etc.)

SG

Science Gateways

- Promote the usage of Science Gateways as a means for attracting new communities and promote the use of eInfrastructures for every researcher

IDF

Federations of Identity Providers

- Foster the creation of Identity Federations in cooperation with Certification Authorities; promote and coordinate their usage. Support integration of different AA approaches

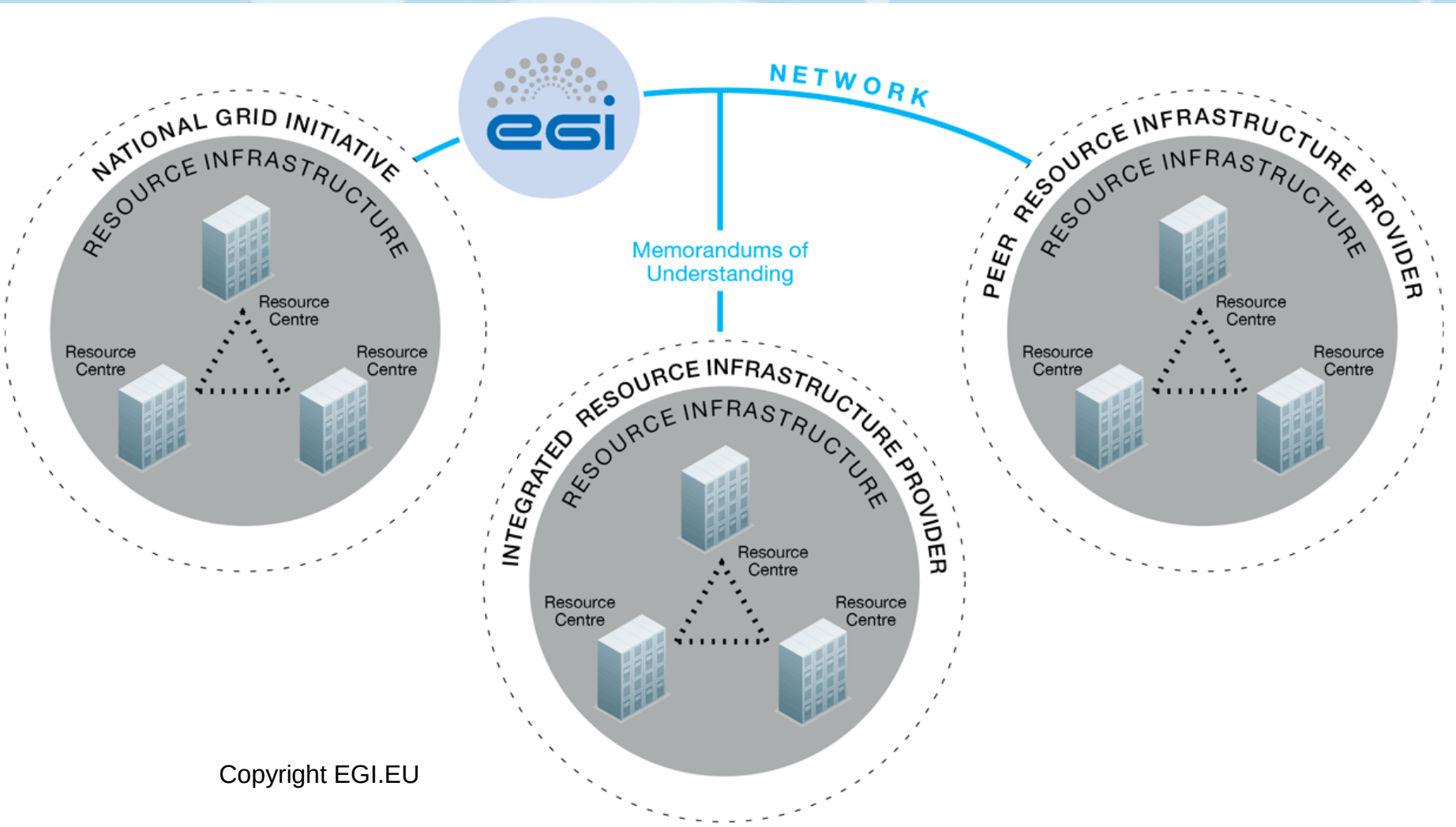
Grid: ROC guidelines

- Set of guidelines how to build and support ROCs for Grid computing
- Full action plan for their establishment and support.
- 6 regions have been identified: Africa (sub-Sahara region)&Arabia, Asia&Pacific, China, India and Latin America.
- Functionality: authentication and authorization, monitoring, user and operational support, management of Service Level Agreements, helpdesks, etc.

