

## SEE-GRID eInfrastructure for regional eScience

[www.see-grid-sci.eu](http://www.see-grid-sci.eu)

**EGEE09 conference**

**Dr. Ognjen Prnjat, GRNET**

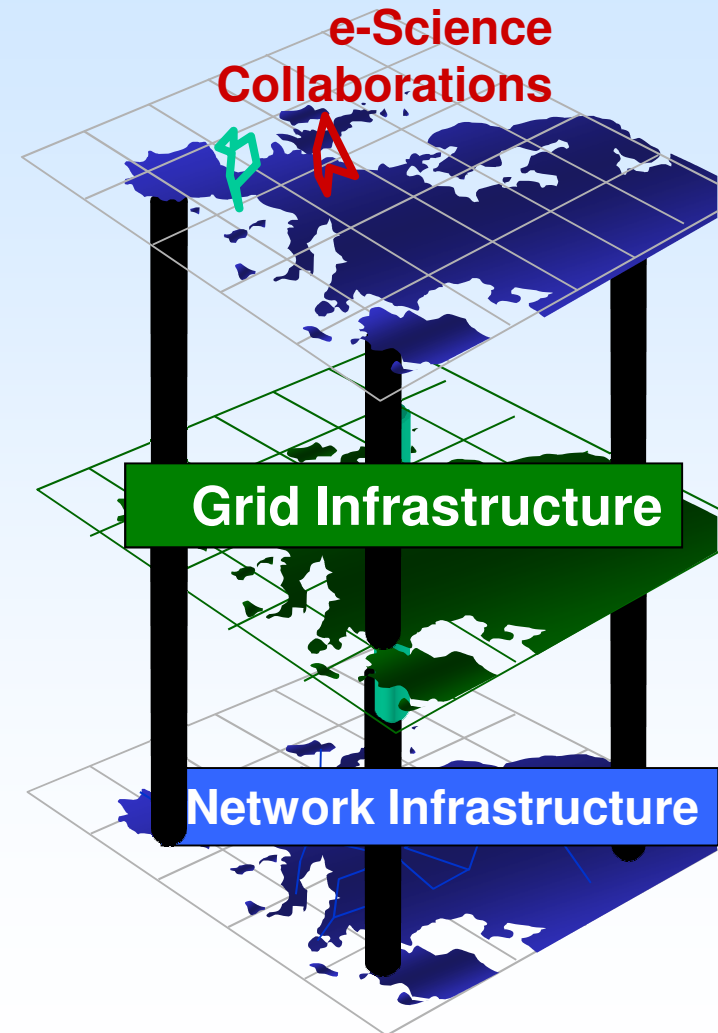


# Pan-EU e-Infrastructures vision



SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

- The Research **Network infrastructure** provides fast interconnection and advanced services among Research and Education institutes of different countries
- The Research **Grid 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
- Integrated networking & grid 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**



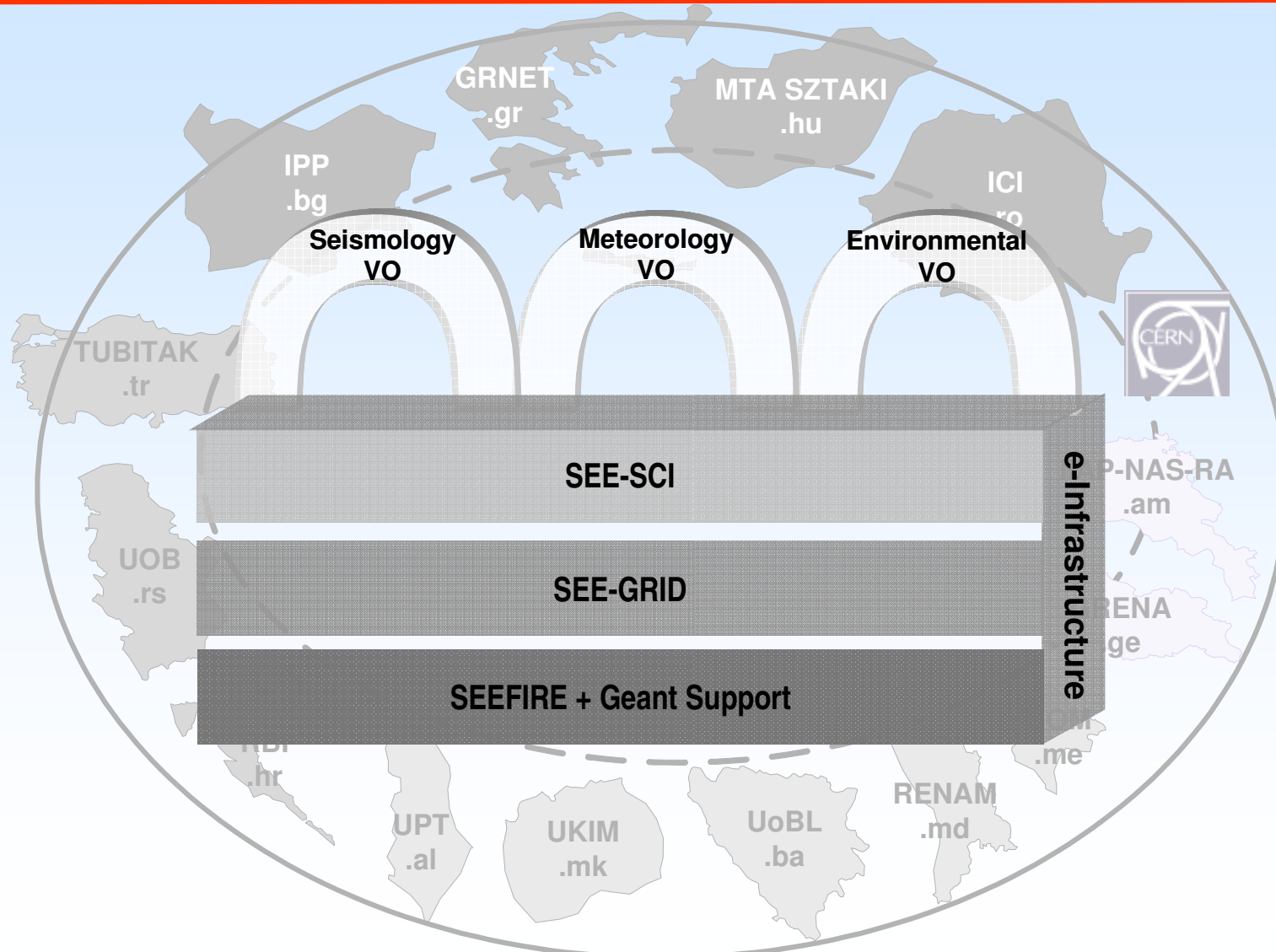
# Project context and timeline: 6 years of close collaboration in SEE

