



Enabling Grids for E-science

Applications de Science de la Planète dans EGEE

Monique Petitdidier
IPSL/CETP

Coordinatrice du domaine Science de la Planète
monique.petitdidier@cetp.ipsl.fr

www.eu-egee.org

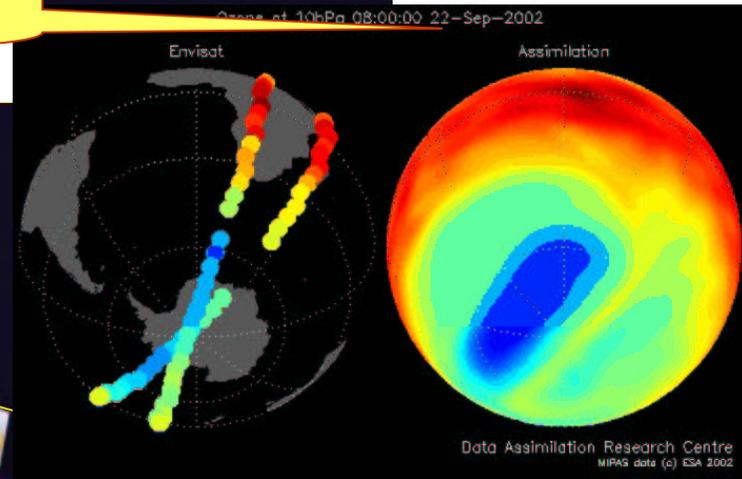


Information Society
and Media



GOME total ozone assimilation

Stratospheric Ozone

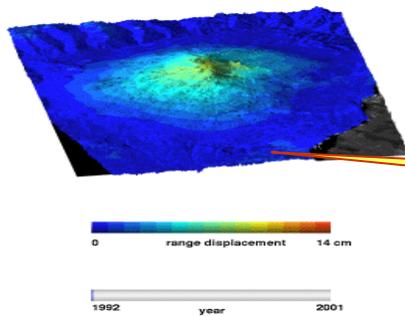


Topography & Motion

Land cover & vegetation

Marine SST, SSH& colour

10 y displacement of Etna 1992-01



M. Petitdidier – coordinatrice ES

- **ESR (Earth Science Research)**

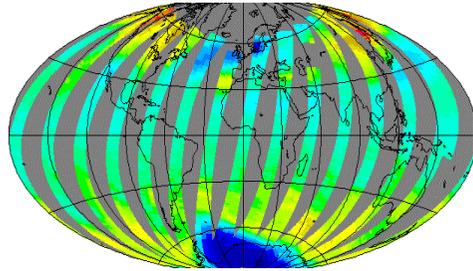
VO Managers: W. Som de Cerff (KNMI), D. Weissenbach (IPSL)

- ~50 personnes, Académie, Organisation, industrie pour tests
- FR, BG, CH, DE, IT, NL, RU, SK, SP
- Critère: application portable, apport de ressources
- Collaboration avec EUMEDGrid, EELA, EUChinaGrid, SEEGrid
- Aide au portage d'applications, développement d'outils
- Partage du savoir

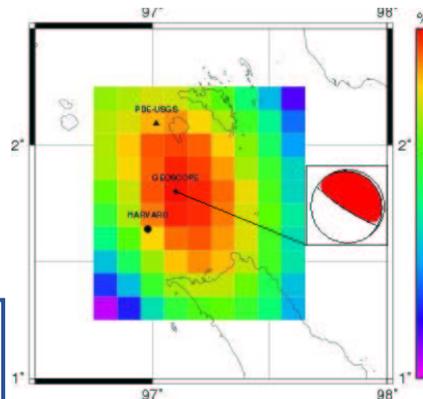
- **EGEODE (Expanding Geosciences On Demand)**

VO manager : Gael Youinou, G. Vetois

- ~30 personnes CGGVeritas + CNRS/INSU
- Geocluster
- e-collaboration



ESA, UTV(IT),
KNMI(NL), IPSL(FR)-
Production and
validation of 7 years
of Ozone profiles from
GOME

