Earth Sciences

activity and perspectives

Jean-Pierre Vilotte and Monique Petitdidier
Institut de Physique du Globe de Paris
Institut Pierre et Simon Laplace

and G. Moguilny (IPGP), D. Weissenbach (IPGP)

Introduction

DataGrid (2002-2004) → EGEE I & II & III (2004-2010) → EGI (SA3) & NGI France Grille (2010-)

DEGREE (2008-2010)

EELA II (208-2010)

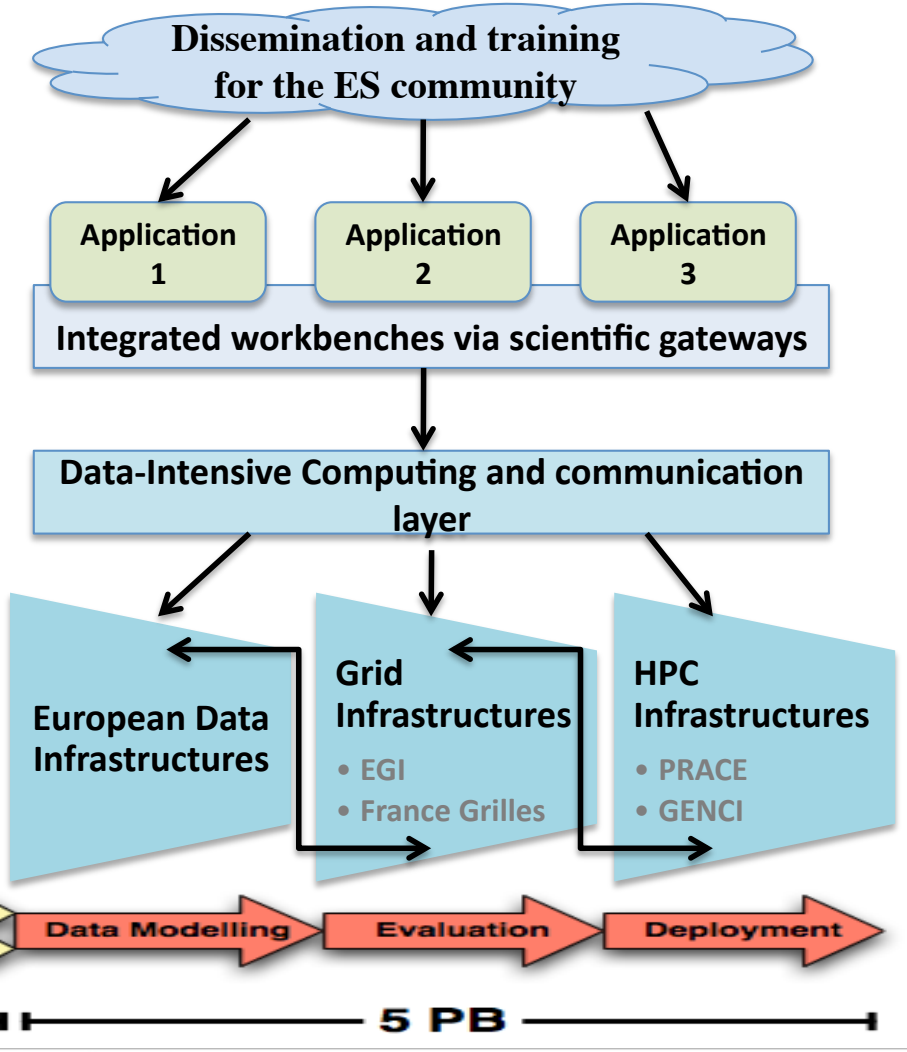
GISELA (2010 – 2012)

A continuous involvement in the Grid since 2001

- DataGrid (2002-2004) : IPSL with ESA and KNMI
- EGEE I & II & III (2004-2010) : IPGP & IPSL
  - Cluster Earth Science
  - MPI working group and Training activity
  - Organization committee of the EGEE conferences
- DEGREE (2006-2008): IPSL & IPGP
  - White paper
- EELA II (2008-2010) and GISELA (2010-2012): IPGP (seismology)
  - Chile, Perú and Argentina

Dissemination: booth and special sessions at the EGU conference

ES Data-Intensive Environment



ES Data-Intensive Applications

Data Integration and Data Analysis

- Multi-attribute data sets
- Exploration and visualization of distributed large volumes of data
- Integration of large volume and/or large number data sets from distributed data resources
- Distributed data analysis (HTC) of large data volumes (> 100s TB) -> New large data sets (~ PBs)
- Data-curation and Data-fitness

Data modeling

- Large scale parameter exploration (HTC)
- Large scale simulations (HPC) producing large volumes of synthetic data to be analyzed
- Large scale Inversion and assimilation (HPC) applications (high memory and storage capabilities)

Orchestration workflows

- Orchestration of Data analysis and Data modeling applications with efficient data transfers between Data, Grid and HPC infrastructures

