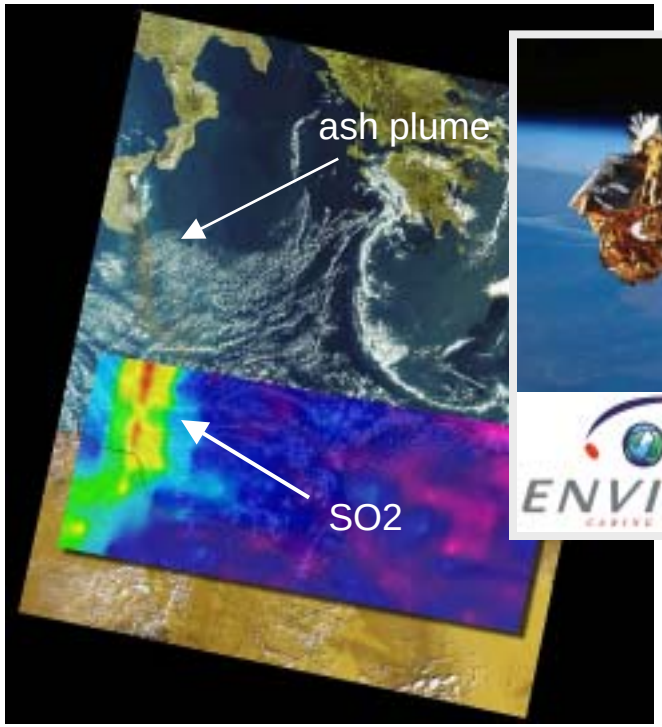
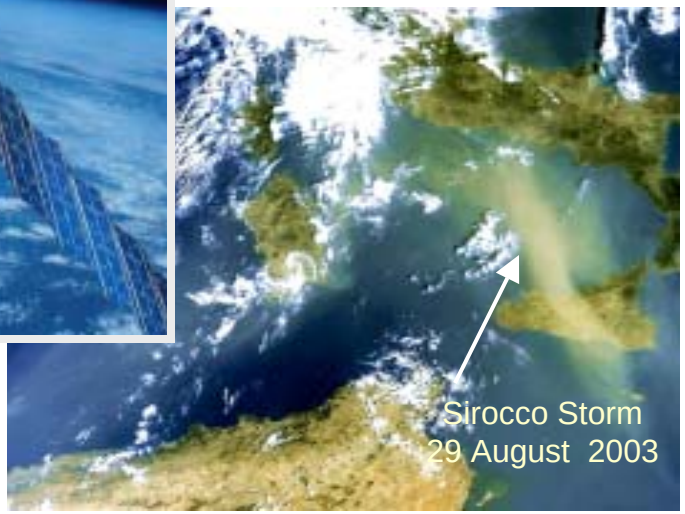
GRID technology and Earth Sciences Applications