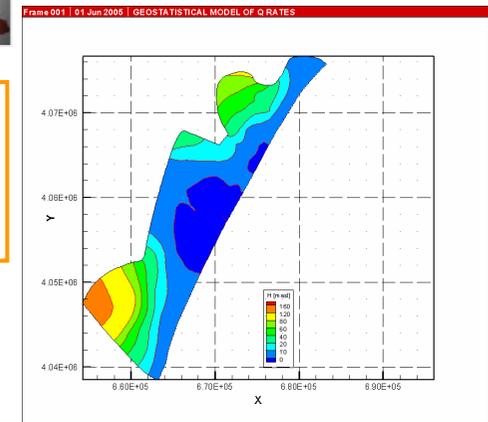


Rapid Earthquake
analysis
(mechanism and
epicenter)
50- 100CPUs
IPGP(FR)



Flood of a Danube river-
Cascade of models
(meteorology,hydraulic
,hydrodynamic....)
UISAV(SK)-

Geocluster for
Academy and
industry CGG(FR)-



Data mining
Meteorology &
Space Weather
(GCRAS, RU)

Modelling seawater
intrusion in costal
aquifer (SWIMED)
CRS4(IT),INAT(TU),
Univ.Neuchâtel(CH)-

DKRZ(DE)- Data access
studies, climate impacts
on agriculture

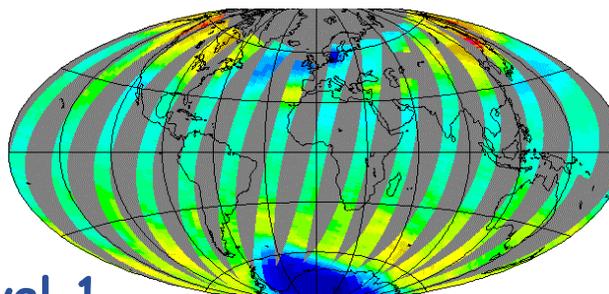
Specfem3D:
Seismic
application.
Benchmark for
MPI (2 to 2000
CPUs) (IPGP,FR)

Air Pollution
model- BAS(BG)

Mars atmosphere
CETP(FR):

- **GOME/ERS – Production de profils d’ozone et validation par des données Lidar**
 - ESA, UTV, KNMI, IPSL
 - 2 publications dans JQSRT
- **GEOCLUSTER – plateforme sismique développée par CGGVeritas**
 - Mis à disposition des utilisateurs INSU: IPGP, ENS
- **ELMER (CSC – Finlande)**
 - **Modèles physiques de Mécanique des fluides, Mécanique structurelle, Electromagnetisme, Transfert de chaleur ,Acoustique**
 - Mis à disposition sur EGEE pour d’autres utilisateurs
- **SPECFEM3D**
 - En projet pour les utilisateurs

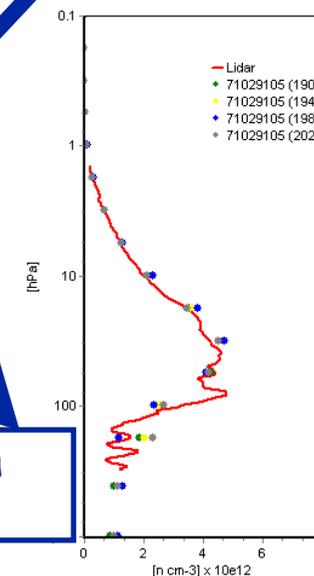
Raw satellite data from the GOME instrument (~75 GB - ~5000 orbits/y)



Level 1
(example of 1 day total O₃)

ESA(IT) - KNMI(NL)
Processing of raw GOME data to ozone profiles.
2 alternative algorithms
~28000 profiles/day

IPSL(FR)
Validate some of the GOME ozone profiles (~10⁶/y)
Coincident in space and time with Ground-Based measurements