- **SEEREN1/2:** establishing the regional inter-NRN interconnectivity and GEANT links [DGINFSO]
- **SEEGRID1/2:** building the regional Grid infrastructure within and beyond EGEE, building NGIs and user communities [DGINFSO]
- **SEELIGHT:** implementation of the lambda facility in the region [Greek HiperB]
- **BSI:** Caucasus region connections [DGINFSO]
- **SEE-GRID-SCI:** eInfrastructure for large-scale environmental science: meteorology, seismology, env. protection. Inclusion of Caucasus. [DGINFSO]
- **SEERA-EI:** regional programme managers collaboration towards common eInfrastructure vision and strategy [DGRTD]

# Vision and context: converged communication and service infrastructure for SEE



SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience



# The project

- **Contract n :** RI-211338
- **Project type:** I3
- **Start date:** 01/05/2008
- **Duration:** 24 months
- **Total budget:**  
3 214 690 €
- **Funding from the EC:**  
2 500 000 €
- **Total funded effort, PMs:** 676.5
- **Web site:** [www.see-grid-sci.eu](http://www.see-grid-sci.eu)
- **Contact person:** Dr. Ognjen Prnjat, GRNET



SEE-GRID-SCI  
SEE-GRID eInfrastructure for regional eScience



# The partnership



SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

## Contractors

GRNET

CERN

SZTAKI

IPP-BAS

ICI

TUBITAK

ASA/INIMA

UoBL

UKIM

UOB

UoM

RENAM

RBI

IIAP-NAS-RA

GRENA

Greece

Switzerland

Hungary

Bulgaria

Romania

Turkey

Albania

Bosnia-Herzegovina

FYR of Macedonia

Serbia

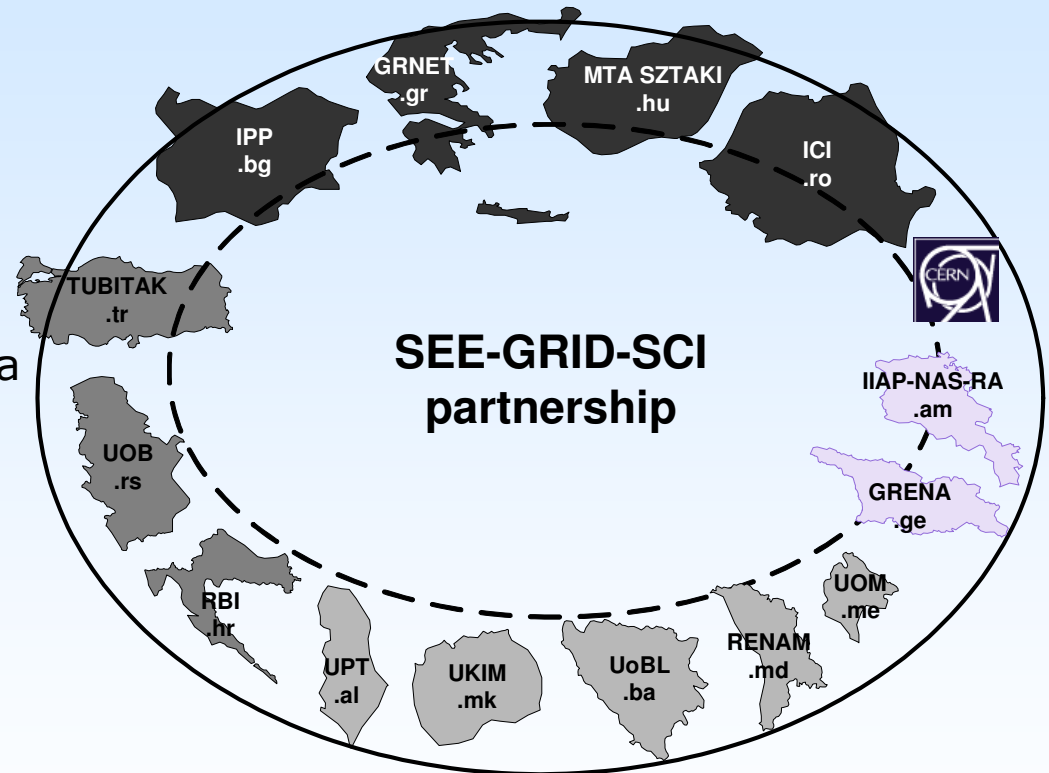
Montenegro

Moldova

Croatia

Armenia

Georgia



## Third Party / JRU mechanism used

associate universities / research centres

# Project objectives



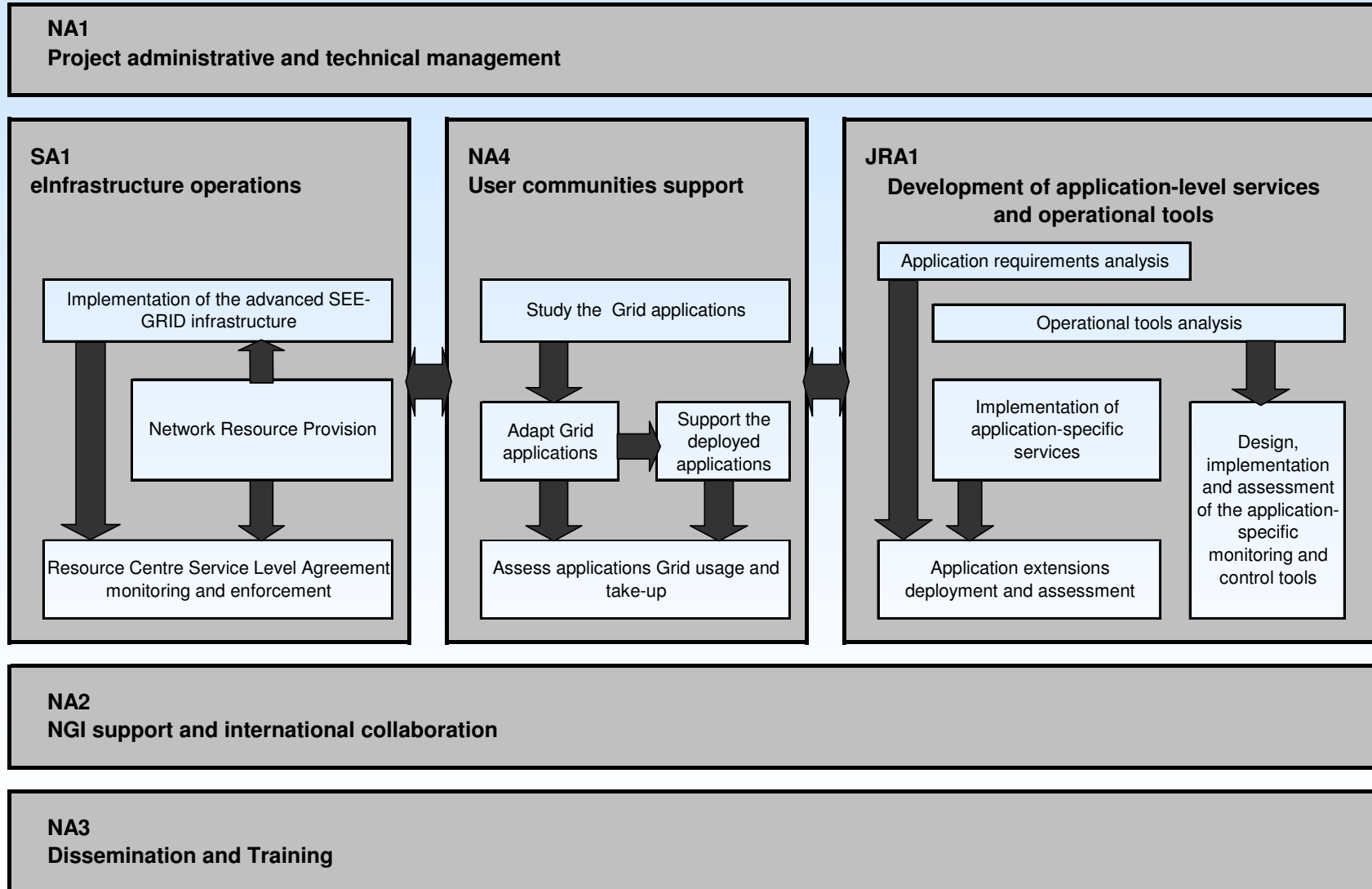
SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