luigi.fusco@esa.int

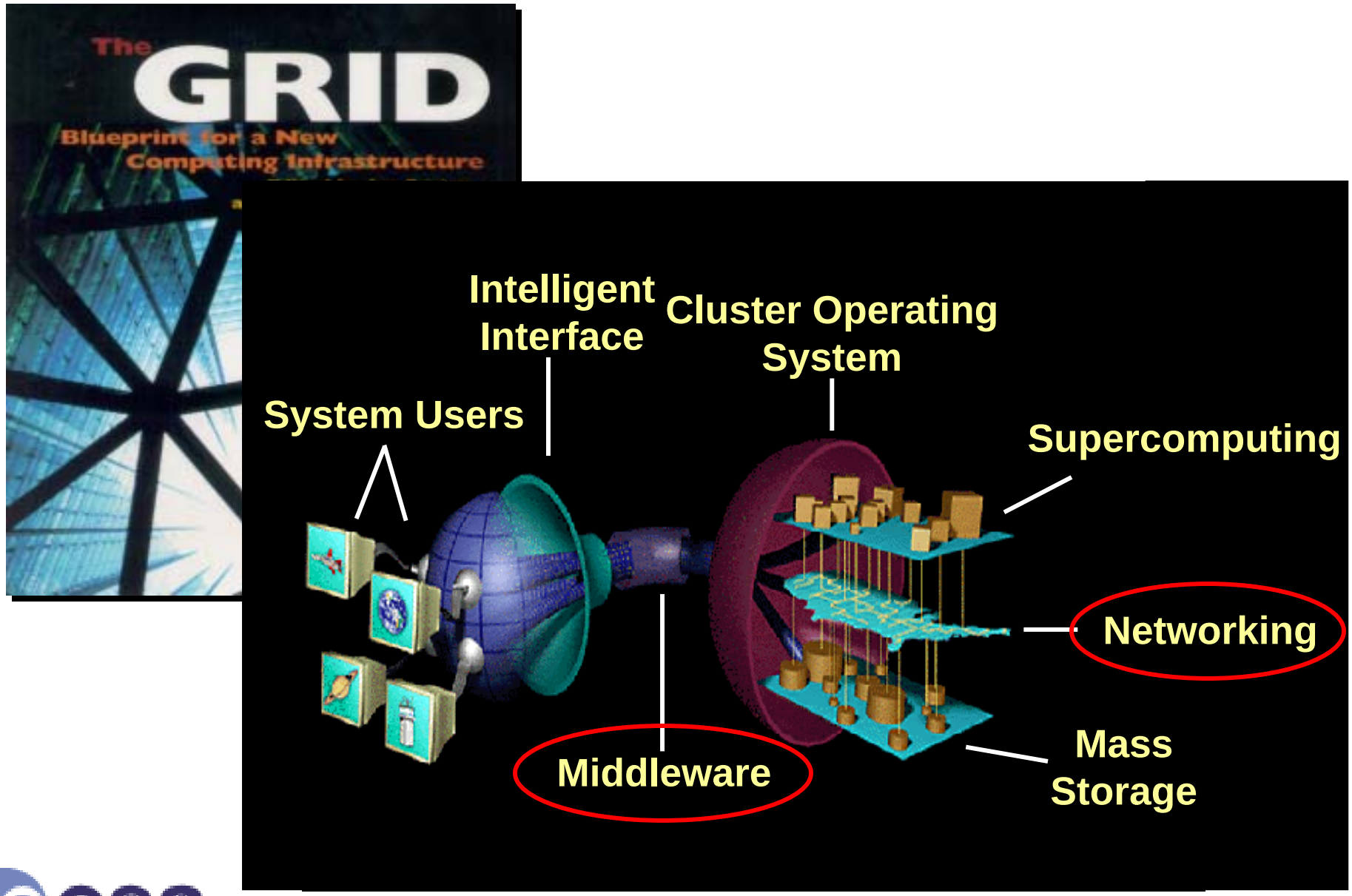
ESA EO Science and Application Department

EO DOSTAG, ESTEC 2-3 Sept 04

Summary

- Brief background on Grid technology
- Grid experience in ESA and EO community
- Current status and planning:
  - Move to operation of mature capabilities
  - Supporting an emerging science community
- Practical Demonstration

