

Sustainable Regional Development for e-Infrastructures: a Case Study

**Dr. Ognjen Prnjat,
Greek Research and Technology Network,
Greece**



HP-SEE

High-Performance Computing Infrastructure
for South East Europe's Research Communities



seera-ei

South East European
Research Area
for eInfrastructures



SEE-GRID-SCI

SEE-GRID eInfrastructure for regional eScience

Why this talk?

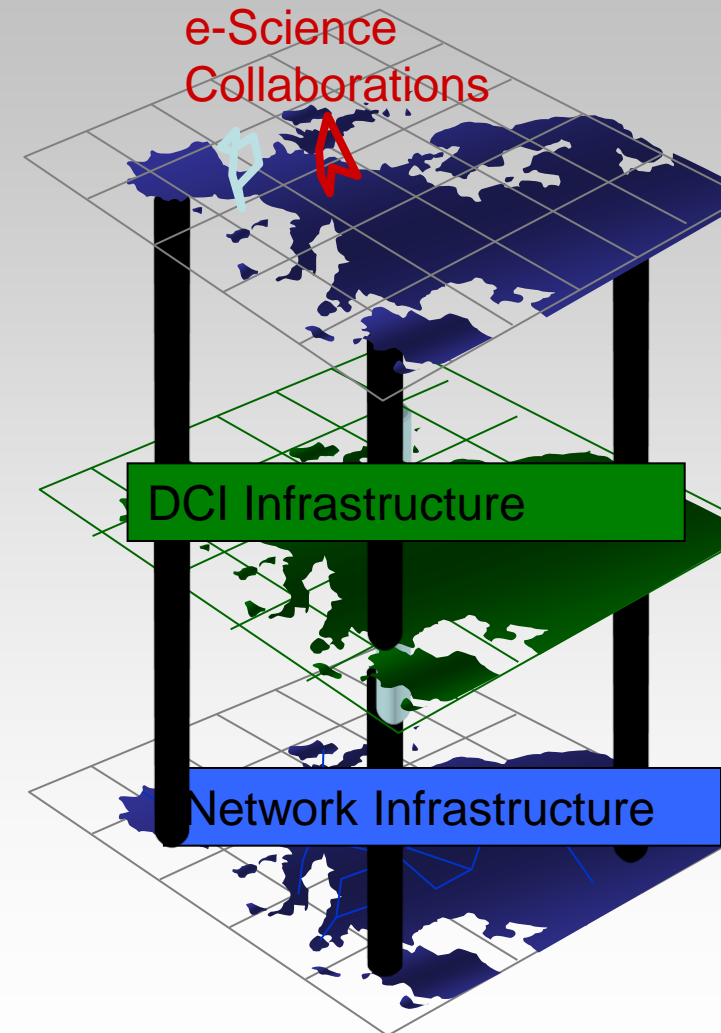
- EC strategy is to promote Pan-EU eInfrastructures across the world and link with peers; creating global collaboration model
- South-East European region is the first example where a set of coordinated and complimentary actions carried this out
- Before 2004: Region with high human potential and low infrastructural and financial resources – to some extent similar to Africa
- Now 2010: Role-model for regional eInfrastructure development; fully engaged in European eInfrastructure activities; wide range of user communities.

- Pan-EU eInfrastructure vision and its relationship with SEE initiatives
- Networking
- Grid
- High-Performance and Supercomputing
- Policy support and programme management

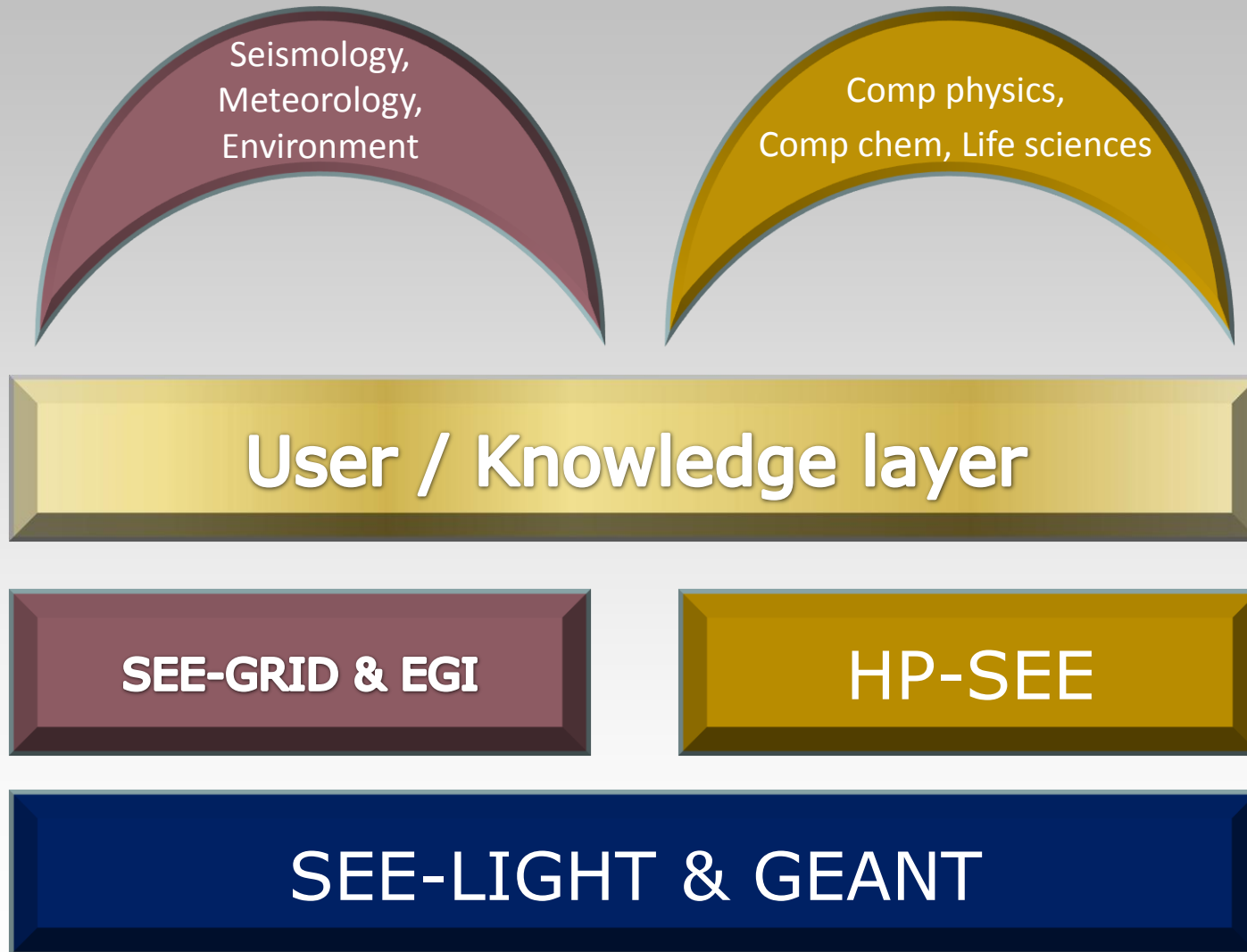
- **Pan-EU eInfrastructure vision and its relationship with SEE initiatives**
- Networking
- Grid
- High-Performance and Supercomputing
- Policy support and programme management

