

# Earth Observation Virtual Organisation applications: a proposal to EGEE

Monique.Petididier@cetp.ipsl.fr  
Luigi.Fusco@esa.int

EGEE NA4 Generic Application Kick-off meeting  
Paris, 18 December, 2003

# Summary

1. The EO VO in EDG
2. Ideas and initial applications for participating in EGEE
3. The ESA EO Grid plans and ESA interests in supporting the EO science community in EGEE

# Description of EO Applications

- **Data:**

- From different instruments and platforms (ground\_based, aboard satellite, airplane, balloon, rockets)
- Different space scales: one location, a trajectory, global(satellite)
- Different time scales: routinely or during thematic campaigns
- Uniqueness: geophysical conditions are never the same
- Large variety of volumes: image to a single parameter

# Description of EO Applications

- Data Processing

- Quality control (visualisation, statistics..)
- Production and validation of elaborated parameters
- Assimilation into models
- Climate Research

- Data Retrieval operations:

- To implement new algorithms for the estimation of geophysical parameters from measured data
- short term prediction for the civil society

# Description of EO Applications

- **Modelling:**

- Different space scales: global, mesoscale, local (overlapped boxes)
- Different time scale: climate to short-term prediction
- Initiation by data sets
- Post-treatment : visualisation, comparisons with other models or set of data
- Storage and time consumers

- **Simulation:**

- To interpret observations, to test theoretical approaches

# EO Virt Org in EDG

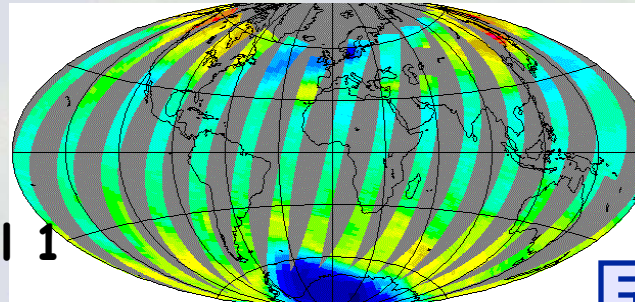
- The EO Virtual Organisation
  - IPSL, IPGP, KNMI, ESA, TV, ENEA, Dutch Space - total some 15+ people
- The EO VO infrastructure
  - @ IPSL/IPGP: local node + EDG
  - @ KNMI: EDG via NIKHEF node
  - @ ESA: local node + i/f ENEA + EDG
  - @ TV: uses ESA
  - @ ENEA: proprietary GRID infrastructure
  - @ Dutch Space: uses NIKHEF node

# EO Virt Org in EDG

- The EO applications and VO in EDG
  - **GOME L2 processing** 2 versions (TV-ENEA-ESA)
    - 7 years GOME (30k files, 500GB) processed 2 times (a- early 2003 in EDG 1.4 and in local grid; b- new algo EDG 2.0 - to be completed) Integration of IDL (cots) licences.
  - **GOME processing OPERA** (KNMI)
    - Use of application metadata DB (Sptifire and RMC) under Grid for large number of files
  - **GOME validation portal** (IPSL-ESA)
    - 1 year integrated with LIDAR profiles. 2 portal versions.
  - **GOMOS reprocessing and validation** (ESA-IPSL)
    - 5 months data (ACRI) in ESA local grid. 40k files, 120GB
  - **GREASE: OMI simulation** (Dutch Space-ESA)
    - 1 months simulation in NL nodes
  - **CEOS-GRID** (ESA-NASA-DS): Under final specification

# High level use case

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

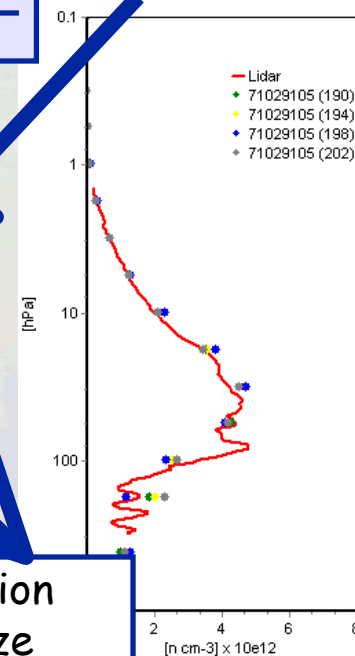


(example of 1 day total O<sub>3</sub>)

Level 1

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 (~10<sup>6</sup>/y)  
Coincident in space and time with Ground-Based measurements