VO Earth Sciences

Earth Science Research (ESR) : Generic VO

- Located since 2004 at SARA (NL)
- VO Managers: D. Weissenbach (IPGP), W. Som de Cerff (KNMI), K. Kassirer (SCAI)
- Coordination: M. Petitdier (IPSL) & H. Schwichtenberg (SCAI)
- 56 registered users, ~20 registered applications from IPGP, IPSL, LOA, BRGM, LGE/E. Mines Alès
- 50 sites in Europe (12 countries, ~ 20 000 CPUs)
- Atmosphere and Climate, Hydrology, Seismology, Environment

Related VOs (~7) in Europe

- Limited projects: CYCLOPS, SEEGRID; or with only a national dimension: in Ireland or Russia
- Integration to the ESR VO is under study within EGI

EGEODE – Geocluster VO (CGG Veritas)

- Since 2004 at CC-IN2P3, no more part of EGI and the NGI France Grilles
- VO managers: J.B. Favreau, G. Vetois and L. Scienini (GeoAzur)
- CGG seismic processing and imaging software
- ~30 users (?) – IPGP, GeoAzur, IPGS-EOST, ISTEP, UPMC-Sisyphe,, ISTEP-P6, ENS Paris

~ 4 450 000 h (CPU norm)

VO Earth Science Research

50 sites in Europe: 12 countries

France, UK, Italy, Greece, Ireland, Netherland, Germany, Slovakia, Bulgaria, Spain, Finland, Russia

2 more sites in China

IPGP/IPSL Grid Node

The only ES node in the French NGI France Grilles

- Moved in 2010 into the new IPGP infrastructures (Cuvier)
- Grid integration within the OSU ecosystem (CNRS-INSU)
- Site administrator: D. Weissenbach (IPGP)
- ES application support: D. Weissenbach (IPGP), G. Moguilny (IPGP)
- Coordination with emerging multi-disciplinary regional nodes of France Grilles
  Bordeaux, Grenoble, Lille, Montpellier …

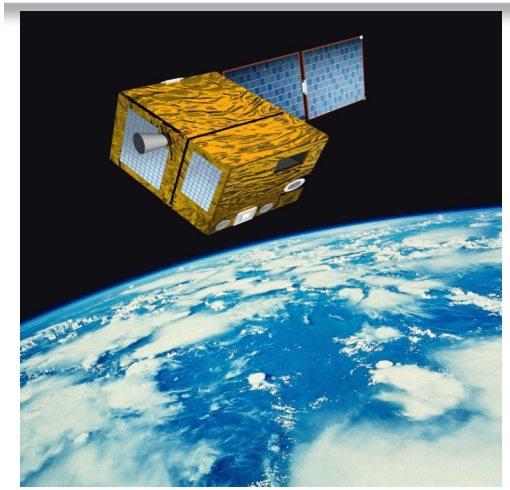
Middleware and Services

- EGI and France Grilles middleware
- Specific services on top of EGI middleware for ES Databases (e.g., GRelC …)
- Will support ES services for Data access and Data Integration (Data and HPC infrastructures)

Resource evolution and contribution to France Grilles

- Until 2010: 34 CPUs and 1 To storage
- 2010: 512 CPUs and 60 To ; 2012: 1024 CPUs and 560 To storage (PPF S-CAPAD and EQUIPEX France Grilles)

Satellite Observation: Atmosphere



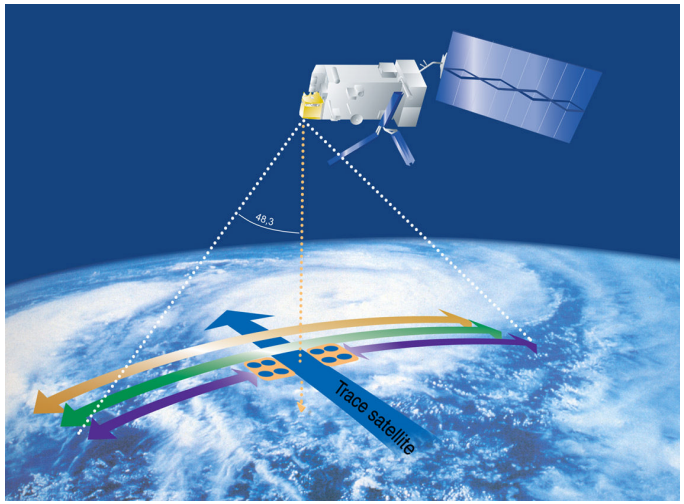
Comparison with the data from the Lidar on Calipso and Modis algorithm validation

- Full resolution processing of the Parasol data on EGEE by the LOA (University Lille): F. Ducos
- Support: D. Weissenbach (IPGP)
- 12,4TB in input →1,7TB output, 3-4 months instead of ~18 months

Infrared Atmospheric Sounding Interferometer

Studies of photochemical pollution process in the low and middle troposphere

- Analysis of the twice-daily measures on EGEE by IPSL/ LISA: M. Eremenko
- Support: D. Weissenbach (IPGP)