The pan-EU e-Infrastructures vision

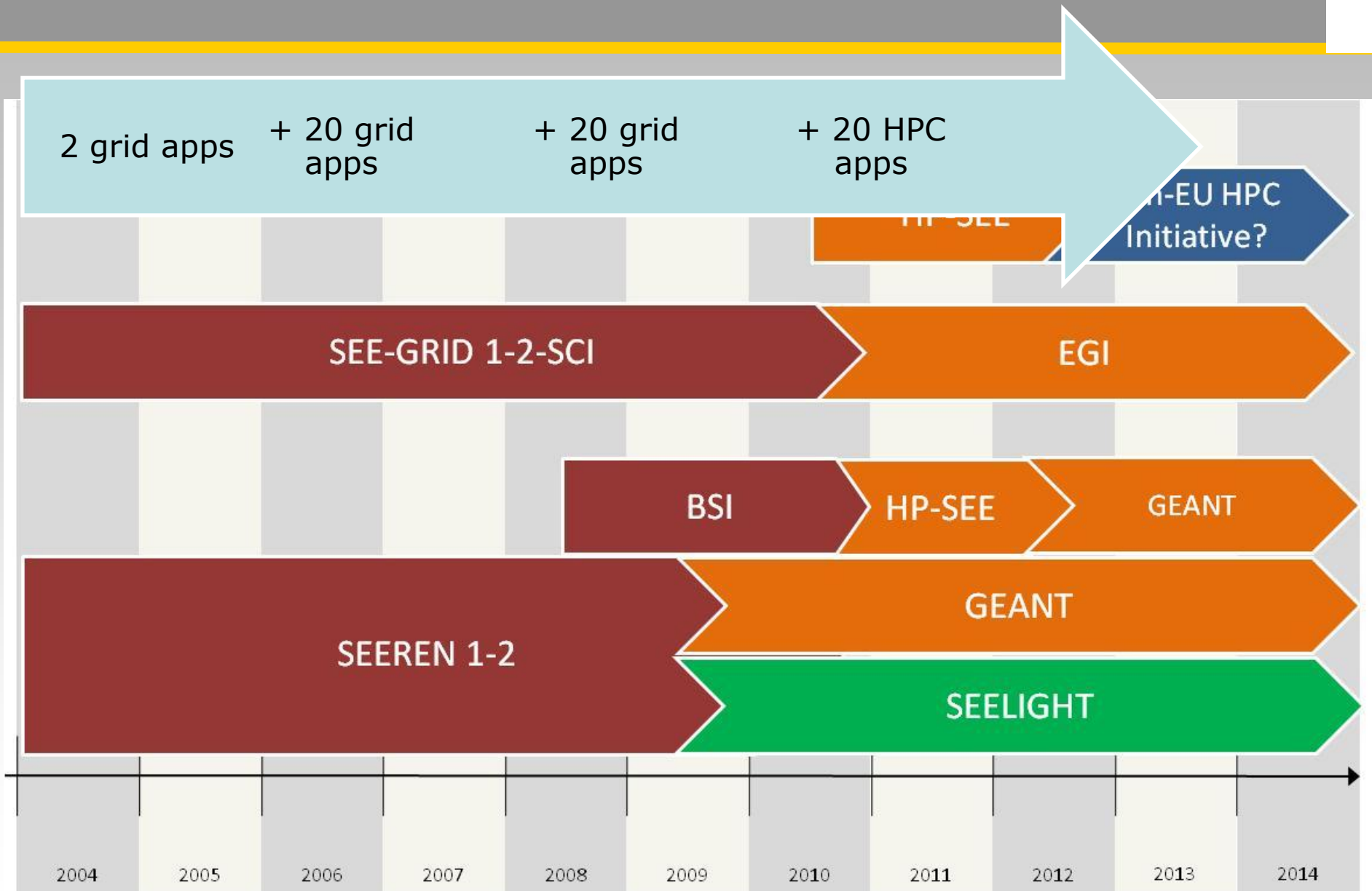
- The Research **Network infrastructure** provides fast interconnection and advanced services among Research and Education institutes of different countries'
 - ❑ Main Initiative: GEANT
- The **Distributed Computing Infrastructure (Grid, HPC)** provides a distributed environment for sharing computing power, storage, instruments and databases through the appropriate software (middleware) in order to solve complex application problems
 - ❑ Main Initiatives: EGI, DEISA, PRACE
- This integrated 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 and Digital Agenda**



The SEE Model: Converged communication & service infrastructure for South-East Europe