Level 2

DataGrid environment

Visualization & Analyze



## • Publications

- M. Petitdidier, C. Boone, C. Leroy, S. Godin, Z. Strachman, J.C. Girard, J. Thibault, C. Hascoet, S. Rohrbach, S. Vautherot and the CCR staff, The First DataGrid Application at Jussieu University Campus: Earth Observation (WP9), 4<sup>th</sup> European DataGrid meeting, Paris, 4-8 March, 2002. (Poster)
- M. Petitdidier, C. Boone, C. Leroy, Z. Strachman with contributions of W. Som de Cerff (KNMI) and J. Van Vegte (KNMI), Requirements about Ozone Gome data format, and proposal about the metadata catalogue. Document Interne, 2002.
- M. Petitdidier, WP9 : Earth Observation Application at IPSL (France), DataGrid News Letter, April, 2002
- Casadio, S., K. Franzens, D. Del Frate, S. Godin-Beekmann, and M. Petitdidier, Grid Technology for the analysis of atmospheric ozone from satellite data. Proceedings of Data Systems In Aerospace (Dasia), Prague, Tchèque, 2-6 June 2003.
- F. Del Frate, M. Iapaolo, S. Casadio, S. Godin, M. Petitdidier, Neural networks for the dimensionality reduction of Gome measurement vector in the estimation of ozone profiles. Submitted to J. Quantitative spectroscopy nad radiative transfer, 2003.
- L Fusco, ... ESA Bulletin May 2003

# EDG Lesson learned

- Possibility to access, process, handle and store large sets of data without transfer to home storage and computers
  - Production of 7 years of Gome L2 products and Validation in several locations instead of one or two
  - Development and test of new algorithms, scientific tools and exploration of new science topics
  - Improved Collaboration -communication:
    - ⇒ No need to develop the same tool by each user
    - ⇒ New approach to scientific work

# Proposed EGEE EO applications

## Application 1 : Ozone studies

- Partners: KNMI, IPSL, SRON, RIVM, Univ. TorVergata, ESA, .....
- Goal: "Operational" use of the Grid by the Ozone community
  - Extend the GOME / GOMOS experience to other instruments (e.g. SCIAMACHY/MIPAS) to produce, store and retrieve Ozone profiles on the Grid
  - Integrate ground information and ancillary data on the Grid
  - Access a larger set of data, perform new processing (assimilation, new parameters, short term prediction, climatology, process studies...), and obtained new scientific results

⇒ To facilitate the collaboration and communication among the international ozone community

# Proposed EGEE EO applications

## Application 2: post-treatment of simulation and model outputs (seismology and climate)

- Initial Partners: IPSL+CEA, IPGP
  - Goal: post-treatment of model & simulation outputs on grid
    - To port on the Grid the seismic simulation (IPGP) ....
      - Test done in DataGrid via MPI, test on a CrossGrid cluster under work
    - Feasibility studies for model outputs by IPSL and CEA:
      - Large volume of experimental data and model outputs
      - transfer to grid of runs performed on vectorial computer
    - Post-treatments: visualisation, statistics, comparison....
- ⇒ Extension to other European laboratories (involved in European project) to develop and compare models

# Proposed EGEE EO applications

## Application 3: Interface EGEE with ESA operational data access and data handling tools

- Initial Partners: IPSL, IPGP, KNMI, Uni TorVer, ESA, ...
  - Goal: demonstrate routine access to available ESA and data tools
    - Partners will need to formalise participation to ESA AO
    - Extend Grid on Demand portal to specific applications
- ⇒ Extension of the use of EGEE to other EO datasets (e.g. MERIS, SAR...), applications (e.g. interferometry, ocean) and partners

# Specific EO requirements

- Metadata catalogues to identify the concerned files
- Restricted access to data, model output...
- Software: C, Fortran 77 and 90, IDL, Matlab, Mathematical library
- Hardware: single processor, cluster for parallel job, vectorial processor
- Platforms: Linux, SUN-OS, SGI-Irix...

# Coordinating the EO VO

## Identified responsibilities / means

- Coordination with EGEE: Monique Petitdidier (IPSL) with the support of W. Som de Cerff (KNMI), F. Del Frate (Univ. Tor Vergata), TBD (IPSL/ IPGP)
- Means:
  - @ IPSL/IPGP: local node + EGEE
  - @ KNMI: new local node + NIKHEF node + EGEE
  - @ TV: local node (tbd) + ESA node + EGEE
  - @ ESA: local node + EGEE?
  - @ other: ??
- Human and funding:
  - Human resources for operating / maintaining the grid environment is assumed made available by EGEE at national level. Need to be finalised.
  - Are there available funds to support EO applications?