- Engaging international user communities (meteorology, seismology, environmental protection) and providing application-specific service extensions
- Providing infrastructure for new communities
- Consolidating actions towards long-term sustainability and European Grid Initiative inclusion
- Strengthening the regional and national human network

# Work organization - PERT



SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience





# Key results: management



- Stable management structure built: Activity Leaders, Steering Committee Members, Grid Infrastructure Managers, Registration Authority Managers, etc.
- Project notebook as an efficient management tool
- Project Objectives and Activities kept constantly up-to-date.
- High complexity of management with 15 partners and 30 3<sup>rd</sup> parties

Project Steering Committee (SC)		
Company	Primary Member	Substitute
GRNET	Ognjen Prnjat	Valia Athanasaki
CERN	Frederic Hemmer	Florida Estrella
IPP	Todor Gurov	Emanouil Atanassov
ICI	Gabriel Neagu	Alexandru Stanciu / Vladimir Florian
TUBITAK	Burcu Ortakaya	Cevat Sener
SZTAKI	Péter Kacsuk	Miklos Kozlovsky
UPT	Gudar Beqiraj	Neki Frasher
UoBL	Milorad Bozic	Mihajlo Savic
UKIM	Aneta Buckovska	Margita Kon-Popovska
UOB	Zoran Jovanović	Antun Balaž
UOM	Bozo Krstajic	Luka Filipovic
RENAM	Peter Bogatencov	Veaceslav Sidorenco
RBI	Karolj Skala	Davor Davidovic
IIAP NAS RA	Hrachya Astsatryan	
GRENA	Ramaz Kvatadze	George Kobishvili

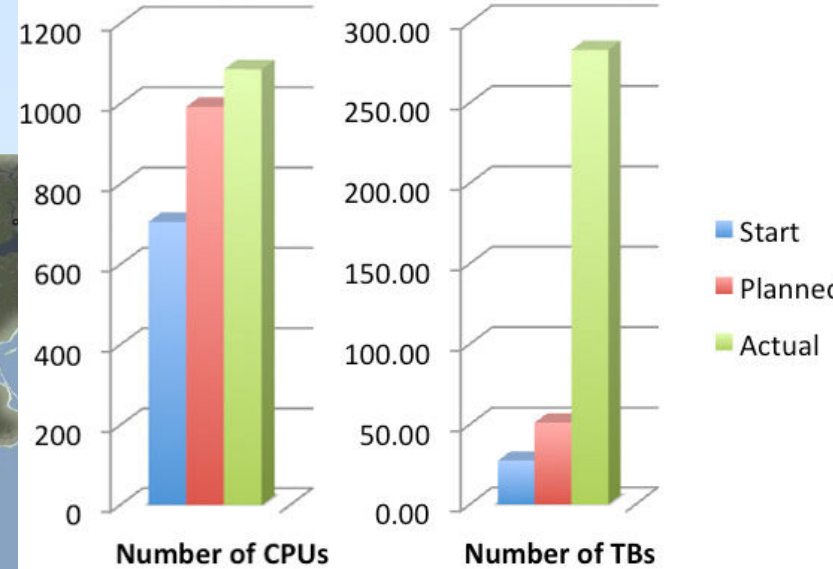
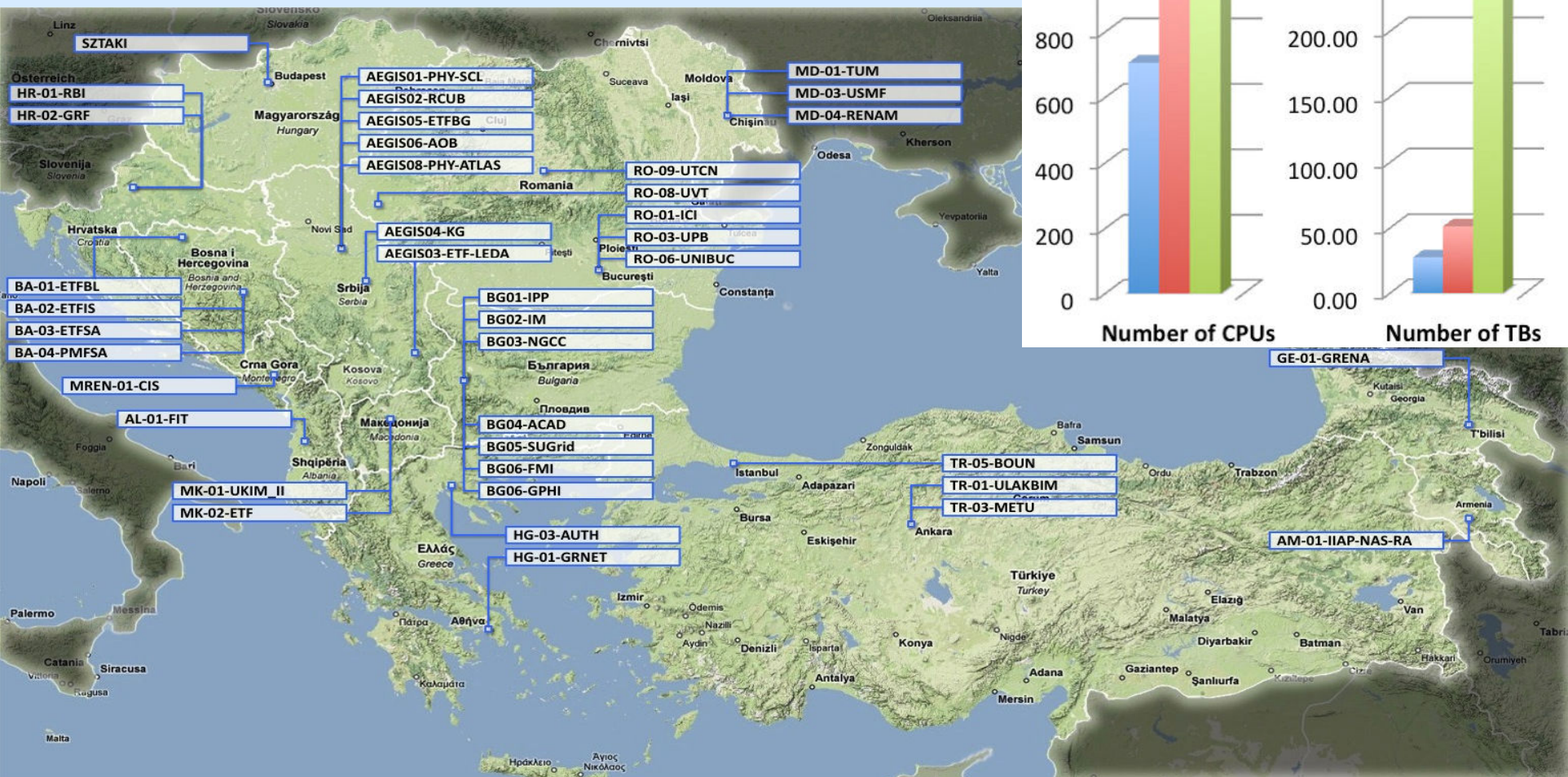
UoBL	112.016	56.008	0,00 €	4.842,00 €
UKIM	103.985	51.992	14.135,00 €	15.363,00 €
UOB	210.529	105.265	39.372,00 €	32.367,00 €
UOM	80.075	40.038	3.904,00 €	0,00 €
RENAM	259.501	29.750	6.448,00 €	8.276,00 €
RBI	112.558	56.279	15.798,00 €	7.681,00 €
IIAP NAS RA	69.016	34.508	4.000,00 €	10.000,00 €
GRENA	72.210	36.105	3.814,00 €	6.464,00 €