The Timeline



Timeline: 6 years of close collaboration in SEE

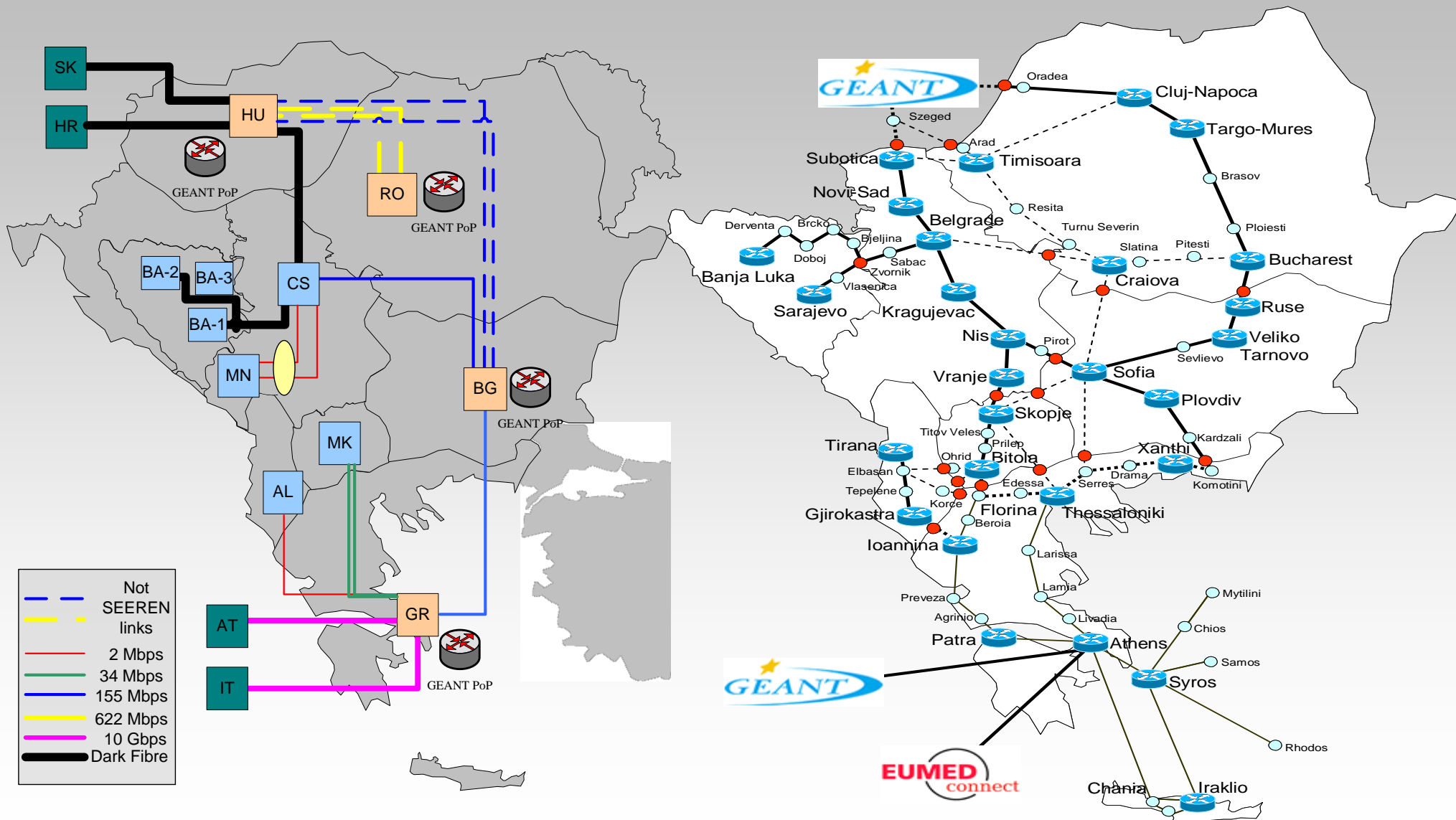
- **SEEREN1/2:** regional inter-NREN connectivity and GEANT links [DGINFSO]
- **BSI:** Southern Caucasus links [DGINFSO]
- **SEELIGHT:** lambda facility in SEE [Greek HiperB]
- Result: partly sustainable national & regional networks, EC funding for GEANT like rest of EU (50% national funds)

- **SEEGRID1/2:** regional Grid infrastructure within and beyond EGEE, building NGIs and user communities [DGINFSO]
- **SEE-GRID-SCI:** eInfrastructure for large-scale environmental science user communities. Inclusion of Caucasus. [DGINFSO]
- Result: partly sustainable national Grids, EC funding for EGI like rest of EU (with 75% national funds)

- **HP-SEE:** regional HPC interconnection and 2nd generation Caucasus link (50% EC)
- Expected result: sustainable national HPC centers, long-term sustainable (hierarchical) model with PRACE and DEISA, mirroring GEANT/EGI?
- **SEERA-EI:** regional programme managers collaboration towards common eInfrastructure vision, strategy and regional funds [DGRTD]
- Expected result: ensuring long-term national-level funds and regional funds to complement EC funds

- Pan-EU infrastructure vision and its relationship with SEE initiatives
- **Networking**
- Grid
- High-Performance and Supercomputing
- Policy support and programme management

Network: SEEREN2 and SEELIGHT



Network Sustainability: SEELIGHT

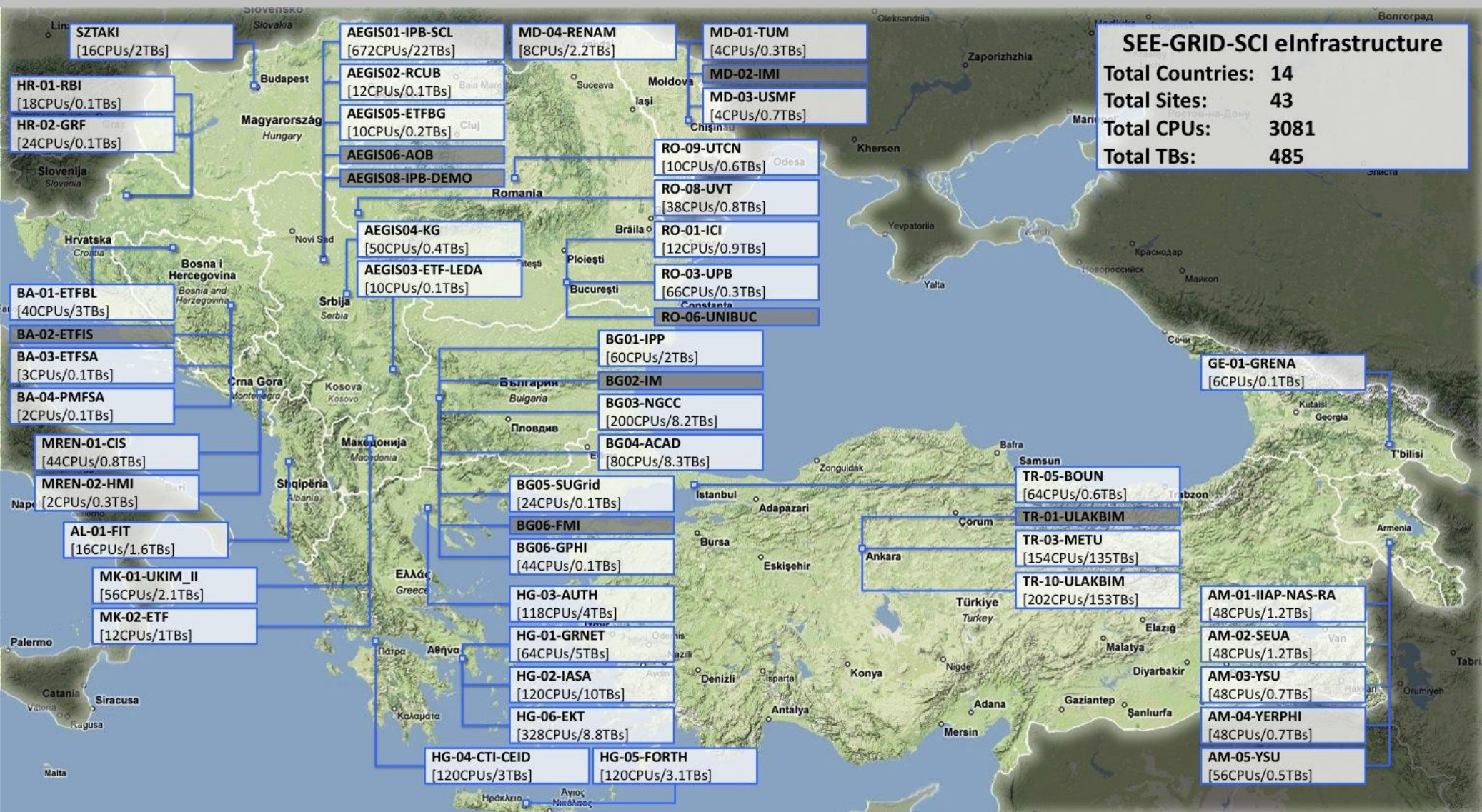
SEE-LIGHT: South-East European Lambda Network Facility for R&E