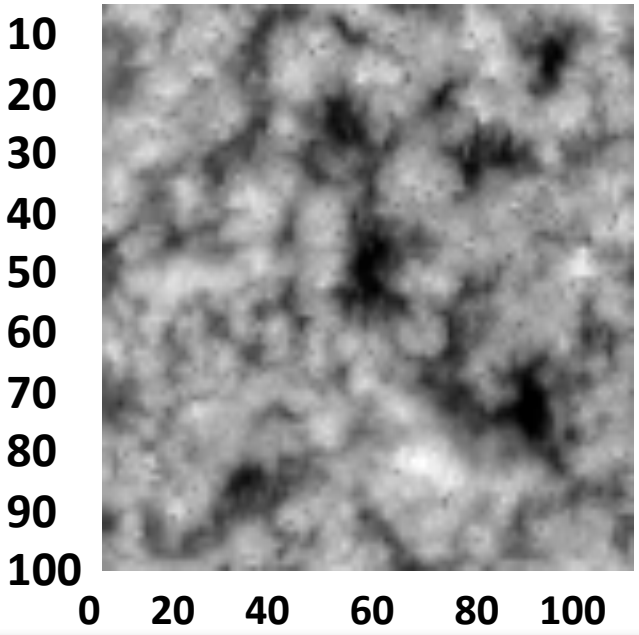


Spatial Teledection: Radiation Transfer Model

Radiation Transfer Model 3DMCPOL (LOA):

Evaluation of the clouds 3D effects on the radiation measured by the A-train sensors (French-American 6 satellite constellation)

- Monte-Carlo method very efficient on the Grid
- Project funded by the PNTS program (CNRS-INSU), PhD fellowship by the DGA and the CNRS



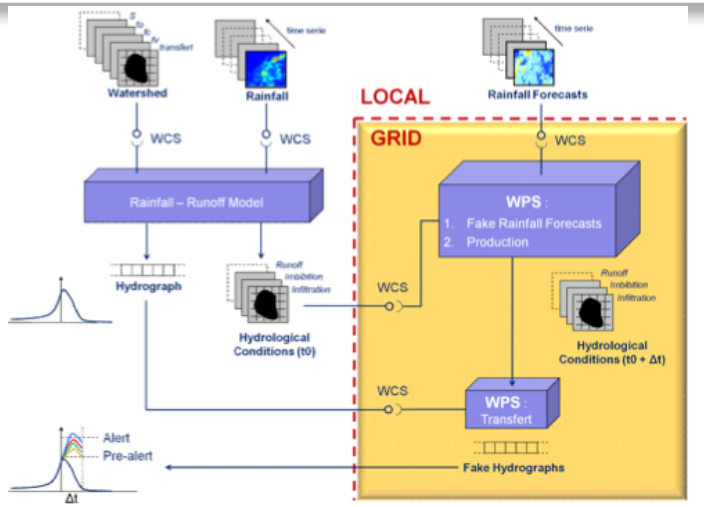
Simulation d'une luminance visible renvoyée par un nuage de type strato-cumulus et mesurée par un radiomètre de type MODIS ou POLDER embarqué sur un satellite.