Grid technology: the concept



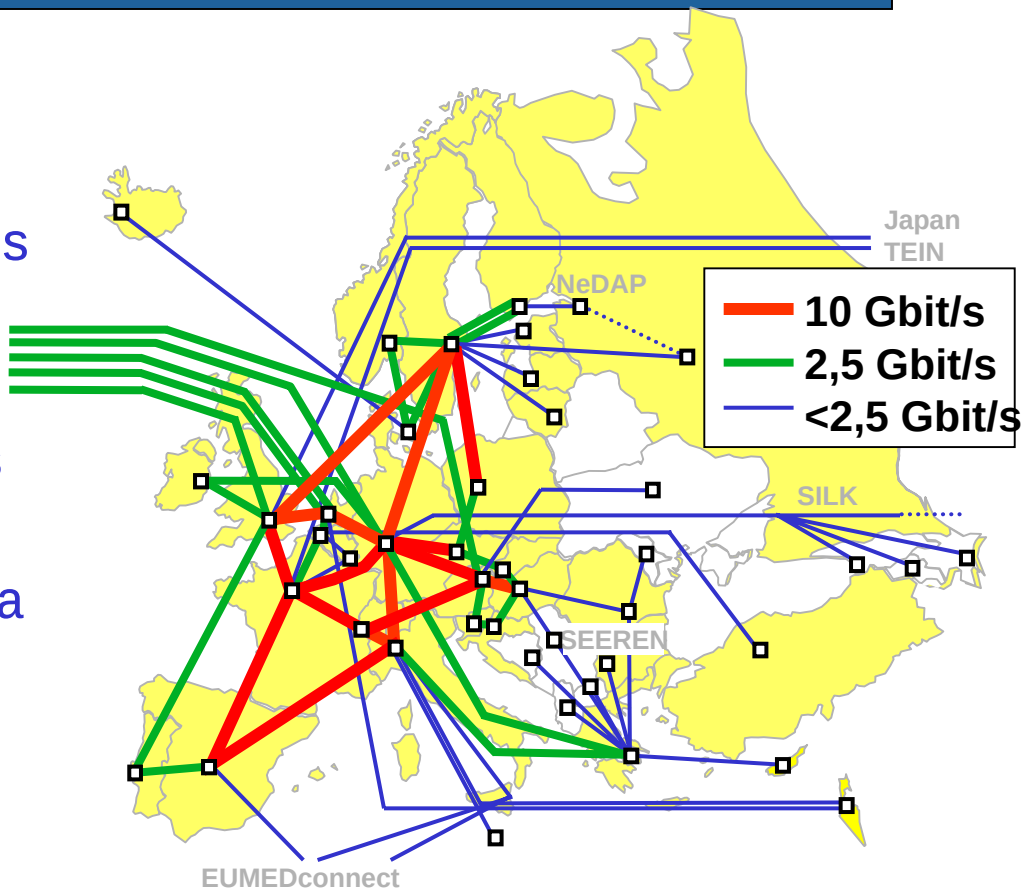
Grid in a Nutshell

- Facilitate *inter*-collaborative sharing of resources across geographic, institutional and scientific boundaries
  - Compute power, data, applications & services …
- Develop infrastructure standards & middleware to interconnect fragmented IT infrastructures
  - Improve accessibility and exploitation of online resources
- Develop comprehensive infrastructure to handle common issues:
  - Security and “single sign on”
  - Cross-community workgroups formation - “Virtual Organizations”
  - Dynamic discovery and utilization of shared resources and services
  - Network monitoring and interconnectivity optimisation
  - Location transparency (of users, computing resources, data etc.)
  - Workload scheduling and load-balancing
  - Accounting, auditing and traceability
- Make it available as a community-wide computing utility

Grid technology: the network

The European High Speed Connectivity

- Pan-European coverage (43 Countries/NRENs)
- Gigabit connectivity at 10 Gb/s
- Linking more than 3900 Universities
- Total 200 MEuro over 4 years (93 MEuro from EU)
- Extend to Mediterranean, Asia Pacific Rim, Latin America ...



Grid technology: the middleware

2001 2004

EC FP6 funding for GRID development

- € 100 m for Grid Research Infrastructures
- € 100 m for the new GÉANT
- € 50 m for Test-beds
- € 125 m for Research on Grids
- Additional budget in Application areas (e.g. Health...)

From GRID to e-infrastructure

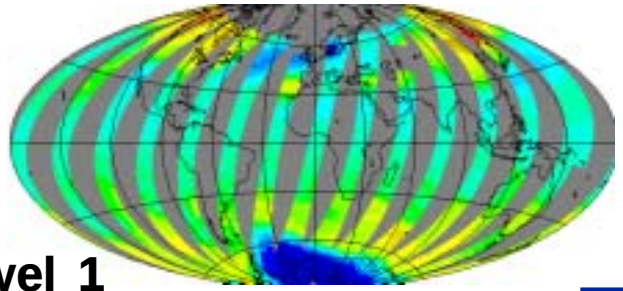
Research e-Infrastructures, the EC IST vision:

- Part of the process to implement the European Research Area (ERA)
- Powerful "instrument" for International Cooperation
- … cohesion, cooperation, standards, industry, etc

This is of interest for the EO community!