- Deployment of an advanced regional network infrastructure (lambda network facility) for R&E, under the Hellenic Plan for the Economic Reconstruction of the Balkans (HiPERB).
- 4 countries committed, others in progress
- SEENet: a management body for SEELIGHT
- National-level support for GEANT is crucial



- Pan-EU infrastructure vision and its relationship with SEE initiatives
- Networking
- **Grid**
- High-Performance and Supercomputing
- Policy support and programme management

Grid: SEEGRID1-2 and SEEGRIDSCI



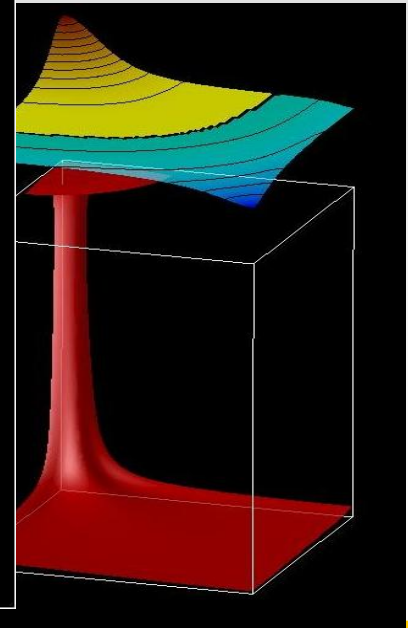
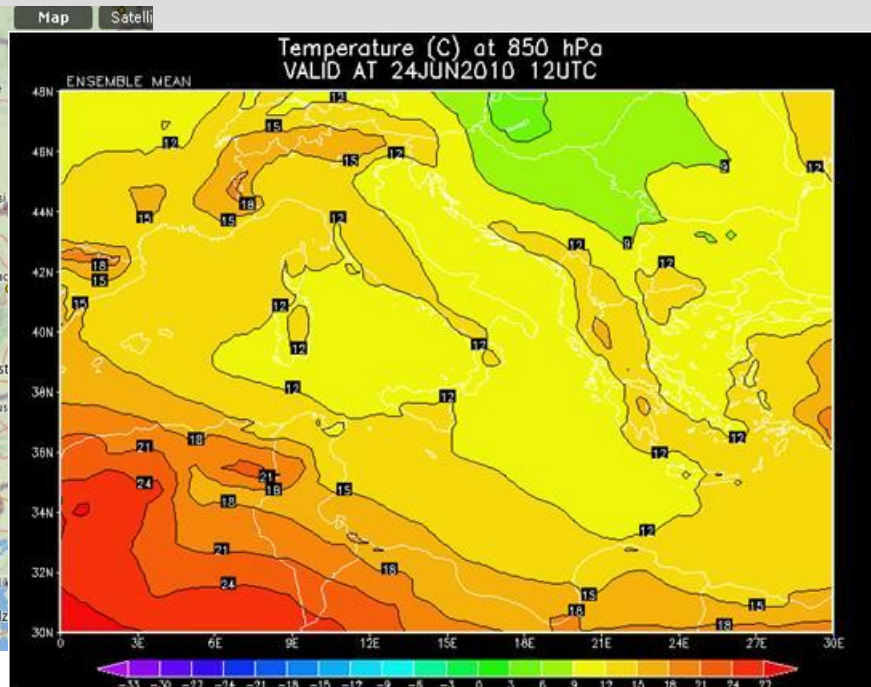
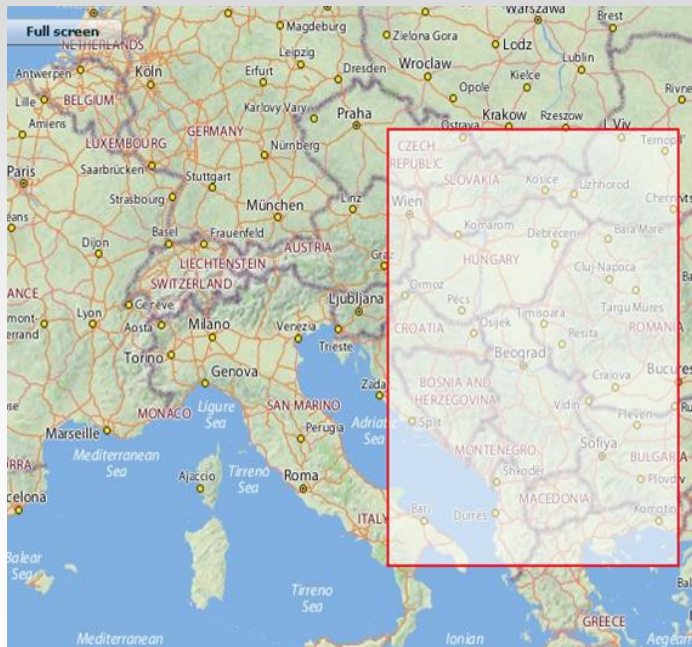
Grid Sustainability

- Large infrastructure built through 3 projects in 6 years
- User community buy-in secured
- National structuring via National Grid Initiatives
- All countries in European Grid Initiative, constant national-Ministerial support is necessary
- National-level projects important
- International support continues with CHAIN project