Level 2

Visualization & Analyze

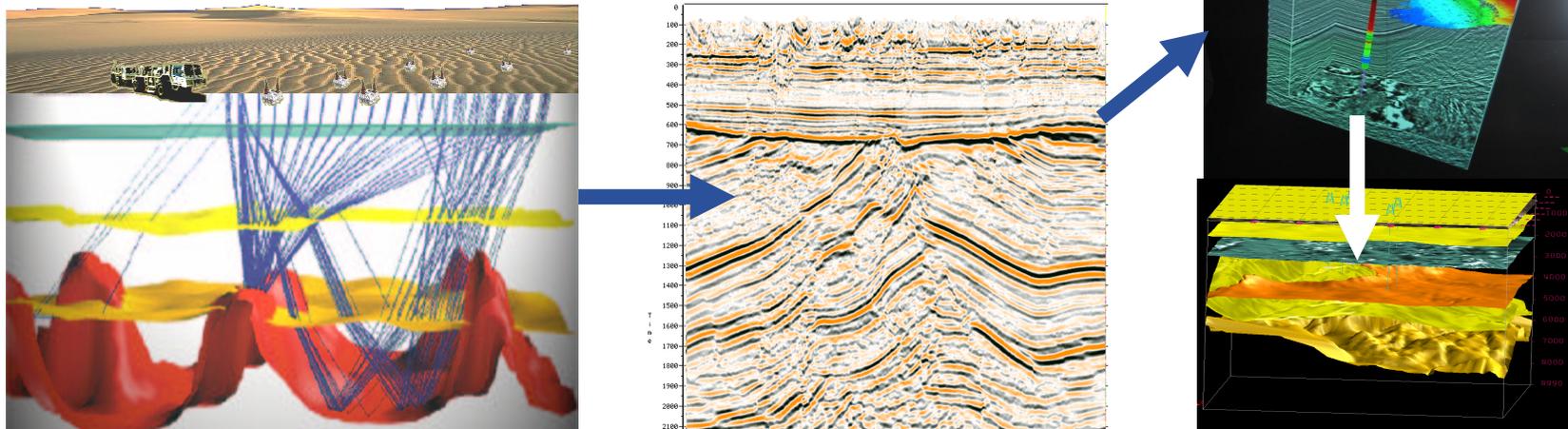
Meta Database server

PosgreSQL – geospatial search

EGEE environment

- **7 années de données, 14,5 orbites/jour**
- **Algorithmes:**
 - Réseaux de Neurones, NNO, (ESA, UTV) using IDL - 2 versions
 - Inversion Algorithm (KNMI) –données O3 climatologie, CEPMMT..
- **Données Lidar (NSDC)**
 - 7 stations au maximum (IPSL)
- **Nombre de fichiers: 70000 pour les 2 versions de NNO**
- **Développement commun**
 - une base de métadonnées sur un serveur sécurisé
 - Recherche par Geolocalisation temporelle et spatiale des orbites passant au-dessus d'une station lidar avec PosgreSQL)
- **Résultats: 2 articles, une thèse**
 - Cas unique de validation de tout un jeu de données satellite avec les données sol disponibles
 - Une fois l'application mise en place on peut valider très rapidement toute nouvelle version de l'algorithme

- Plateforme sismique generique developpee par CGG (400 modules) portee sur EGEE pour la recherche (R&D Industry + Academy) inclue des outils standard pour le traitement des données, la simulation et l'inversion (modèle d'optimisation).
 - Serveur de license pour autoriser les utilisateurs
- Négociation avec l'INSU pour accès à Geocluster depuis une dizaine d'années
- **Mis à disposition de ces utilisateurs français de Geocluster (IPGP, ENS, ENSG, Geoazur, LGS...) + autres**
 - Dernière Version
 - A plus grande échelle due aux moyens de calcul de la grille



Exploitation et Production de Grands jeux de données

- **GOME**
- **Geoscope:**
 - étude du bruit sur les 25 années de sismographes
- **Etude de l'évolution de pesticides dans le sol**
 - BRGM –Orléans projet Européen Footprint
- **Traitement d'une base de données sur les nuages (CETP-ABM)**
 - En projet