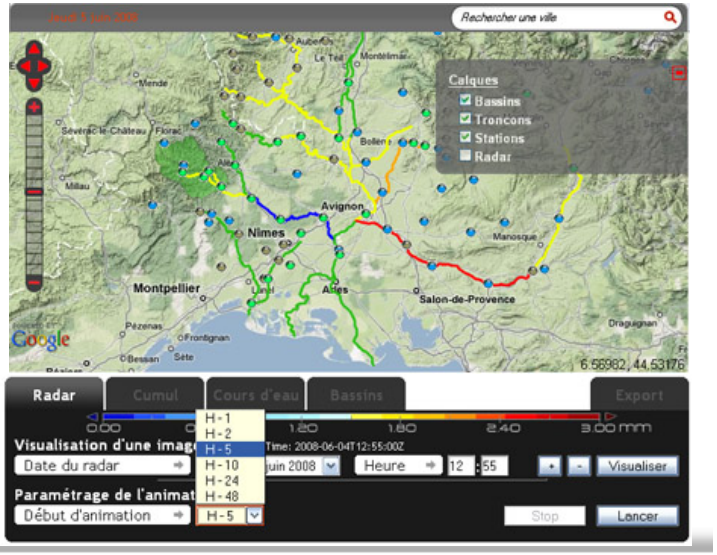
LOA (Universté. Lille 1) : C. Cornet

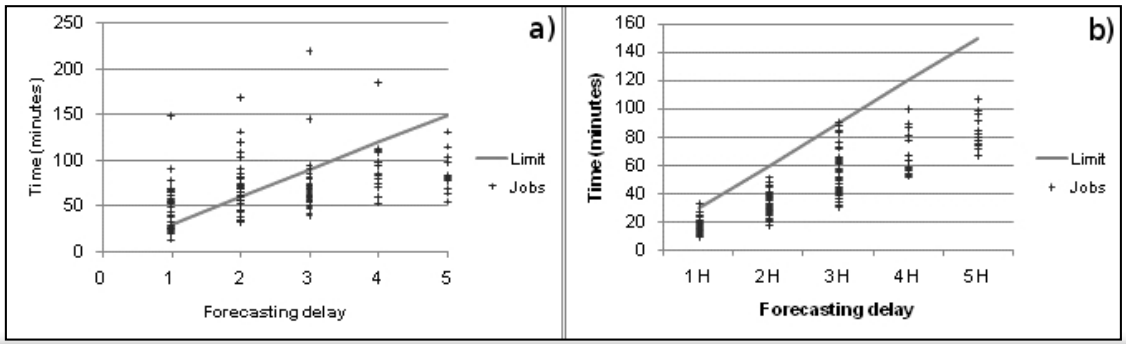
Hydrology: G-AHLTAÏR platform

G-ALHTAÏR: a fast flood prediction platform

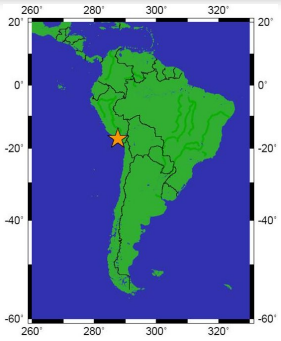
- Generate hydrological previsions in real time
- Platform for operational applications (Service de Prévision des Crues SPC, Service Central d'appui à la Prévision des Innondations SCHAPI, bureaux d'étude)
- *Contribution à l'amélioration de l'expertise en situation de crise par l'utilisation de l'informatique distribuée : Application aux crues à cinétique rapide, V. Thierion, Thèse de Doctorat, Ecole des Mines de Paris, 330p.*
- Ecole des Mines d'Alès: V. Thierion, P.-A. Ayral