User communities support

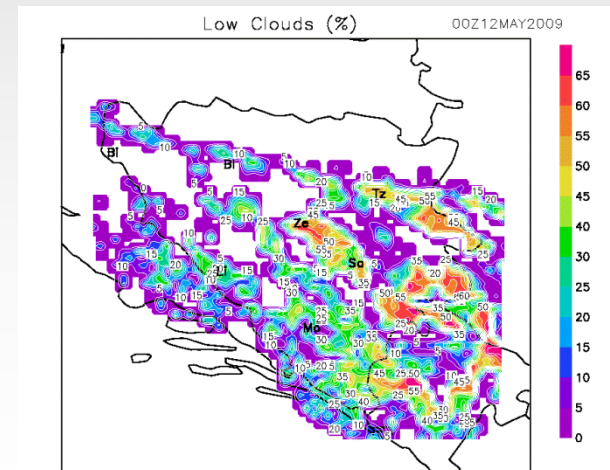
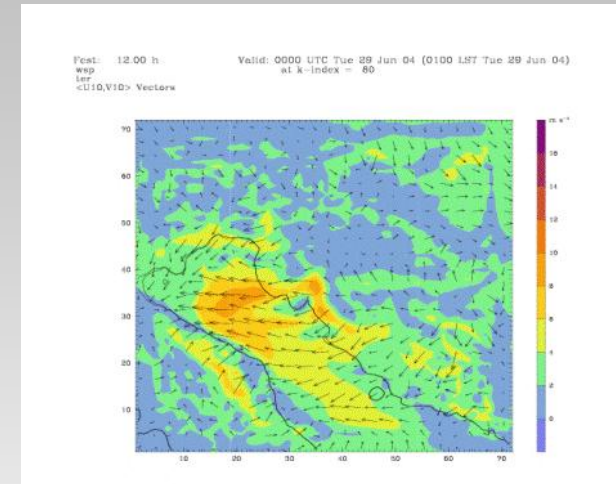
- Seismology (6 major applications), meteorology (2) and environmental protection (8); all in production
- Cross-border user communities and beneficiaries