The DataGrid Ozone processing and validation test-case

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

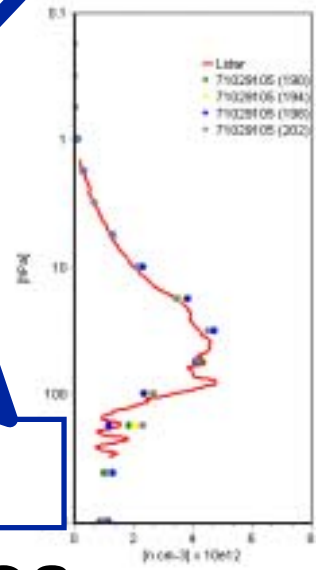


Level 1
(example of 1 day total O_3)

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

Level 2

IPSL(FR)
Validate some of the GOME ozone profiles (~10^6/y)
Coincident in space and time with
Ground-Based measurements



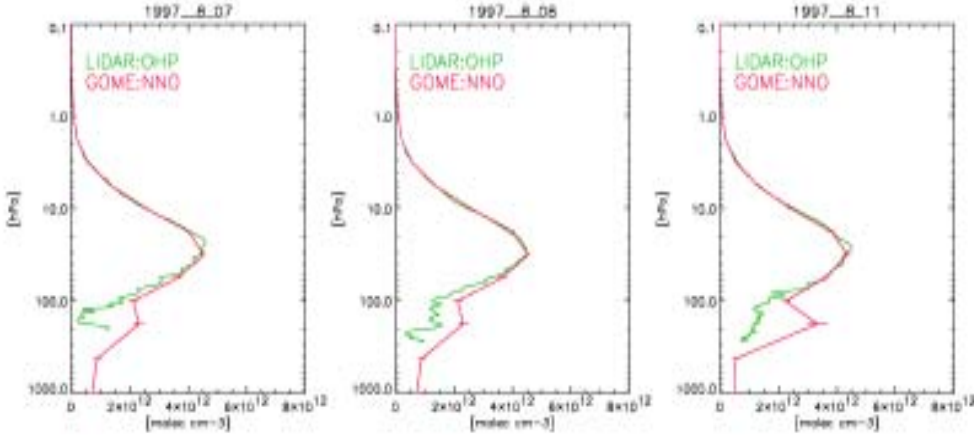
DataGrid environment

Visualization & Analyze

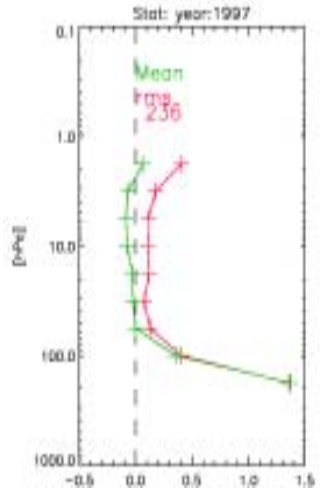
Additional GRID demonstration: GOMOS, OMI, CEOS, …

The science GOME user results

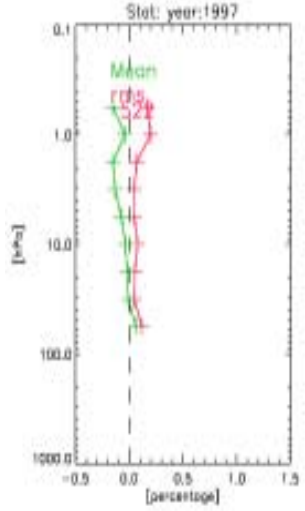
IPSL Validation using several Lidar stations



1 year comparison of Gome (red line) and Lidar data



Haute Provence



Mauna Loa

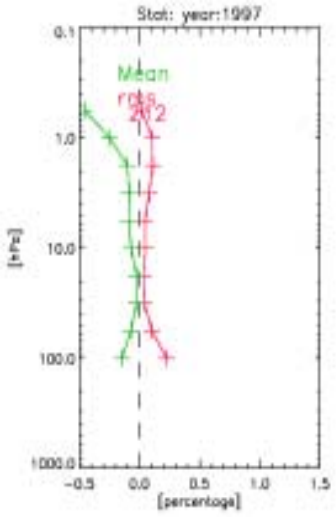


Table Mountain

1997 validation

Green line: mean of [O3(Go)-O3(Li)]/O3(Li)