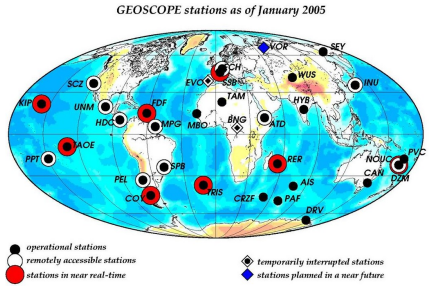


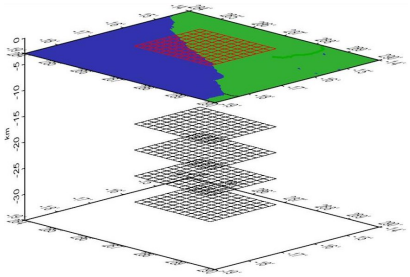


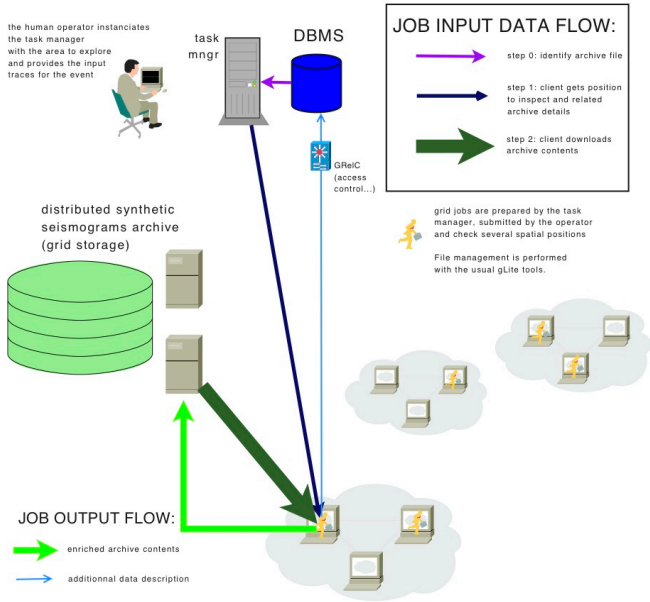


gCSMT: Grid earthquake CMT determination



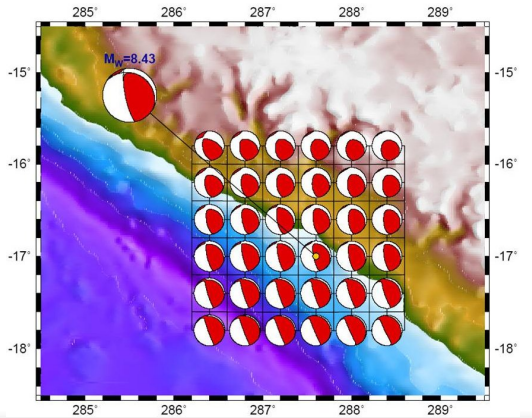






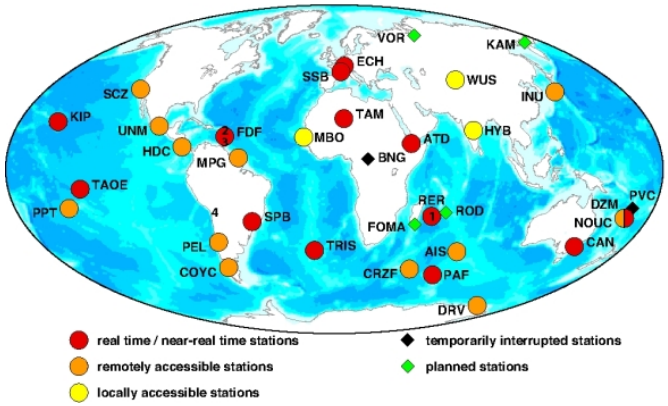
gCSMT: rapid seismic moment tensor solution at the centroid of great earthquakes (M > 5.5)

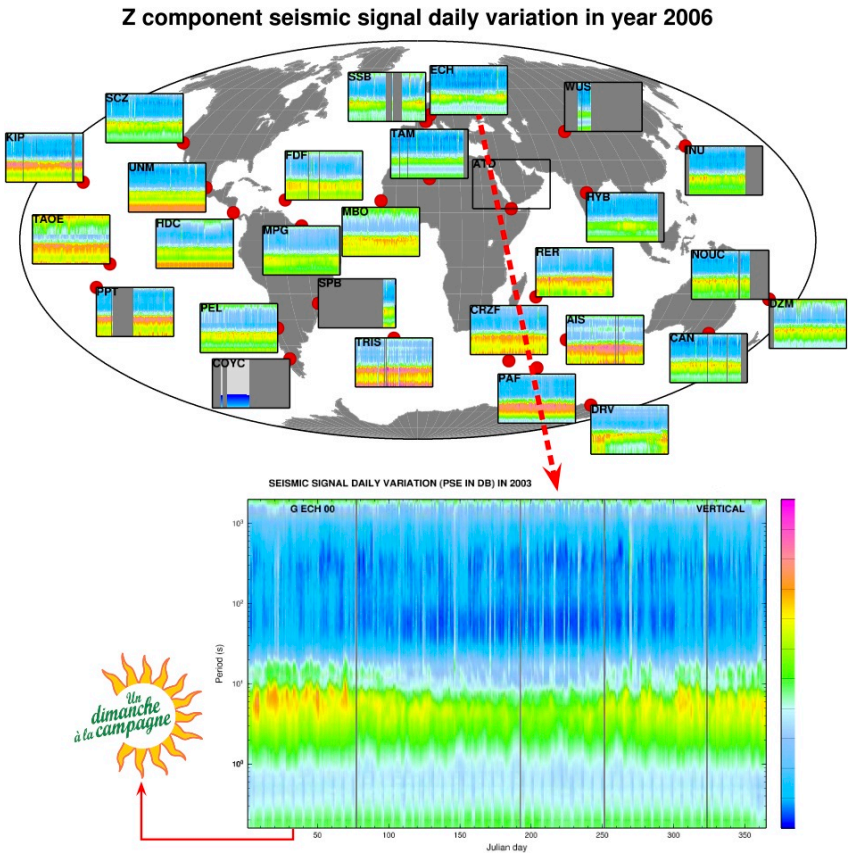
- Operational platform in production in the French data centre GEOSCOPE: E. Clévédé (IPGP), G. Patau (IPGP)
- Embarrassingly parallel inversion with parameter space exploration
- Synthetic seismograms database dynamically built and stored on the Grid (GRelC data service software)
- Support: D. Weissenbach (IPGP)
- Operating since 2004, and continuous developments since



Noise Analysis of continuous seismic data

Data curation: Daily noise average analysis of the continuous seismic records archived in the Geoscope Data centers