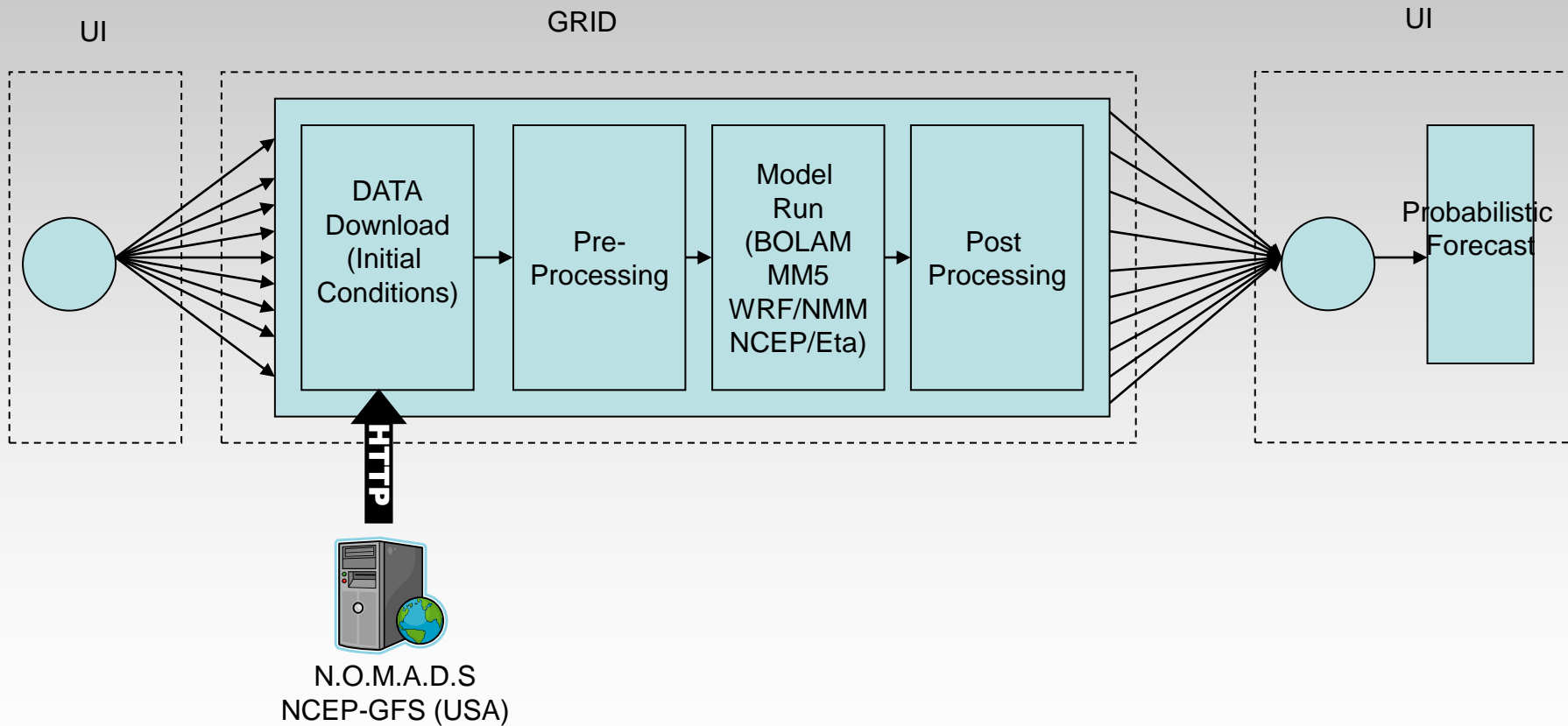
Meteorology: example

➤ Regional Multi-model, Multi-analysis Ensemble Prediction System

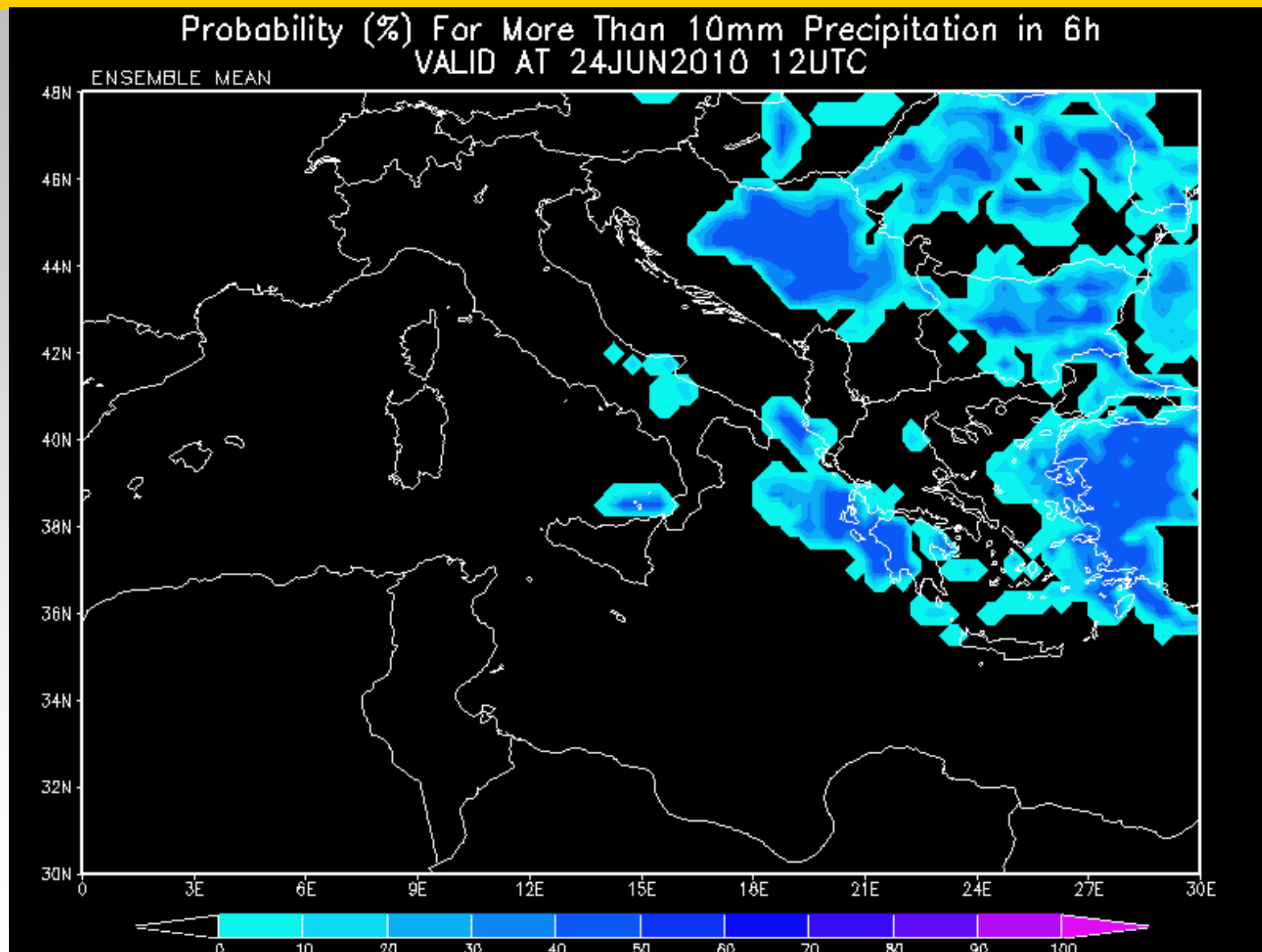
- ❑ BOLAM, MM5, NCEP/Eta, and NCEP/WRF-NMM
- ❑ SEE-wide scale detailed forecasts
- ❑ Coordinate, collect and analyze the outputs from all models for the generation of probabilistic forecasts
- ❑ Complex and CPU-intensive



Meteorology: Ensemble forecasting



Meteorology: Ensemble forecasting



Seismology

- 6 seismology applications
- Serves regional historical seismic data provided by partner national seismology centres using a high level interface (Seismic Data server - SDS) that is easy to use/adapt.
- Collaborates with other seismology related organizations (ORFEUS, EMSC) and provides high level interface to data collections of these organizations



Size of Continuous
Waveform Data

3.3 TB

No. of stations

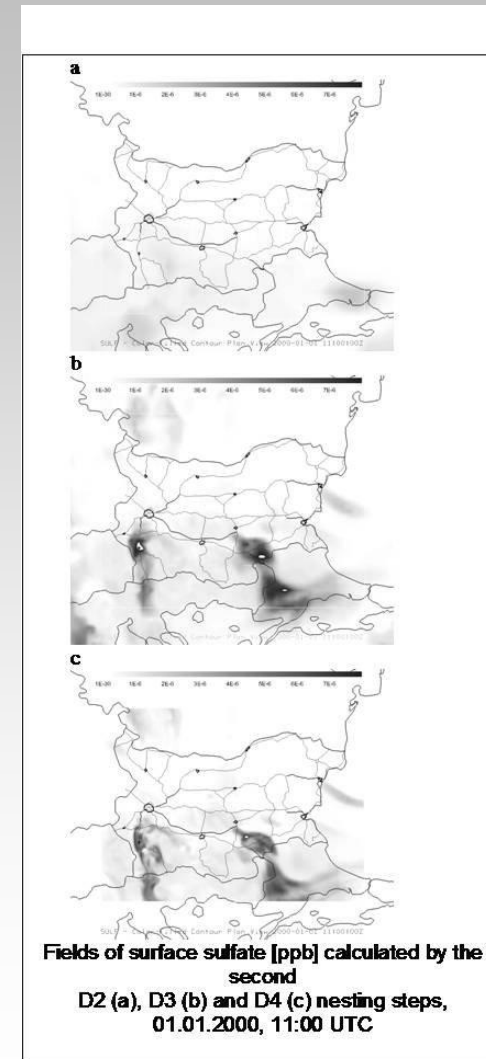
93

➤ Main application domains:

- ❑ Environmental protection/response
- ❑ Environment-oriented satellite image processing
- ❑ Extending to oceanography and climate change impact modeling

➤ 8 core applications:

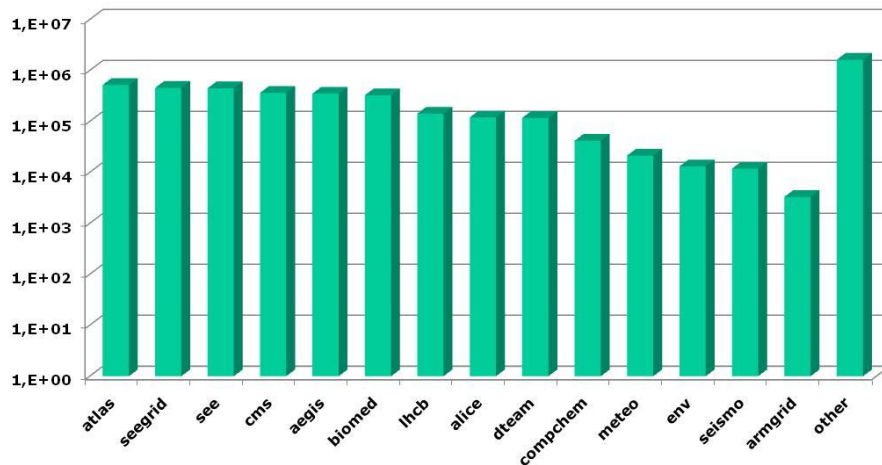
- ❑ Modeling System for Emergency Response to the Release of Harmful Substances in the Atmosphere
- ❑ Multi-scale atmospheric composition modeling (demoed with JTS JRA1 development)
- ❑ Monte Carlo Sensitivity Analysis for Environmental Systems
- ❑ Environment Oriented Satellite Data Processing
- ❑ Groundwater flow simulation system
- ❑ Study of Charges of Environment with Remote Sensing
- ❑ Climate Change Impact on Air Quality
- ❑ Oceanography



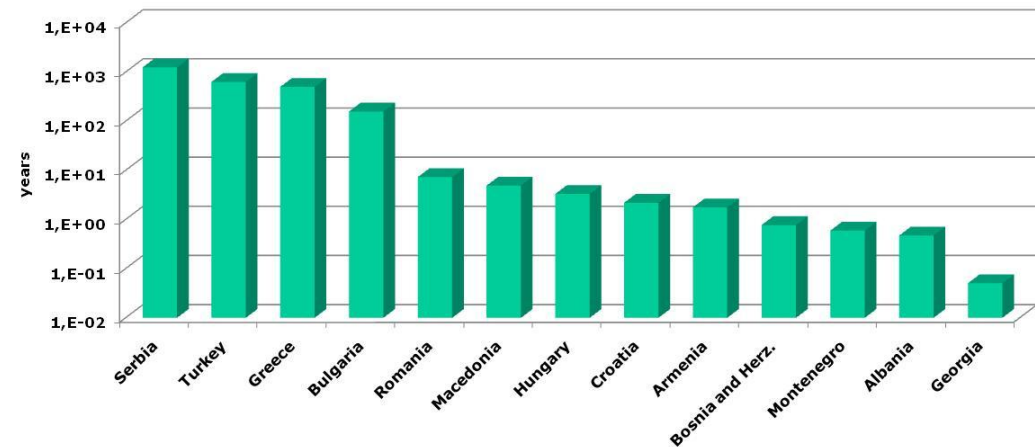
Grid usage by communities

- SEE-GRID sites have provided more than 22.5 million elapsed CPU hours or 2566 CPU years, and more than 4.5 million jobs during the project
- Utilization of dedicated resources: 89%

Number of Jobs per VO



CPU Time per SEE-GRID-SCI Country



- Pan-EU infrastructure vision and its relationship with SEE initiatives
- Networking
- Grid
- **High-Performance and Supercomputing**
- Policy support and programme management

High-Performance and Supercomputing

- **Project:** High-Performance Computing Infrastructure for South East Europe's Research Communities
- **Contract n°:** RI-261499
- **Project type:** CP & CSA
- **Call:** INFRA-2010-1.2.3: VRCs
- **Start date:** 01/09/2010
- **Duration:** 24 months
- **Total budget:** 3 885 196 €
- **Funding from the EC:** 2 100 000 €
- **Total funded effort, PMs:** 539.5
- **Web site:** www.hp-see.eu



HP-SEE Core Objectives

- Empowering multi-disciplinary virtual research communities in collaboration
- Deploying integrated HPC infrastructure for virtual research communities
 - ❑ Including a GEANT link to Southern Caucasus
- Policy development and stimulating regional inclusion in pan-European HPC trends

Existing infrastructure – Blue Gene/P

- IBM Blue Gene/P –two racks, 2048 *PowerPC 450*processors (32 bits, 850 MHz), a total of 8192 cores;
- Double-precision, dual pipe floating-point acceleration on each core;
- A total of 4 TB random access memory;
- 16 I/O nodes currently connected via fibre optics to 10 Gb/s Ethernet switch;
- Theoretical peak performance: $R_{peak} = 27.85$ Tflops;
- **Energy efficiency: 371.67 MFlops/W: Green top 10**
- Smaller HPC machines in Romania, Bulgaria, Hungary
- Upcoming purchases in Serbia and Greece