# Key results: Grid infrastructure



SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

- Dedicated CPUs: 1086 total
- Storage: 288 TB



# Infrastructure management



SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

The image displays a map of the SEE region (Southeastern Europe) with several infrastructure management tool windows overlaid. The tools include:

- What is @ the Grid?**: A web-based monitoring interface showing system status and metrics.
- seegrid GStat**: A dashboard showing regional service metrics and a table of site statistics.
- Nagios**: A network monitoring interface showing host status and service status.
- Helpdesk**: A ticket management system interface.
- BBlMSAM**: A monitoring tool showing a table of site details.
- Application Accounting Portal**: A portal for application accounting and resource usage.
- HGSM (Hierarchical Grid Site Management)**: A table listing grid sites with columns for ID, Country, Name, email, Phone, and Comments.
- Pakiti**: A host management interface for AEGIS01-PHY-SCL.

The map shows countries including Romania, Bulgaria, Greece, Serbia, Bosnia and Herzegovina, and parts of Croatia, Hungary, and Turkey. Major cities like Bucharest, Sofia, Athens, and Istanbul are marked.

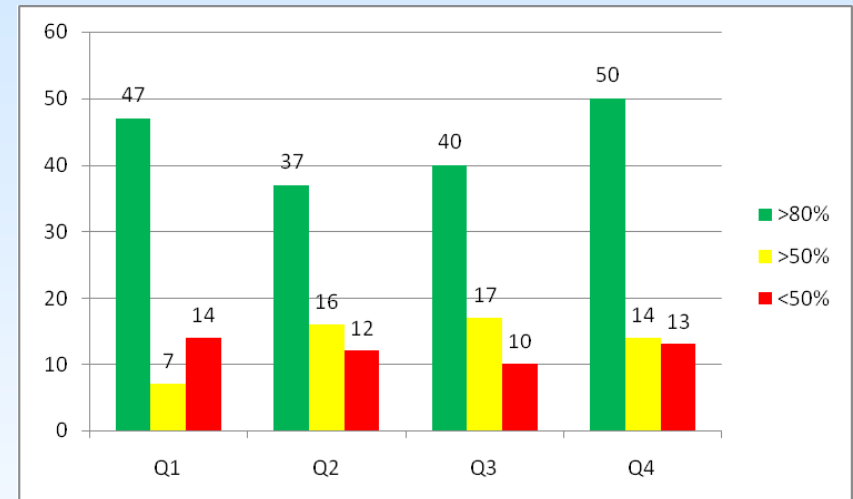
# Infrastructure operations



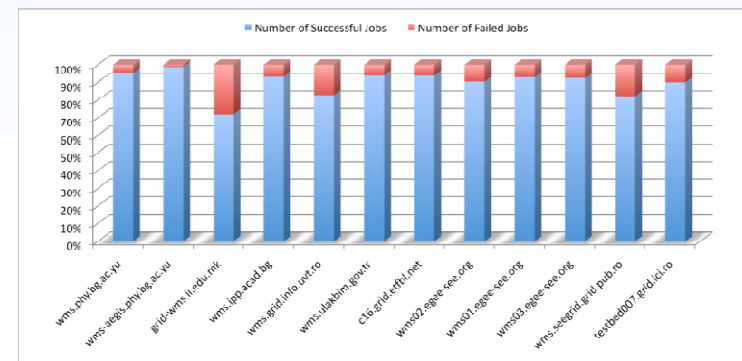
SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

- Core services deployed redundantly, spread out
- Catch-All Certification Authority + National CAs
- Operational tools as shown on previous slide
- Grid Operator on Duty (GOOD), Operations manual
- Fully interoperable with EGEE, overlapping
- SLA definition, monitoring and enforcement and corrective actions

- Overall availability of services 91.17%



- Job success rate 94,27%



# Key results: JRA1 development – Application Services



SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

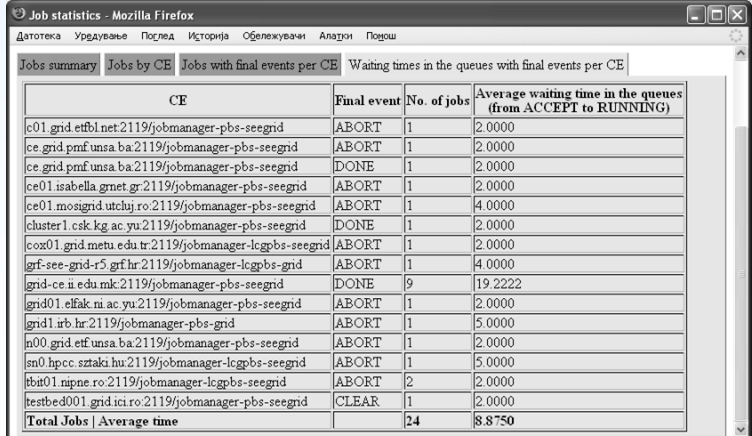
- Objectives:
  - Capture commonalities across scientific fields in terms of application requirements on Grid middleware
  - Define development areas for middleware plug-ins and application-level services
  - Implement these
  