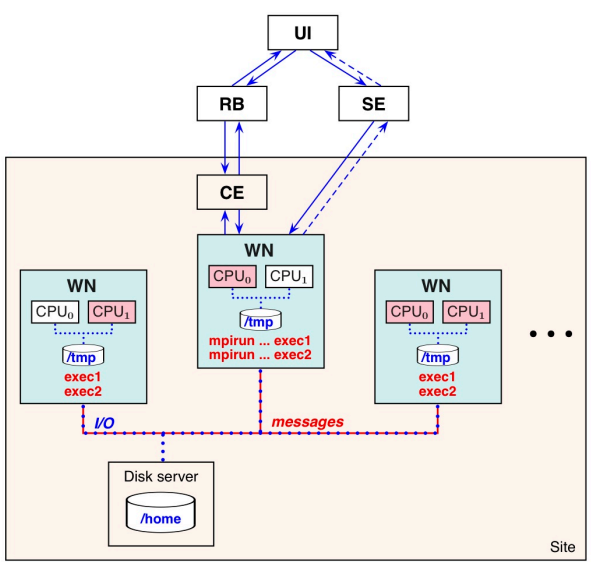


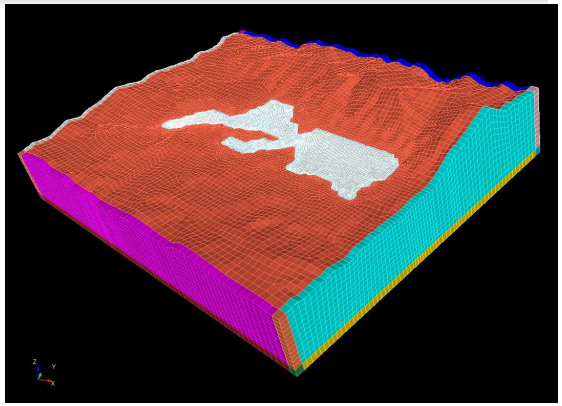


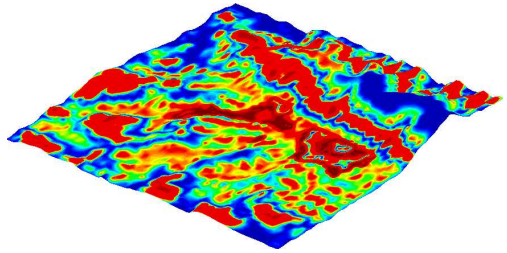
- 30 years of continuous records (~30 stations, ~2 TB)
- Fortran + Octave
- E. Stutzman (IPGP)
- Support: D. Weissenbach (IPGP)

MPI ground motion prediction on the Grid

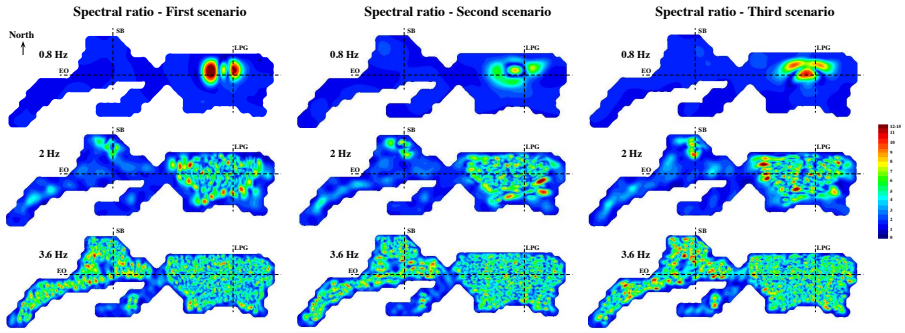
SENUM3D: Simulation of 3D wave propagation and ground motion prediction



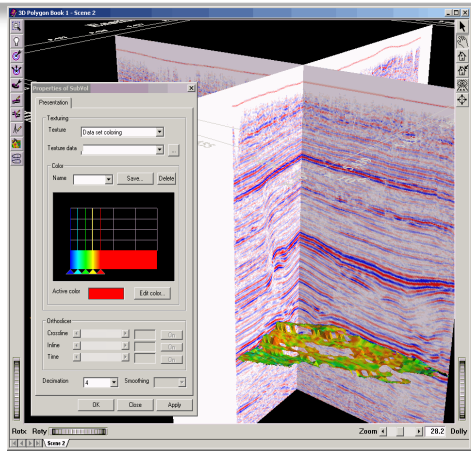




- MPI application
- Parametric exploration for assessing response variability
- E. Delavaud and J.P. Vilotte (IPGP)
- Support: G. Moguilny (IPGP)

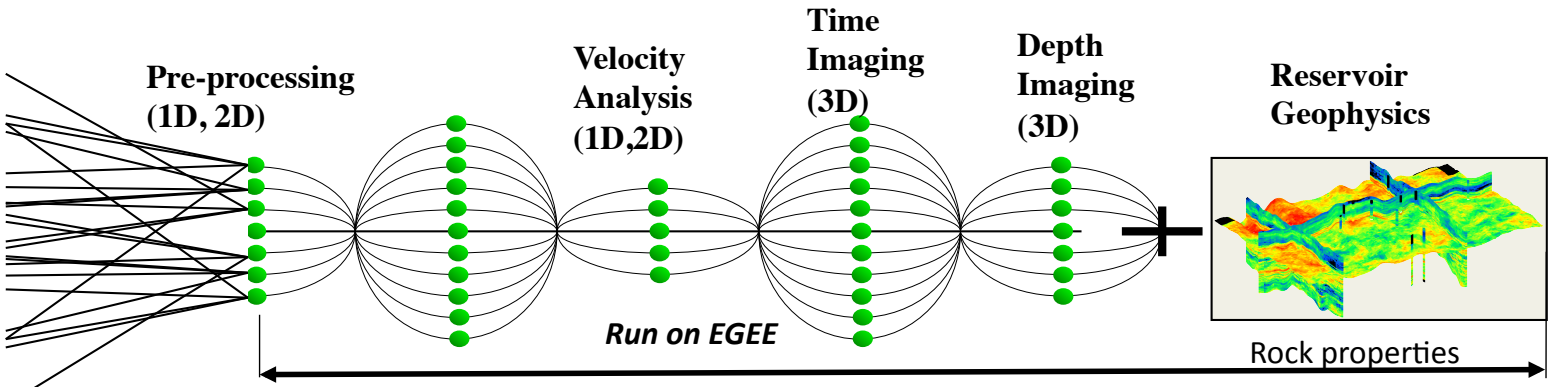


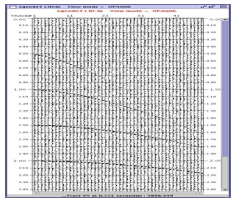
EGEODE Applications: What Next ?