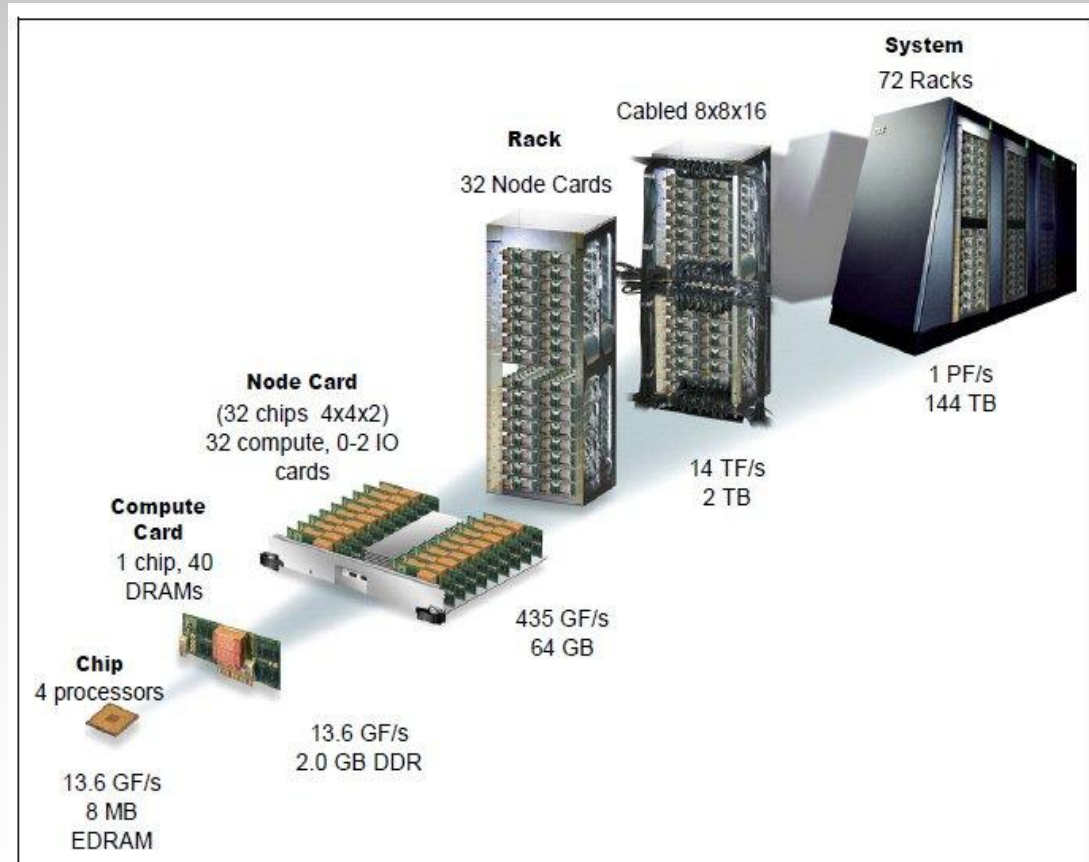


Figure 1-2 Blue Gene/P packaging

Introduction to Virtual Research Communities

➤ **Comp. Physics**

6 countries,
8 applications

➤ **Comp. Chemistry**

6 countries,
7 applications

➤ **Life Sciences**

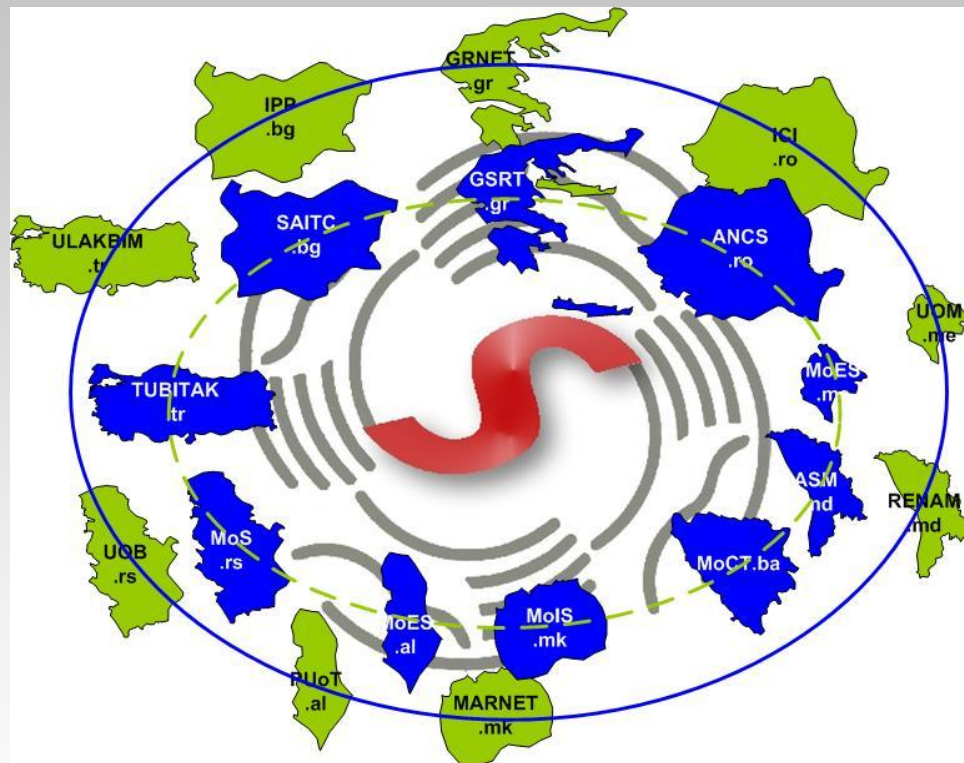
5 countries,
7 applications

Country	Physics	Chemistry	Life Sciences	TOTAL
Albania	1			1
Armenia			1	1
Bosnia-Herzegovina		1		1
Bulgaria	2	2		4
Georgia			1	1
Greece		1	2	3
Hungary			2	2
Moldova	1			1
Montenegro			1	1
FYR of Macedonia	1	1		2
Romania	2	1		3
Serbia	1	1		2
TOTAL	8	7	7	22

- Pan-EU infrastructure vision and its relationship with SEE initiatives
- Networking
- Grid
- High-Performance and Supercomputing
- **Policy support and programme management**

SEERA-EI: the project

- **Area:** "INFRA-2008-3.1: ERA-NET supporting cooperation for research infrastructures in all S&T fields"
- **Start:** 1 April 2009
- **Duration:** 36 months
- **Total budget:** 2 000 000 €
- **Funding from the EC:** 1 500 000 €
- **Contract n°:** RI-228052
- **Project type:** Coordination and support action



- Core Objective – Develop and strengthen the coordination and cooperation of national eInfrastructures programmes in the region of South-East Europe.
 - The core objective of the SEERA-EI project is to engage national key programme owners in the field of Research Infrastructures (eInfrastructures) in SEE in common dialogue and planning, to establish a sustainable and sustained communication platform, and to undertake activities targeting durable cooperation in building the common regional Research Infrastructures vision, strategy and action plan.

- O1 - Enable common dialogue and exchange of information between national programme owners
- O2 - Analyse state of the art in national programmes
- O3 - Reduce fragmentation and improve coordination across SEE segment of ERA in field of RI, on the level of national programmes, and avoid creating niches of non-standardised programmes
- O4 - Facilitate efficient set-up and support for national eInfrastructure programmes, in coordination with common guidelines and regional vision

- O5 - Enable concrete cooperation of programme owners and common planning and development of joint activities, both within project duration and beyond it in long term
- O6 - Establish a long-term platform for sustained dialogue between programme owners that will extend beyond project lifetime
- O7 - Enable information exchange with rest of ERA, carry out international collaboration activities, and contribute to EU policy on eInfrastructures

Conclusion: Regional development

- A set of coordinated and complementary actions enabling rapid regional development in eInfrastructures
- Technical and policy actions
- Leveraging EC, regional and national funds

- The success metric is not the size or speed of the infrastructure
- These are part of wider R&D efforts to sustain regional development
 - ❑ Increasing the retention of talented scientists in the region
 - ❑ Pursuing joint R&D efforts among countries in the region
 - ❑ Making available the benefits of the Information Society for citizens
 - ❑ Easing the digital divide between the region and rest of EC
 - ❑ Regional political stability and cohesiveness

- Ready to help our African colleagues with accumulated know-how, and also via CHAIN project!