- Rigorous requirements capture and design approach
- Architectures vary from libraries to compositions of coordinated distributed services
- 11 Application services / add-ons (4 in production)
- Examples: Seismic Data Server, User-Level Monitoring

# JRA1 development - Operational tools extensions

- Objective: Analyse application-focused features of operational tools and develop new or extend existing tools
  - Tools that collect data relevant to the operations of the infrastructure from the point of view of application developers and users
  - Tools that automate operational tasks and procedures which are particularly related to deployment and running of applications

- 7 tools

- Example: Job Tracking Service



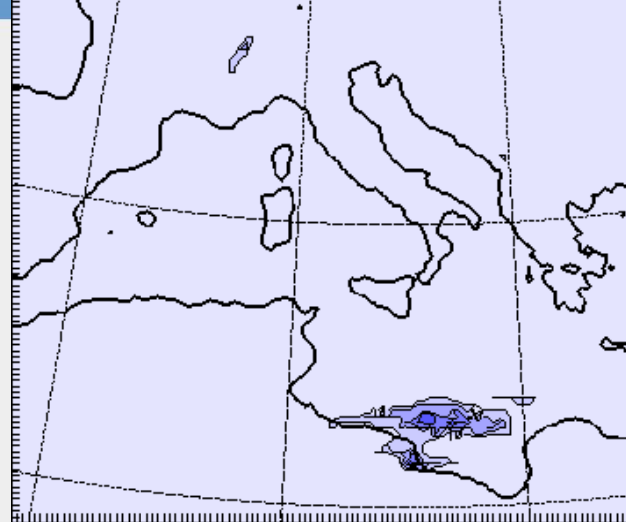
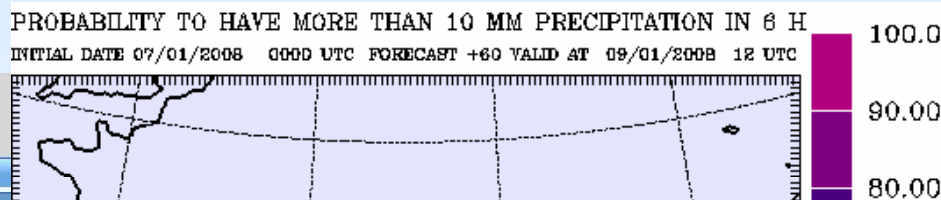
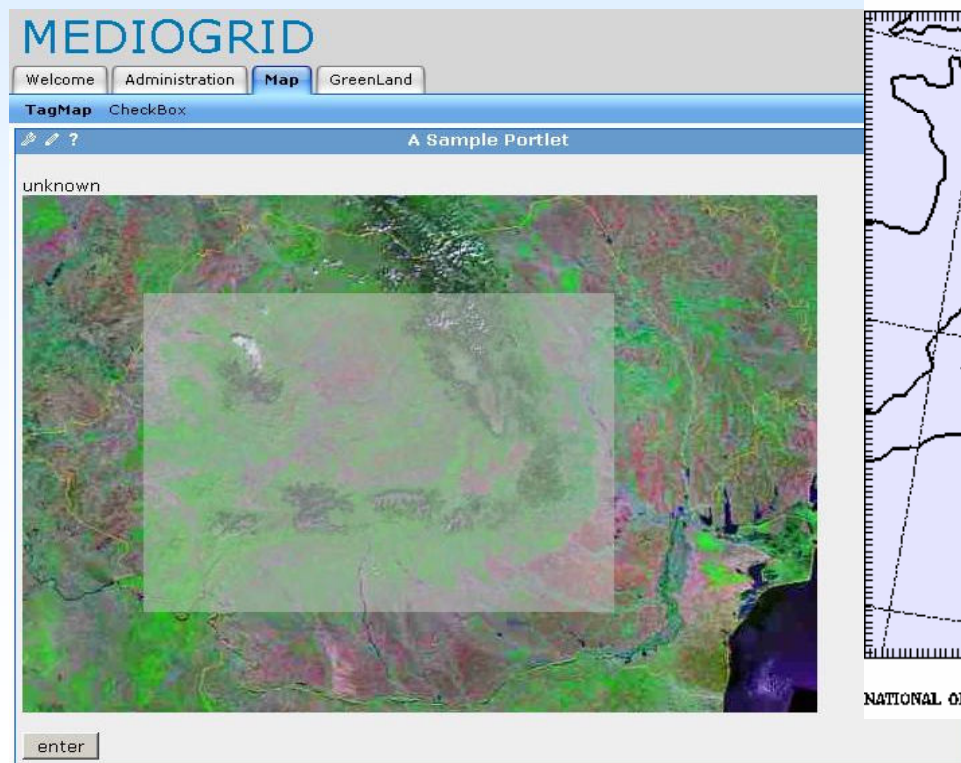
CE	Final event	No. of jobs	Average waiting time in the queues (from ACCEPT to RUNNING)
c01.grid.etfbl.net:2119/jobmanager-pbs-seegrid	ABORT	1	2.0000
ce.grid.pmf.unsa.ba:2119/jobmanager-pbs-seegrid	ABORT	1	2.0000
ce.grid.pmf.unsa.ba:2119/jobmanager-pbs-seegrid	DONE	1	2.0000
ce01.isabella.grnet.gr:2119/jobmanager-pbs-seegrid	ABORT	1	2.0000
ce01.mosigrid.utcluj.ro:2119/jobmanager-pbs-seegrid	ABORT	1	4.0000
cluster1.csk.kg.ac.yu:2119/jobmanager-pbs-seegrid	DONE	1	2.0000
cox01.grid.metu.edu.tr:2119/jobmanager-lcgpbs-seegrid	ABORT	1	2.0000
grf-see-grid-r5.grf.hr:2119/jobmanager-lcgpbs-grid	ABORT	1	4.0000
grid-ce.u.edu.mk:2119/jobmanager-pbs-seegrid	DONE	9	19.2222
grid01.elfak.ni.ac.yu:2119/jobmanager-pbs-seegrid	ABORT	1	2.0000
grid1.irb.hr:2119/jobmanager-pbs-grid	ABORT	1	5.0000
h00.grid.etf.unsa.ba:2119/jobmanager-pbs-seegrid	ABORT	1	2.0000
sn0.hpcc.sztaki.hu:2119/jobmanager-lcgpbs-seegrid	ABORT	1	5.0000
tbit01.nipne.ro:2119/jobmanager-lcgpbs-seegrid	ABORT	2	2.0000
testbed001.grid.ici.ro:2119/jobmanager-pbs-seegrid	CLEAR	1	2.0000
<b>Total Jobs   Average time</b>		<b>24</b>	<b>8.8750</b>