# Summary

1. The EO VO in EDG
2. Ideas and initial applications for participating in EGEE
3. The ESA EO Grid plans and ESA interests in supporting the EO science community in EGEE



# EO grid plans @ESA-ESRIN

## 1. Recent facts - references

- **Frame work for ESA-EC in near future activities**
  - EC-ESA framework agreement ...; (ESA+) EC COM(2003)673 White Paper on "Space: a new European Frontier for an expanding Union. An action plan for implementing the European Space Policy"
  - EC- ESA **Global Monitoring for Environment and Security**, Final Report for the GMES Initial Period (2001-2003):

"A key feature of the GMES information architecture is the need to support collaboration between geographically dispersed GMES users and service providers. Collaboration has to be supported by an electronic infrastructure enabling GMES users not only to communicate but also to access resources such as very large data collections or archived information, scientific experiments and computing power. For the data- and computationally intensive areas of GMES, such as real-time modelling based on Earth observation data or climate modelling, high-performance networks and GRID-based computing are essential for mining, sharing and analysing data and visualising results." ...

"The combination of an ESDI together with high-speed technology networks (GRID&GEANT), space and in-situ monitoring and data collection ..."

- **ESA Agenda 2007** (new ESA DG plan)
  - Technology innovation plans ...; EO "Open & Operational" initiative

# EO grid plans @ESA-ESRIN

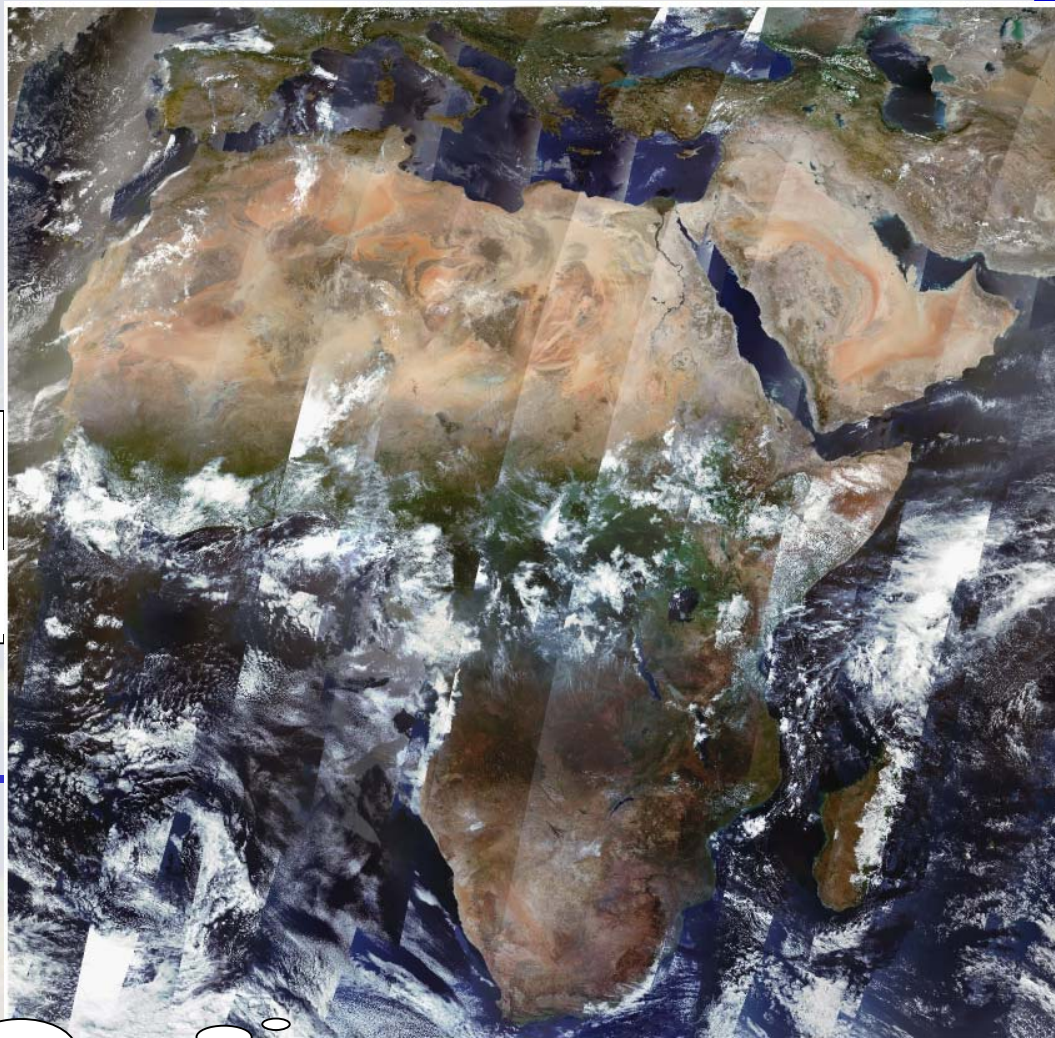
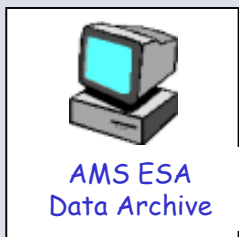
## 2. The ESA EO present Grid infrastructure