Example action plan: ROC Africa&Arabia

Contacts	✓	Bruce Becker
Status	✗	*7 Sites (to be updated with all regional sites). The ROC is operational. Not registered in GOCDB.*
Helpdesk	✓	https://support.africa-grid.org *This is an XGUS instance*
Accounting	✗	*Accounting records are not published.*
Monitoring	✗	*Monitoring information is not published. The ROC runs SAM-NAGIOS but there is no data in it.*
Website	✓	http://www.africagrid.org
Action Points		AP-AAROC-1: Sign MoU with EGI.eu as an Integrated Resource Infrastructure Provider AP-AAROC-2: Provide IGTF Accredited Certificate Services that will cover the whole AA ROC AP-AAROC-3: Create a new Operations Center in the EGi.eu GOCDB and register Resource Centers AP-AAROC-4: Setup and Operate a Grid Monitoring Service AP-AAROC-5: Publish accounting records to the EGI.eu Accounting System from all certified Resource Centers AP-AAROC-6: Adopt and employ Operational Policies and Procedures AP-AAROC-7: Set up dedicated Support Unit in GGUS

Grid: Collaboration Architecture



Grid: Integrated Resource Infrastructure Provider requirements

1. MoU & SLA with EGI.eu
2. Set-up Operations Center providing
   - Accounting/Monitoring Systems
   - a Helpdesk System Integrated with GGUS
   - Core Services as needed
3. Register sites to GOCDB
4. Adhere to EGIs Best Practices and Policies
   1. Respond to tickets
   2. Maintain site availability and reliability high
   3. Always run the recommended versions of middleware and OS

Grid: Organizing the ROC

- Adhering to Grid Security and Operational Policies and Procedures
- Setting up a Helpdesk service integrated with a dedicated GGUS Support Unit
- Organise teams for 1st and 2nd level of support
- Setup Accounting and Monitoring services compatible with the EGI services.(e.g SAM/APEL)
- Registering sites in GOC database:
  - GOCDB is the central contact service of EGI.EU; is used to:
  - Collect ROC management contacts
  - Collect Site contact points
  - Register Services offered by each site (visible or not to EGI)
  - Declare downtimes