# Key results: User communities support

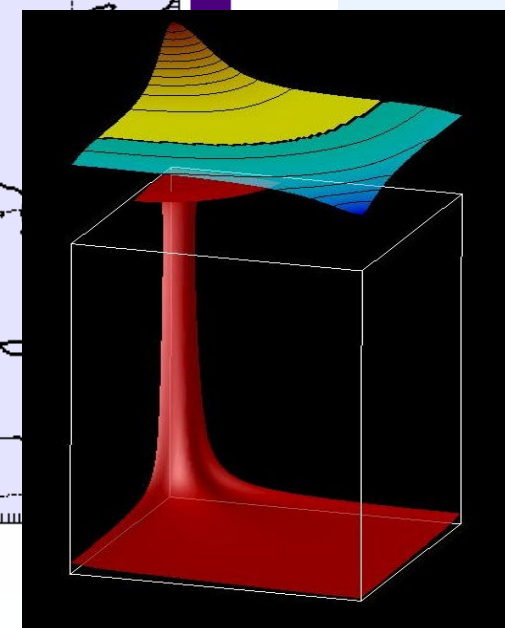


SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

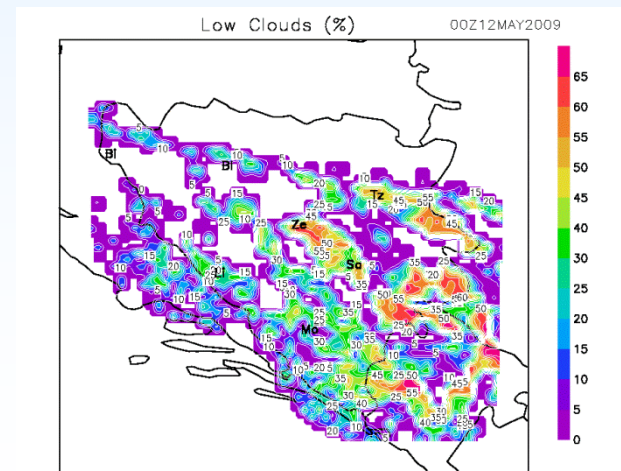
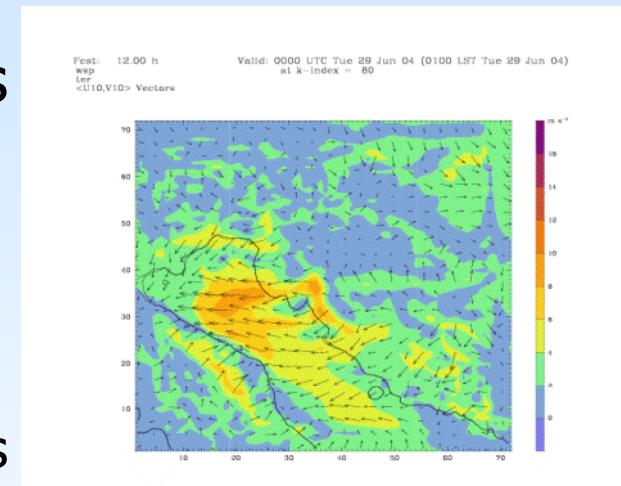
- Seismology (6 major applications), meteorology (2) and environmental protection (8); in satisfactory stages of deployment
- Cross-border user communities and beneficiaries
- Very clear and efficient procedures for support



NATIONAL OBSERVATORY OF ATHENS



- 1. 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
  - Very complex; very CPU-intensive
- 2. Study of interaction of airflow with complex terrain