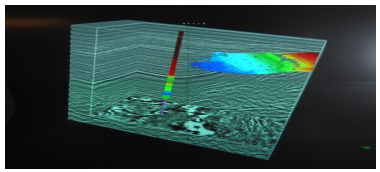


Seismic processing generic platform for research and education

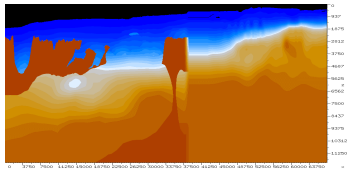
- Based on Geocluster (CGT), an industrial application used by CGGVeritas
- 6 CNRS laboratories: ISTEP (Paris 6), UMR Géoscience Azur, UPMC – Sisyphe, IPG-EOST Strasbourg, IPG Paris, ENS Géologie
- No more on EGI and France Grilles



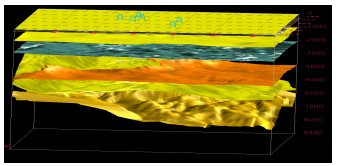




3D seismic



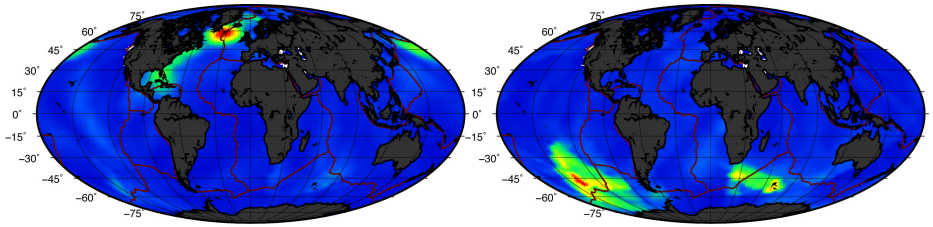
Velocity model



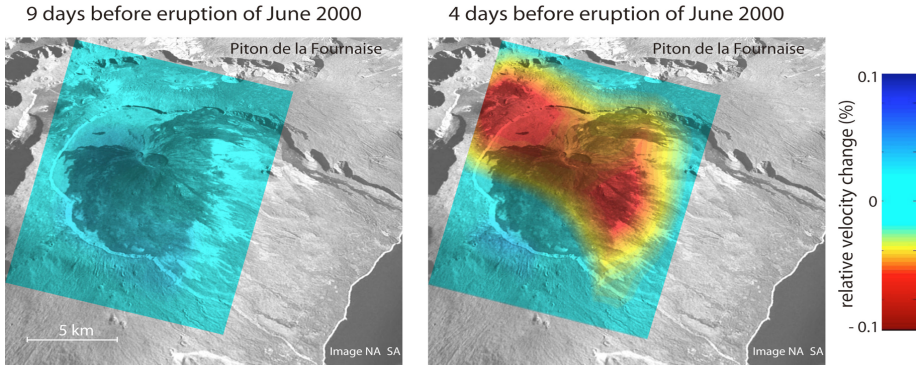
Structural model

ES new applications: Seismology

Studying the coupling between the Solid Earth, the Oceans, and the Atmosphere



Seismic noise correlations: observing precursors to volcanic eruptions

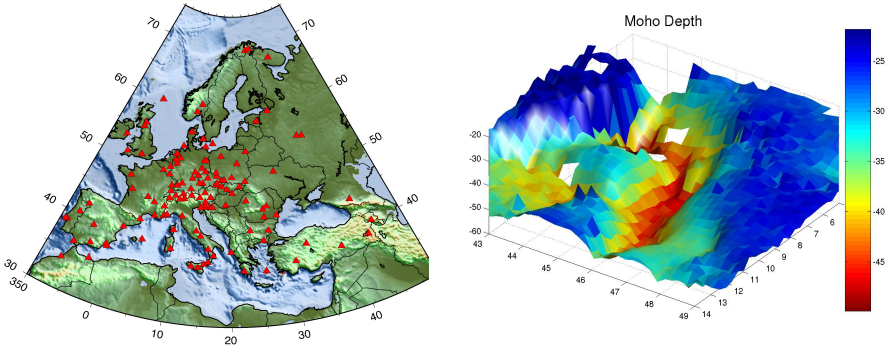


Imaging and monitoring based on correlations and ambient seismic noise

CNRS-INSU: LGIT Grenoble, IPGP
European ERC: project WHISPER
ESFRI project: EPOS-PP
TGIR RESIF