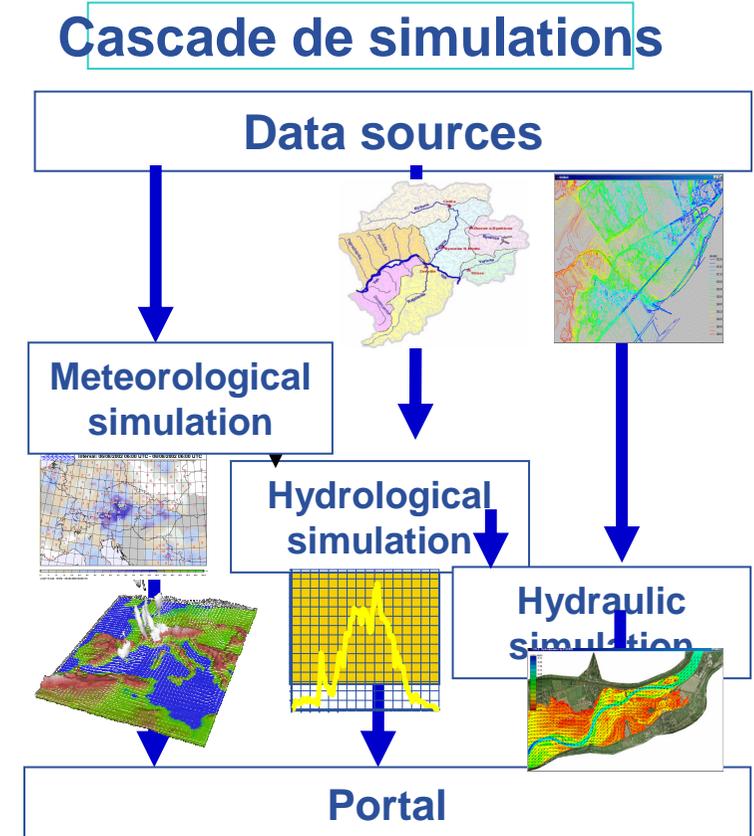
BRGM: I. Dubus, F. Boulayia, F. Dupros

- **Projet Européen avec comme coordinateur le BRGM**
Evaluation du risque de contamination des ressources en eau par les pesticides à différentes échelles de travail.
- **Scénarios climatiques (base de données Meteo France), différents types de sol, et 100 pesticides**
 - ⇒ **12 millions de run 1-2h chacun**
 - ⇒ **10 Toctets – fichiers de sortie (stockés au BRGM)**
- **Interêt de la grille:**
 - **fonctionne 24/24 et 7 jours/7**
 - **Possibilité de lancer dans un temps court des millions de jobs (CF Biomed-Wisdom) –les sites doivent être prévenus.**

- **Crue des fleuves et rivières**
 - Développement d'une plateforme sur EGEE avec cascade de modèles, workflow flexible par l'IISAS (Slovaquie)
 - Collaboration IISAS dans le cadre d'un projet INTAS (CNES, Ukraine et Russie) pour le transposer à des rivières ou fleuves en Ukraine
 - Collaboration IISAS avec l'Ecole des Mines d'Alès (Vincent Thieron)
 - Grand intérêt du projet Européen CYCLOPS (Cyber-infrastructure for Civil protection Operative ProcedureS)
- **Intrusion d'eau de mer dans les aquifères côtiers**
 - CRS4 (Italie), Univ. Neuchâtel (Suisse), INAT (Tunisie), France
 - Très grand intérêt d'autres pays méditerranéens (EUMedGrid)

L. Hluchy, Viet Tran, M. Ciglan ... (II-SAS, Bratislava Slovaquie)

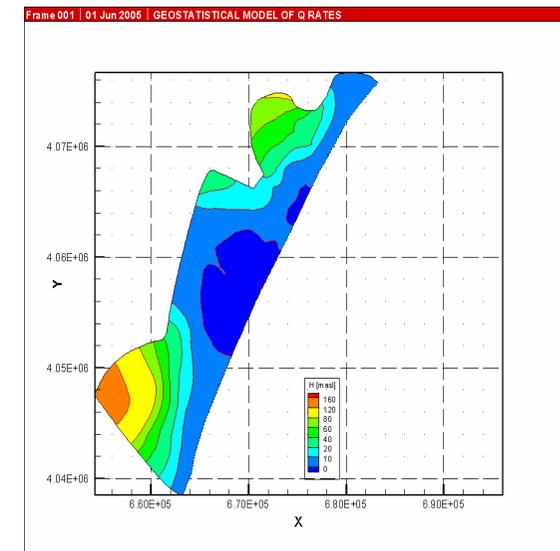
- Prévision des crues d'un affluent du Danube;
- Données expérimentales :conditions aux limites météorologiques, réseau de rivières, Carte numérique de terrain
- Modèles de météorologie ALADIN (MPI-parallel), MM5 (MPI-parallel)
- Hydrologie HSPF (sequential-parametric), NLC (sequential-parametric),
- Hydraulique. DaveF (MPI-parallel), FESWMS (MPI-parallel)
- Sortie: Prévision du temps, des précipitations, hydrographie, niveau de l'eau et vitesse d'inondation
- Gestion des opérations par workflow dynamique



G. Lecca (CRS4 IT), P. Renard (Unine, CH), J. Kerrou (INAT, Tunisia), R. Ababou (IMFT, Fr)

Détermination de cartes de probabilités d'intrusion d'eau de mer

- **Algorithme: CODESA3D**
 - Modèle simulation de Monte Carlo basée sur un modèle de flux d'eau souterraine dépendant de la densité et de transport de sel
 - 100 simulations de Monte Carlo
- **Données:**
 - Géologie, Topographie, Météorologie, eau extraite par le fermier, propriétés de l'aquifère (carte du sol, utilisation du sol)