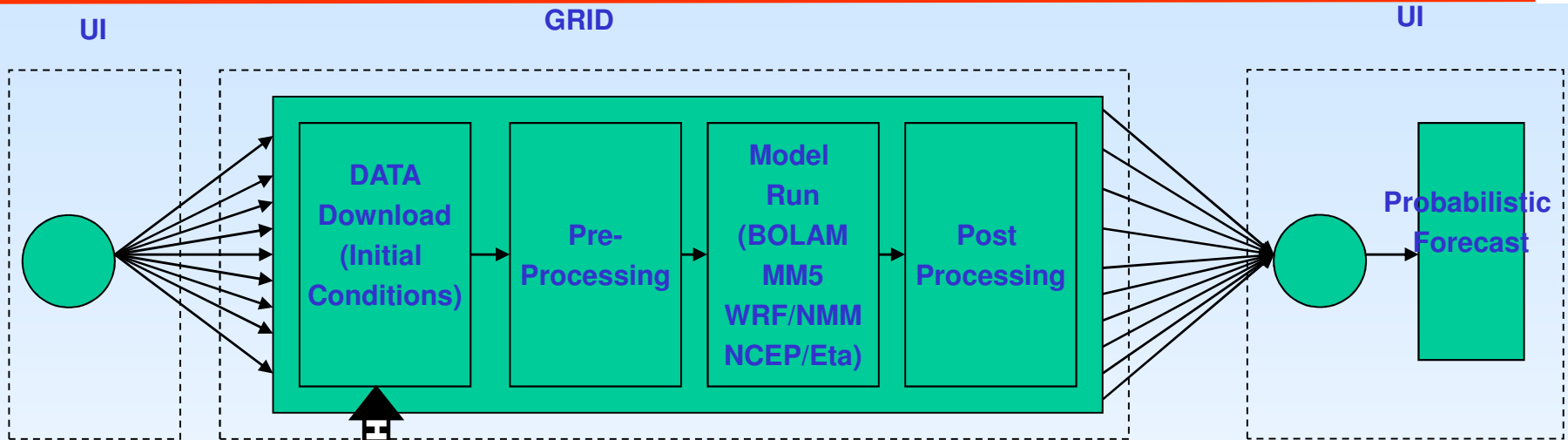




# Meteorology VO: Ensemble forecasting

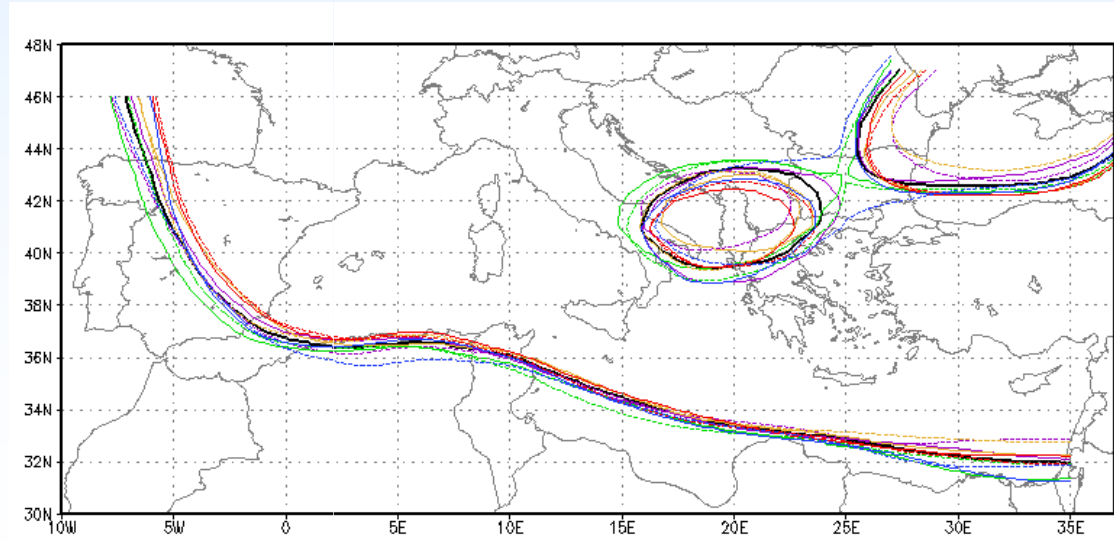


SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience



  
N.O.M.A.D.S  
NCEP-GFS (USA)

HTTP



# Seismology VO



SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

- Seismology VO offers the researchers:
  - Access to seismic data mirrored from national research centers on a timely basis
  - Adequate computing resources close to the seismology data repositories
  - Collaborative working environment with both regional groups and global organizations
- Core: seismic data server serving large seismic data sets from sensors (TBs)
- 6 Applications



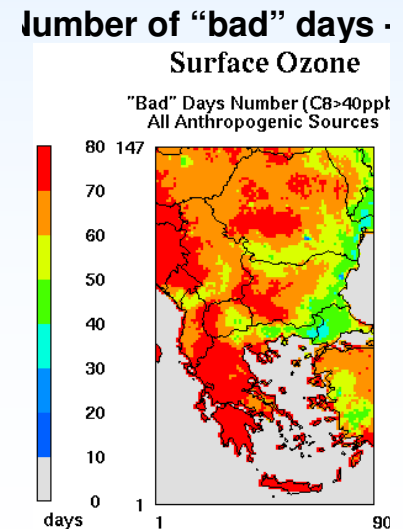
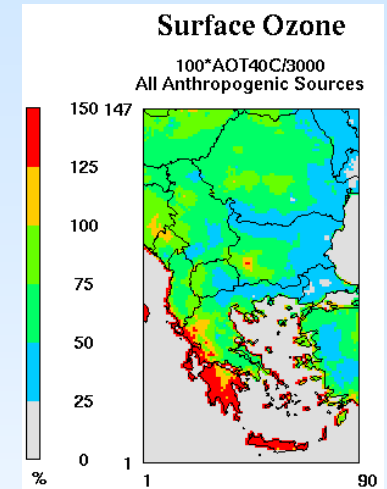
Size of Continuous Waveform Data	2.44 TB
No. of stations	86

## Main application domains:

- Environmental protection/response
- Environment-oriented satellite image processing
- Extending to oceanography and climate change 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 + related applications
- Groundwater flow simulation system
- Study of Charges of Environment with Remote Sensing
- Climate Change Impact on Air Quality
- Oceanography

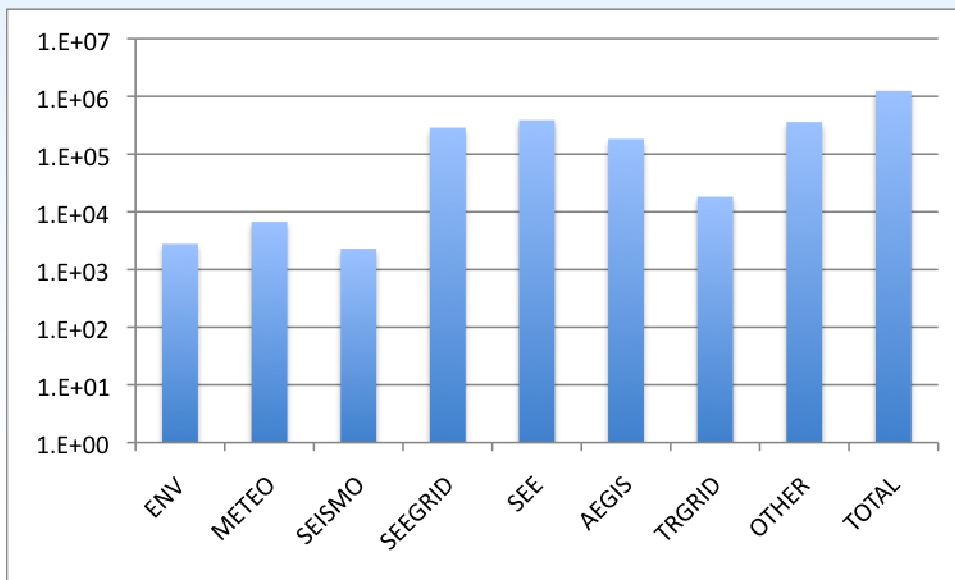


# Infrastructure usage by communities

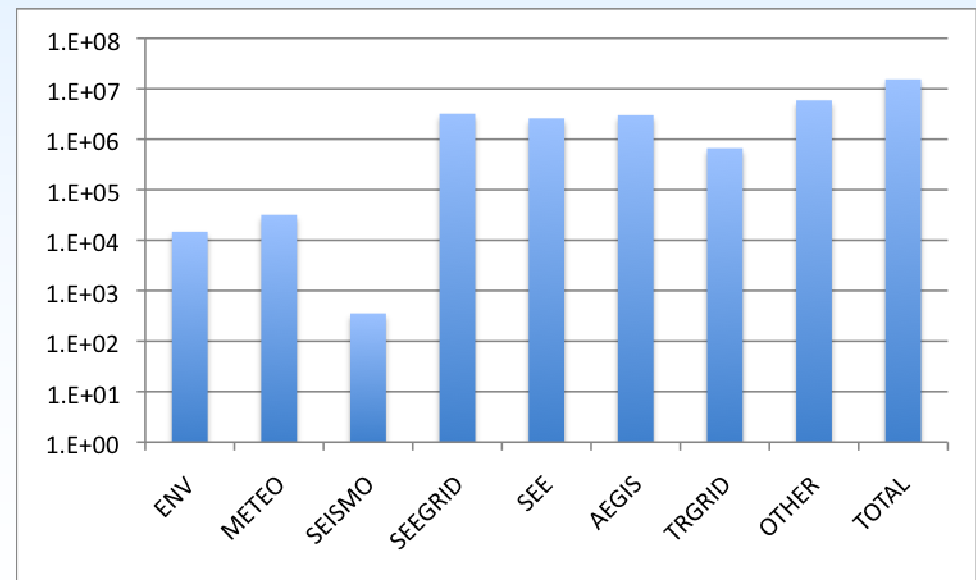


SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

- SEE-GRID-SCI sites have provided more than 15.4 million Normalized Elapsed CPU hours and more than 1.2 million number of jobs during Y1
  - SEEGRID VOs (seegrid, seismo, environ and meteo)
  - National VOs (AEGIS VO, TRGRID VOs)
  - Regional VOs (see)