M. Campillo (LGIT), N. Shapiro (IPGP), F. Brenguier (IPGP)

Noise-based seismic tomography: Application to Alps



New ES applications: Climatology

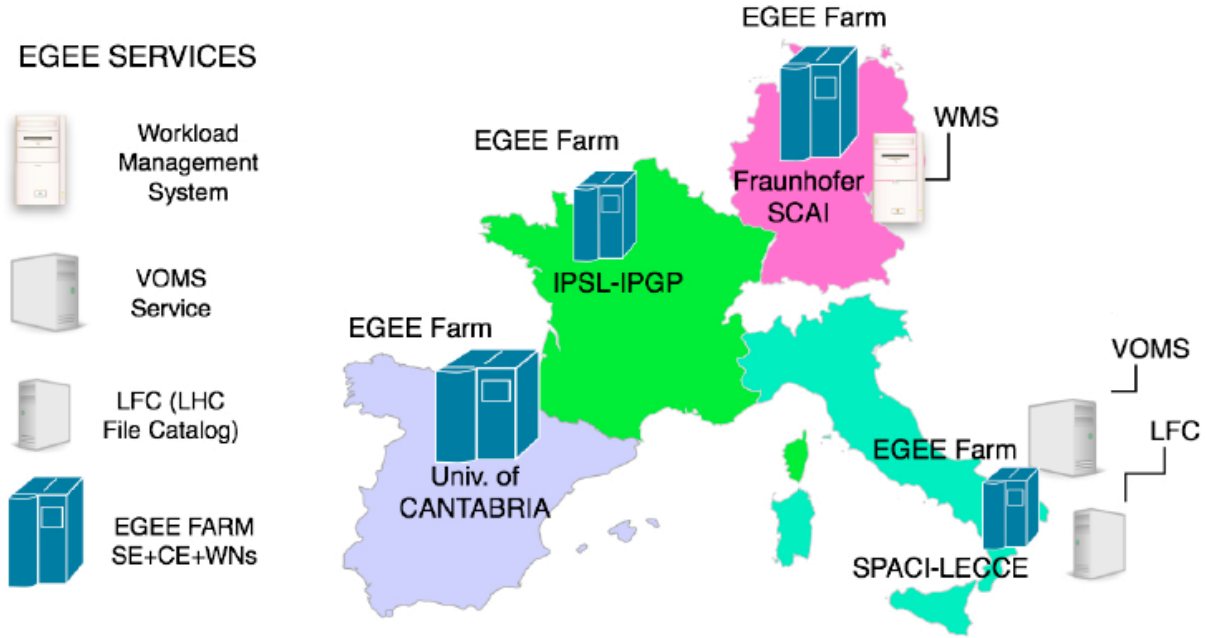
Increasing interests for the Grid

- For data analysis and sharing in relation with the "Earth System Grid" used for the climate data management and analysis within the GIEC framework
- Investigation of a solution for the IPSL "climate" data centers in the framework of EGinsPIRE (GRelC and Gensi.DEC) & ESG
- Deployment of a solution and evaluation for data analysis computing of the climate data

Test bed deployed for evaluation of the G-climate data distribution

This solution was not selected in the GIEC

Support: D. Weissenbach (IPGP)



ES Conclusions and Perspectives (1)

ES community in France Grilles:

- Contribution to the NGI France Grilles infrastructure **should be wrapped with the Computing and Data Infrastructures of the "Observatoires des Sciences de l'Univers" (OSU) of the CNRS-INSU inserted in the PRES ecosystem**
- In addition to regional multi-disciplinary NGI nodes, it is **essential to have one or more ES NGI nodes to evaluate and deploy services and gateways that meet the ES applications needs** bridging the ES Data infrastructure to the Grid and HPC infrastructures in France and Europe
- ES research infrastructures are intrinsically Global at international (US, Japan) and European scales requiring an active European ESR VO coordinated with US and Asian initiatives

ES emerging Grid related scientific production

- 15 papers in international journals
- 20 extended abstracts and reports
- 10 PhD
- Special sessions at the EGU and AGU annual conferences, booth at the EGU

ES conclusions and perspectives (2)

The Grid is now in the ES research landscape in France and Europe:

- Data integration and data analysis of large volumes
- Data intensive simulation, inversion and assimilation

ES community challenges:

- A service-oriented architecture integrating Data, Grid and HPC infrastructures
- Lower the barrier of uptake through scientific gateways

ES European community integration:

- ES ESFRI projects: Atmosphere (ICOS, IAGOS, EISCAT3D); Ocean (EUROARG, EMSO), Solid Earth (EPOS-PP) …
- IT initiatives: Environment (ENVRI), Seismology (VERCE), Data infrastructures (EUDAT) …
- EGI and PRACE and other standardization bodies

National infrastructures:

- New ES TGIR projects (e.g., RESIF …)
- ES Implication in the national EQUIPEX projects (GENCI, France Grilles)