- EDG Computing Element, Storage Element
- Interface to Middleware for Earth Observation Applications (**Grid Engine**) developed by ESRIN
- Interface to operational infrastructure (MUIS, AMS) and OGIS compatible user portals.

## 3. On going upgrading

- Upgrading of local grid environment (UI, CE, SE) as platform for development of new applications
- Enlargement of the "campus grid" to ENEA, CNR-ISAC and UniTV
- Procurement of new "operational" local grid CE+SE (128 PCs, 20 TB) for "EO science and applications services"

# Present "GRID on Demand" application environment



mand

29-Aug-2003

Observation here are three  
are specific  
now more.

3:4:4
... 4:5
processed : 0 5:6
processed : 0 6:7
processed : 0 7:8

WCS / WFS / WMS

Catalogue (CSS)

European Grid

geant ←

ESA Distributed  
G/S (future)

Local Grid

# EO grid plans @ESA-ESRIN

## 4. New short term RTD projects

- "The VOICE" - e-collaboration environment -funded by ESA (under evaluation - will start in Jan/Feb 04)
- Digital Libraries ... and GRID (FP6 funded?)

## 5. Preparation of "operational environment"

- Integration of compatible and available ENVISAT software processing tools
- Procurement of Research Network Bandwidth

## 6. Preparation of ESA internal technology plan

- Involvement of industry
- Consider internal operational requirements for coming 3-4 years

# Grid on Demand new Services

## Based on available ESA developed tools:

- On-demand geographical orbit projection and mosaicking
- Format Conversion Services: Envisat standard, GeoTiff, HDF-EOS ...
- GOMOS/GOME level 1-2 processing and validation (demonstrated in EDG)
- MERIS tools AATSR
  - SMAC Processor: Simplified Method for Atmospheric Correction
  - MERIS Smile Correction Processor
  - MERIS/AATSR FLH/MCI : Fluorescence Line Height and Maximum Chlorophyll Index algorithm
- All Envisat Products L3/Binning processors

# Which operational services? – ...raising interests

## 1. Support to science users

- Support science communities for focused collaborations, e.g. cal/val, global products, new algorithms
- Support selected AOs and PIs

## 2. Support to application projects

- Provide reference application processing environment for generation of products

## 3. Support to communication needs

- Generation of periodic global and regional products for immediate availability at ESRIN

## 4. Support to Specific Reprocessing

- Allow multiple re-processing of same dataset
- Consider long term evolution of ground segment

# Vision ... in 1 year

- Consolidated system in "operational environment"
- Consolidated application service environment
- Demonstrate sharing of processing in other 1-2 sites (Kiruna, DLR, ...) where data are locally archived
  - prepare plan for a full distributed GRID environment in EO facilities
  - Move processing to archives sites

# Proposed EGEE EO applications

## Application 3: Interface EGEE with ESA operational data access and data handling tools

- Initial Partners: IPSL, IPGP, KNMI, Uni TorVer, ESA, ...
  - Goal: demonstrate routine access to available ESA and data tools
    - Partners will need to formalise participation to ESA AO
    - Extend Grid on Demand portal to specific applications
- ⇒ Extension of the use of EGEE to other EO datasets (e.g. MERIS, SAR...), applications (e.g. interferometry, ocean) and partners



# Possible working model with EGEE

- **Objective:**

- Support EO scientists to access ESA data and processing tools
- Support EO scientist to develop applications and new collaborations

- **Initial issues:**

- Definition of EO VO
- Agreement on data access by EO VO (e.g. via AO proposal)
- Agreement on interfaces to EO Portal and "GRID engine"
- Compatibility/synchronisation of ESA dedicated resources with EGEE resources

- **Resources issues**

- Assume ESA does not require funding (if no special development is required)
- How to support development of new applications by EO scientists?
- How to define priorities?