Number of jobs



Normalized CPU time (hours)

# Key results: National Grid Initiatives

- Initiating and structuring NGIs
  - SEE-GRID pioneer of the NGI concept
  - Significant developments in a number of areas: organisational stability, operational maturity, national-level support
- Providing active support for NGI establishment in other developing regions
- Deliverables:
  - DNA2.1 - NGI metrics specification (44 detailed metrics, quarterly reporting and analysis)
  - DNA2.2 - NGI cookbook (detailed NGI setup guidelines, distributed to all other regions)
  - DNA2.3a - NGI intermediate assessment (detailed assessment and recommendations)

# NGI status



SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

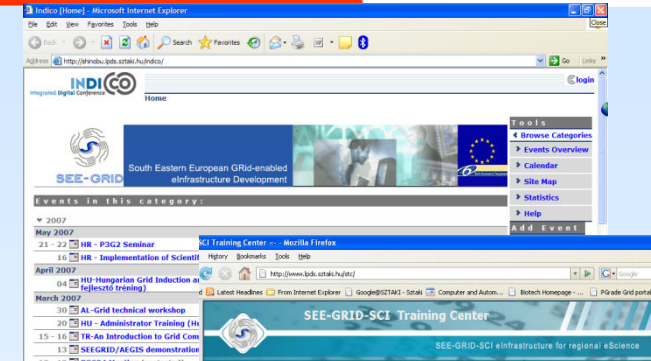


# Key results: dissemination and training



SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

- Dissemination Event Agenda:
  - project website see-grid-sci.eu
- Training Event Agenda
  - Training portal <http://www.lpds.sztaki.hu/stc>
- Core promotional material
- Additional brochures and posters
- Harmonization of the training material
- Training material repository
- Trainings: generic Grid and VO specific
- Training materials in local languages
- tInfrastructure: 18 sites, core services, mock CA
- A pool of 42 SEE-GRID-wide trainers, growing
- Quality assurance procedures



**SEE-GRID-SCI SEISMOLOGY - VIRTUAL ORGANIZATION**

OVERVIEW OF SEISMOLOGY VO

Seismology is the study of seismic phenomena that appear in the form of earthquakes and vibrations of the earth caused due to natural or artificial causes. Seismology also studies both instrumental resources to solve equations that arise in the mathematical modeling of seismic phenomena as well as storage resources to store and access historical seismological information and seismic activity data and without other continuous or integrated follow-up from several geographically distributed sensors.

Before the operation of the infrastructure, the two main sources of seismic data were data banks. Data providers would place their data in the archive, which would then provide either direct or download the data that were forwarded to clients. This was a very slow, tedious and expensive process. Today, the data dissemination process is becoming increasingly automated. It is necessary to have a flexible, reliable and manage all the data. Automation of this process may require the writing of various scripts which may be quite difficult to do so. Besides, it is necessary to ensure that the data are in the right format and protocol. There have been efforts also to make seismic data available by web services which will help to disseminate the data. This requires some data base and system programming and to link each service with their applications. Grids refer these and other services research and other research centers.

SEISMOLOGY VO provides an available of resources to high performance. All grid users are encouraged to use the infrastructure for access distributed data just like the existing local data with the help of middleware and other tools.

The SEE-GRID-SCI Seismology Virtual Organization is in line with the general vision of the SEE-GRID-SCI infrastructure to harmonize and integrate the resources from the South Eastern European countries. Besides, there are other organizations and projects working on related topics, such as the SEE-GRID-SCI infrastructure and Harmonization of the training material repository, which will be essential and complementary work with the infrastructure that is being developed by the SEE-GRID-SCI infrastructure.

SEISMOLOGY VO PLATFORM

In Seismology VO, data, tools, utilities, distributed applications, and VO member seismic waveform data from various South Eastern European countries are planned to be collected and stored.

To facilitate the use of Seismology VO, the following applications are implemented:

- 1. Data base server: retrieves data from different partner countries.
- 2. Local operations: including the update of data base.
- 3. Programming tools that will provide easy access to data.
- 4. CA/Middleware of various seismology applications: SCA, SCA/CA, PLS, SCA, MEDS, VO, and SCS.

**SEE-GRID-SCI e-INFRASTRUCTURE for regional eScience**

Infrastructure to enable new scientific collaborations among user communities and providing advanced capabilities to more researchers, with its in seismology, meteorology and environmental protection. The major and structuring effort on target user communities that from the available infrastructures.

SEE-GRID-SCI infrastructure to cater for demands of the computing and storage resources and avoiding new partner to help mature and stabilize the National Grid Initiatives in the new era of longer-term sustainable Grid infrastructure in Europe.

to contribute to the stabilisation and development of South-East of stimulating infrastructure development and adoption by innovative high-quality research across target scientific fields.

see-grid-sci.eu

# Dissemination and training



SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

- Large number of dissemination and training events
- Focused and widespread campaign:
  - 1 regional dissemination event (meteorology workshop)
  - 17 national level dissemination events
  - 2 regional training events (grid administrators; meteorologists)
  - 17 country-level training events (9 dedicated to Meteorology / Environmental Science / Seismology communities)
  - 290 people trained
  - 880 people attended focused dissemination events
- Other dissemination-related activities: 2 newsletters; disseminations in public media: TV (3); Press Releases (3); Newspapers (3); E-newspapers (3); scientific papers (30+) and presentations (40+).



**Regional event**





# Conclusion: Regional vision



SEE-GRID-SCI  
SEE-GRID infrastructure for regional eScience

- Vision:
  - Being on the technological par with the rest of Europe
  - Enabling local scientists to use their potential
  - Role-model for regional developments
  - Leading the way in wider contexts
- Strategic success metrics of regional initiatives:
  - not Gbps/sec; number of nodes; TBs of storage
- The initiatives are puzzle pieces of RTD 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
  - Improvement of regional competitiveness in all market sectors
  - Regional political stability and cohesiveness



- Join us at the User Forum in Istanbul!
- <http://sgsuf.ceng.metu.edu.tr/index.php?conference=sgsuf&schedConf=sgsuf09>