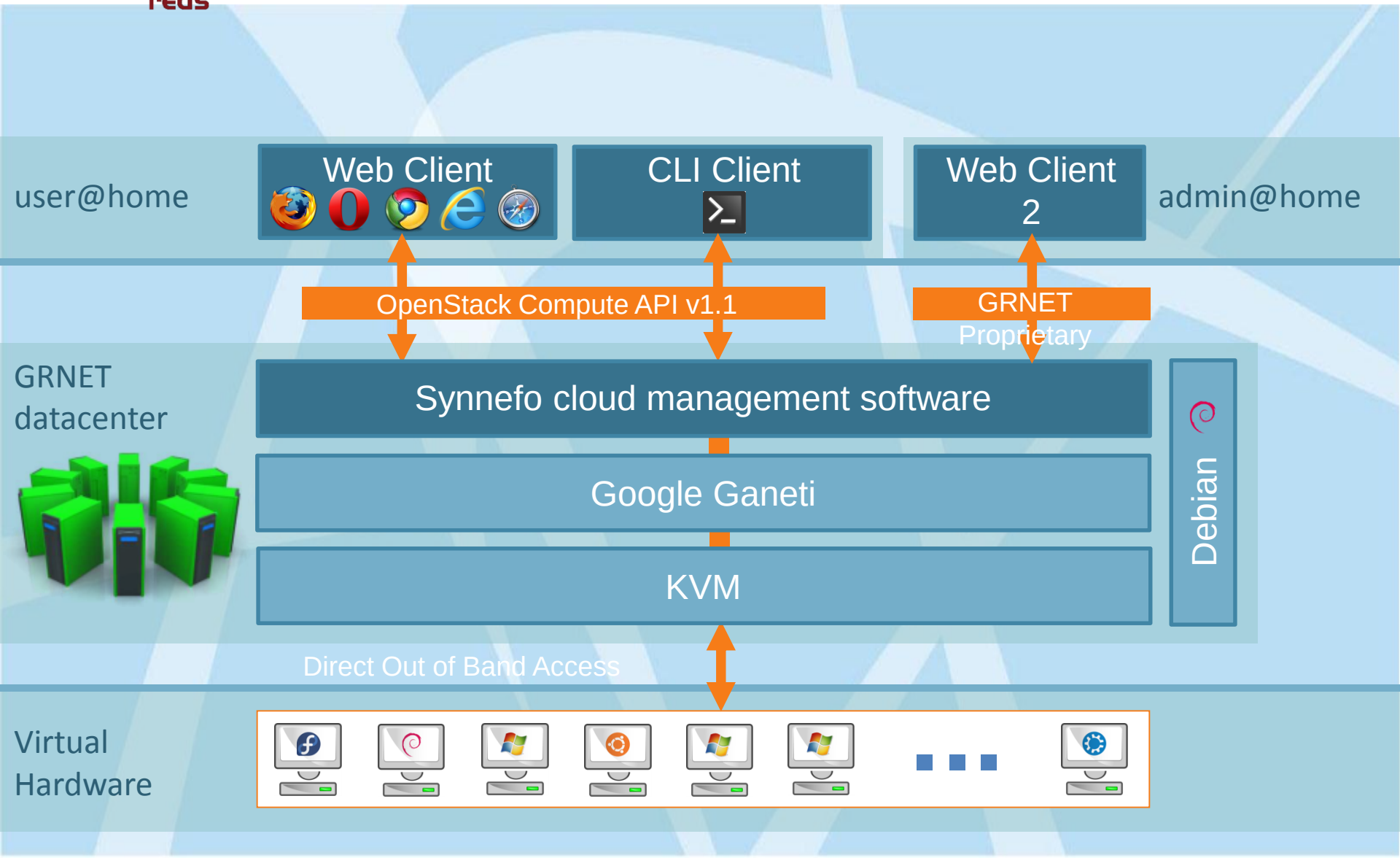
HPC: Intercontinental resource sharing

- CHAIN-REDS will identify HPC installations on the world-wide level
- Installations will be available in Europe/Greece, also initial agreement with Tianhe, would like to collaborate with CHPC
- Possibility to have a test pilot call for intercontinental HPC applications – as proof of principle mainly
- Following the HP-SEE model presented, and PRACE examples