Intérêt de la grille de calcul

- Améliore le partage des données avec les géologues, modélisateurs et gestionnaire des ressources en eau
- Permet aux gestionnaires de l'eau en Tunisie d'explorer l'impact de leur décision par utilisation de la grille à partir d'une interface web :

EUMEDGrid

- **Contrôle et optimisation du temps d'exécution de 50 à 100 jobs soumis simultanément**
 - Caractéristiques des tremblements de terre
- **Jobs paramétriques**
 - Géomorphologie
- **Chaînage de jobs**
 - Ozone polaire
 - Jobs longs (>24h)
- **MPI**
 - SPECFEM3D

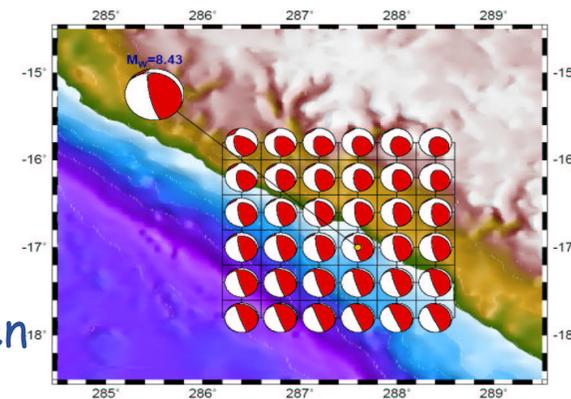
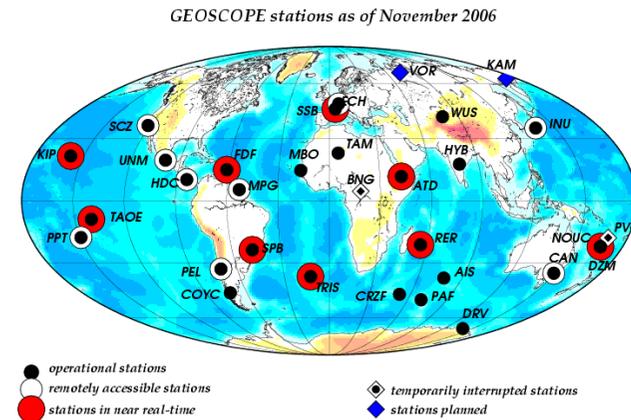
- IPGP: E. Clévéde, G. Patau

Application fonctionne sur alerte

- Données: 30 stations de GEOSCOPE
- Sélection des stations et des données
- Définition d'une grille 3D + temps centrées sur l'estimation de l'épicentre
- En chacun des points calcul de la fonction de Green, un job regroupe plusieurs points => ~ 50-100 jobs simultanés

Résultats obtenus ~6hr après le tremblement de terre et publiés sur le site de Geoscope

Tous les tremblements de terre majeurs étudiés en 2006: 21/24



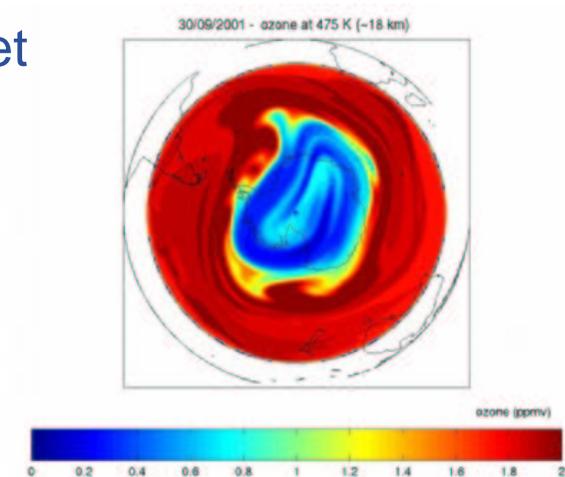
- S. Godin-Beekman (IPSL), C. Boone (IPSL), J. Gazeaux (IPSL)

Objectif:

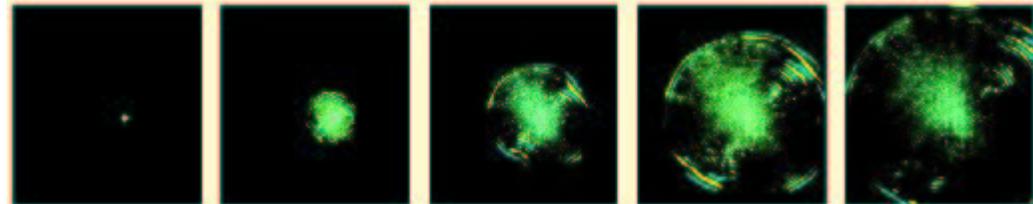
A partir de données, de sorties du modèle ECMWF et de simulation déterminer pour la perte d'ozone dans les régions polaires pour 10 ans

- **Caractéristiques:**

- Mémoire vive 2,3Go
- Long job (6-7 jours) Chaînage
- Confidentialité des données



- Résolution des problèmes de propagation d'ondes sismiques à l'échelle régionale pour modéliser la propagation d'ondes à hautes fréquences dans des structures géologiques complexes en utilisant une méthode aux éléments spectraux.
- Application d'abord écrite par D. Komatitsch (Université de Pau), utilisé dans plus 75 laboratoires au monde, en particulier pour des études sismiques, Portée par G. Moguilny (IPGP)
- Application utilisant F90 + MPI
 - Peut fonctionner de 2 à 2000 CPUs (Earth Simulator)
 - Nécessite les « homes » montés en NFS,
 - demande le lancement de 2 mpirun sur les mêmes noeuds, alloués dans le même ordre
- **Projet: être distribué sur EGEEII aux utilisateurs autorisés**

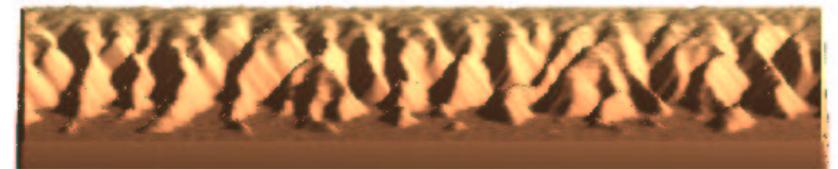


- IPGP: C. Narteaux, O. Rozier
 - Comprendre la formation et l'évolution des terrains (érosion des montagnes, dynamique des dunes...)
 - Algorithme : 3D Automaton Cellulaire avec différents types d'interaction entre cellules voisines (érosion, déposition, transport)

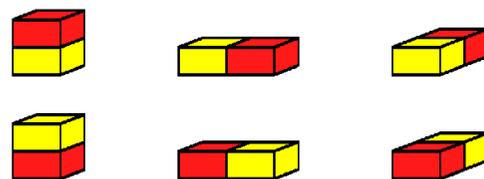
Tien Shan (Chine)



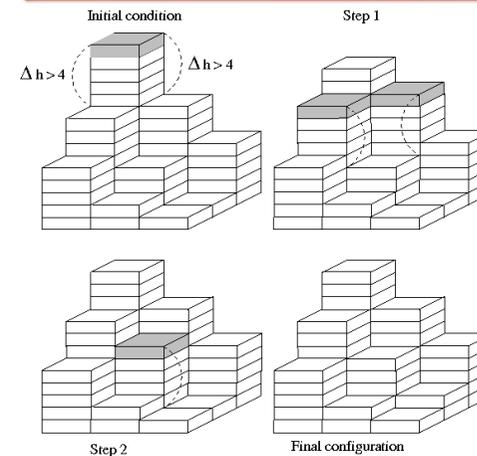
Modèle



Physical processes are associated with transition of doublets



+

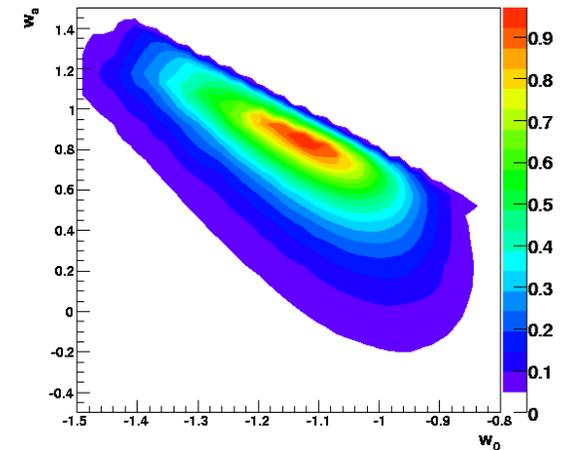


avalanche mechanism

- **Mars : Etude en fonction de différents modèles**
Non-thermal source of hot oxygen in the Martian upper atmosphere
 - JGR Planète
- **Modèle de pollution à grande échelle (BAS, Bulgarie, NERI, Danemark)**
 - Modèle Eulérien Danois – MPI
- **Climat (DKRZ)**
 - Gestion des metadonnées pour accéder aux sorties de modèle
 - Calcul sur les sorties
- **Data mining sur des sorties du modèle NCEP (GCRAS, Russie)**
- **Modèle sismologique pour la région de Thessalonique (Grèce)**
- **Cosmologie (IN2P3-Marseille, Annecy, + Chinois)**

CPPM: A. Tilquin, Zuxian Qian Univ. Beijing, China EU-ChinaGrid

- Goal: Solve the dark energy enigma
- How: Combine various cosmological probes
 - (SN1a) Supernovae of type 1A
 - (CMB) Cosmological Microwave Background
 - (BAO) Baryon Acoustic Oscillation
 - (WL) Weak Lensing
 - Large Scale Structure...
- Method : Full likelihood computation (Minuit)
 - Full simulation of probes
 - Frequentist statistic
- Problem : Construction of a confidence level contour take 2000 hours
- MPI Jobs



- **EGEE adapté pour:**
 - Partage de données et d'algorithmes
 - Exploiter et produire de grands jeux de données
 - Statistiques, calcul intensif – faire tourner beaucoup de fois un job
 - Job sur alerte - CPUs disponible
 - Possibilité opérationnelle de gestion des risques
 - Modélisation MPI
 - -
- **Difficulté**
 - Job très long conçu pour tourner en séquentiel sur un calculateur particulier
 - Beaucoup de mémoire vive
 - Utilisation de certains softwares sous licence (RESPECT)

Ce qui est en projet pour EGEEIII par Earth Science

- Rendre générique des outils déjà existant
- Interface générique avec les web services
- Quelques interfaces existent déjà avec les web services GIS
- E-collaboration
- Workflow dynamique pour applications complexes

Pour que la nouvelle génération de middleware Grille soit adaptée aux besoins spécifiques des applications des Sciences de la planète: le projet DEGREE

- Strategic objectives
 - Bridge the ES and GRID communities throughout Europe
 - Ensure that ES requirements are satisfied in next Grid generation
 - Ensure the integration of emerging technologies for managing ES knowledge

The DEGREE team:

IISAS, Slovakia
(Coordinator)

CNRS (IPGP+IPSL), France

KNMI, The Netherlands

UNINE, Switzerland

CRS4, Italy

SCAI, Germany

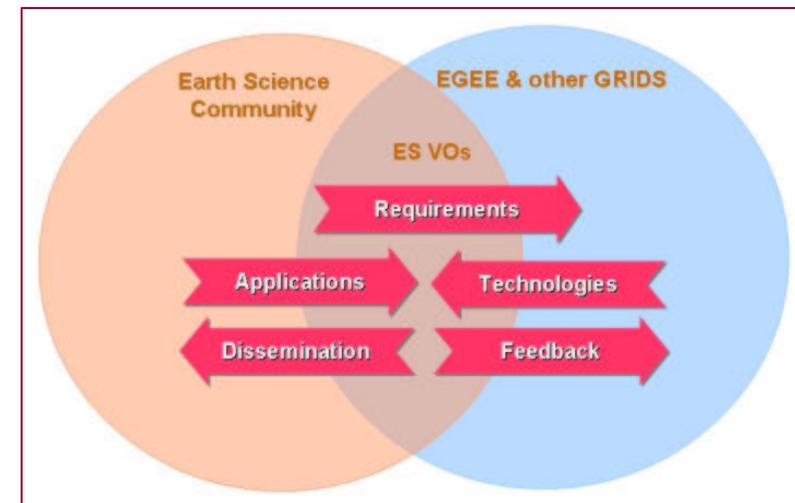
GCRAS, Russia

ESA-ESRIN, Italy

CGG, France

Dutch Space, The Netherlands

<http://www.eu-degree.eu>



Project Vision

Build a bridge linking the ES and Grid communities

La Collaboration interactive entre les équipes permet d'avoir plus de temps pour les découvertes

La grille ouvre de nouveaux champs de recherche et repousse ainsi les frontières de la connaissance