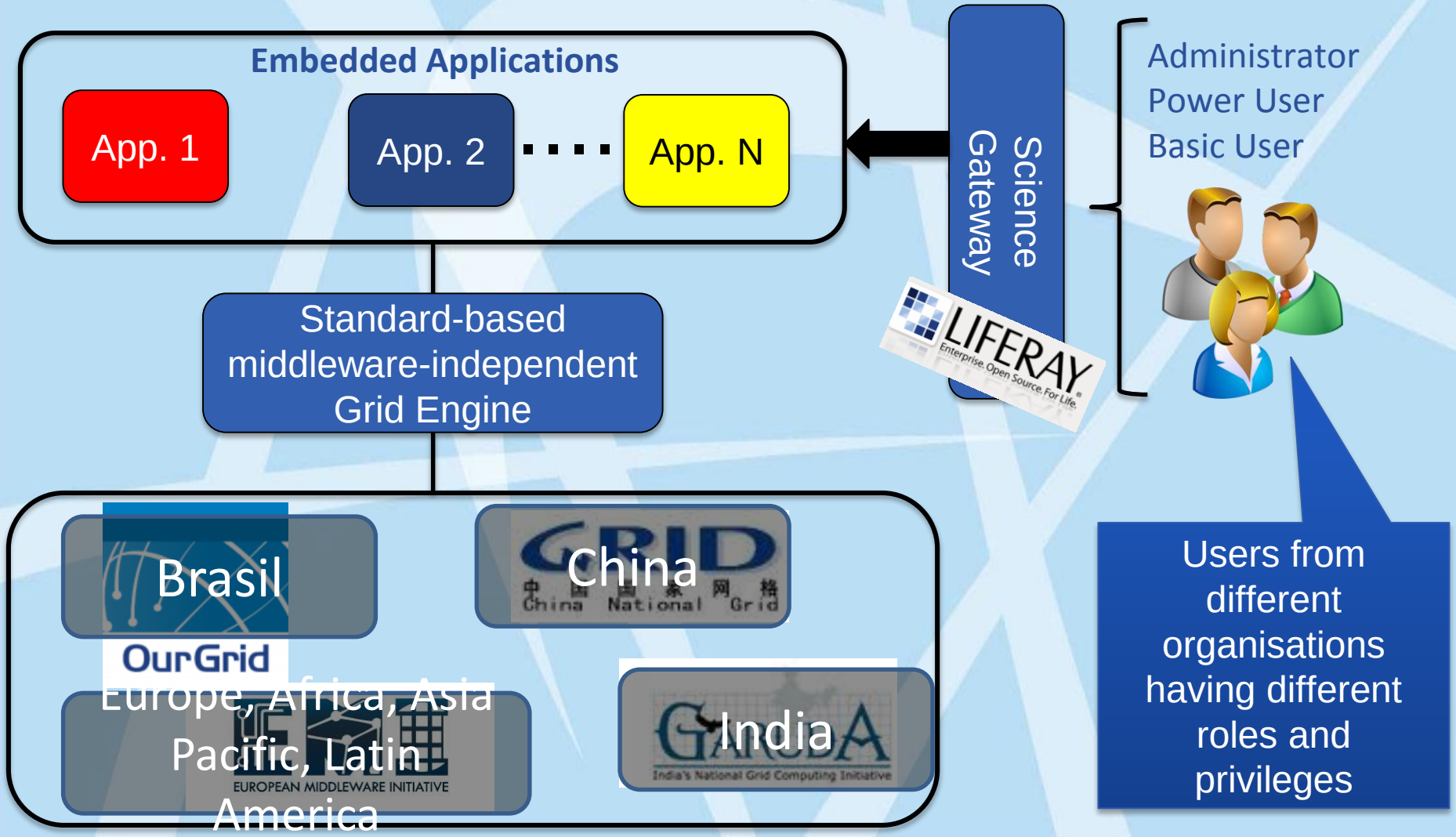
Cloud computing

- A number of new solutions
- GRNET: service ~okeanos; software synnefo
- ~okeanos is set to deliver IaaS
  - Compute (Virtual Machines)
  - Network (Virtual Networks, private Ethernets (L2) or public IPv4/6)
  - and Storage (Virtual Disks) / Pithos service
- Supports also project-like access for scientific computing
- 5K VMs, 2,2k users
- Synnefo: openstack compatible, uses google ganeti
- Details at okeanos.grnet.gr , trial at okeanos.io
- GRNET's customers: IT depts of connected institutions, university students, researchers in academia
- Users manage resources over a simple, elegant UI, of a REST API, for full programmatic control
- Increasingly accepted in the European NREN community

Cloud computing: ~okeanos



Access to heterogenous infrastructures via pre-embedded applcations: SG model



Outline: summing up

- State of the art in eInfrastructures in Europe
- Regional collaboration models in networking, Grid computing and HPC - a case for international collaboration in Africa
- CHAIN-REDS: worldwide collaborations
- CHAIN-REDS: Intercontinental Grids and HPC collaboration opportunities

Conclusion: layered developments & resource sharing

- National-level developments a baseline
- Regional integrations very important as a vehicle for interoperation and collaboration
- Path to worldwide infrastructures and integrations
- CHAIN-REDS provided a platform for DCI integration of Europe and other continents
  - ROCs for Grid interoperation model
  - HPC and SC: resource sharing and common calls for access
  - Clouds: diverse solutions, production-level available
  - Also SG model for heterogeneous access

- Call to contribute to AA ROC
- Call to contribute HPC/SC resources

- Thank you!


grnet
Διασυνδέοντας την Έρευνα και την Εκπαίδευση


chain
reds

CHAIN – Strategic Vision

- World-wide Distributed Computing Infrastructure can address big scientific challenges that are not manageable with local/national computing systems
- Virtual Research Communities can transparently access different kind of resources: scientific applications and tools, Data Repositories, CPUs and Disks. The vision is that of VRCs sharing resources ubiquitously across different administrative domains
- Regional e-Infrastructures should be made interoperable among each other. CHAIN is committed to promote and validate a proof-of-concept that addresses this.