Red line: rms

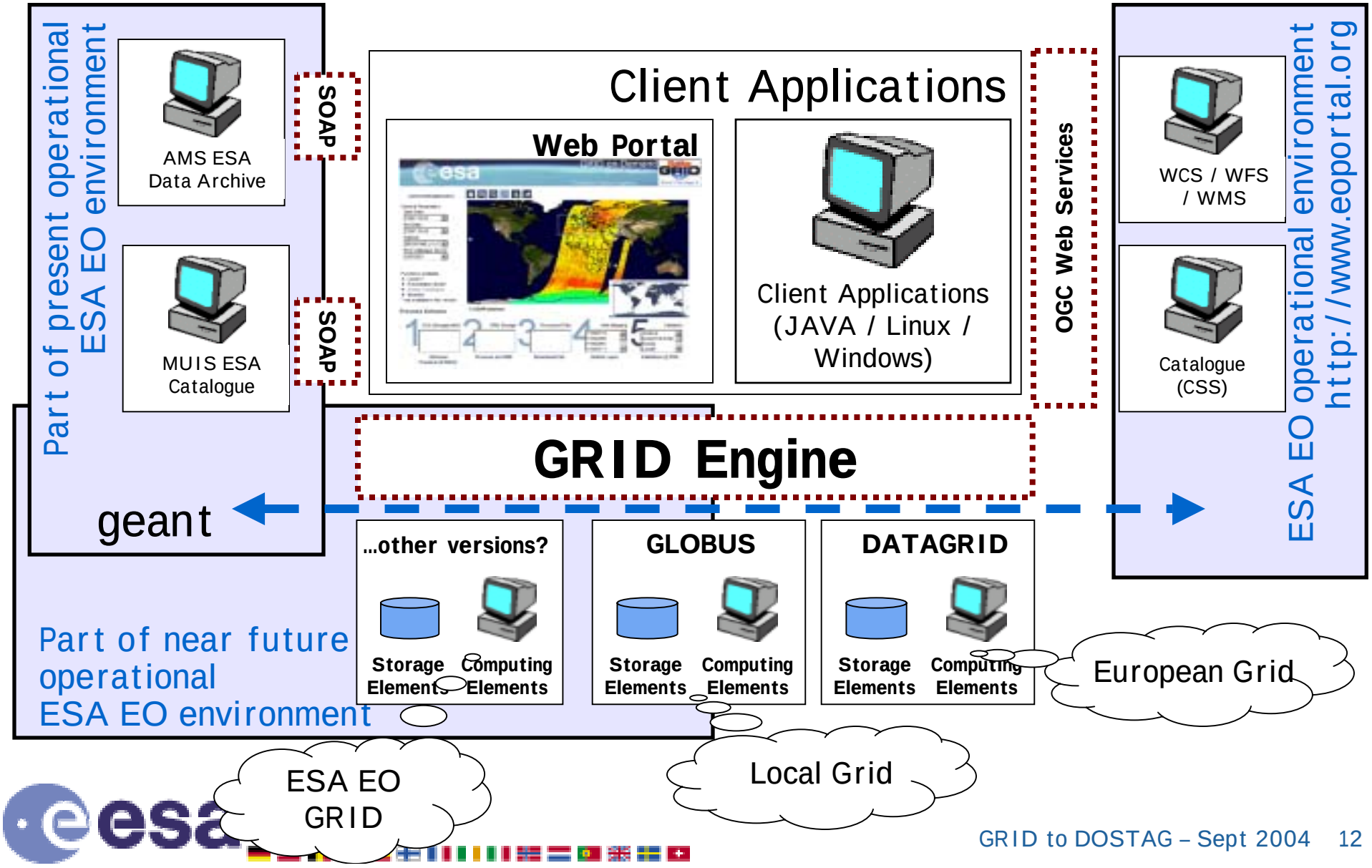
Grid technology: ESA early experience

- **ESA Grid Interest Group**
  - (2001) – focused to assess/share ESA common infrastructure
  - **http://esagrid.esa.int/**
  - **SpaceGrid** study (GSP 2002)
    - analyse needs and opportunities for various space applications (space science, solar weather, simulation spacecraft/plasma interaction, radiation transport, mechanical engineering, CDF…)
    - Analyse Grid-aware EO Ground Segment
  - CEOS WGISS GRID Task Team
- **Improve connectivity and Grid services across ESA**
  - ESTEC linked to high speed NL academic network (SurfNet)
  - **Certificate Authority** with SCI at ESTEC
- **Supporting industrial developments**
  - Dutch Space OMI simulation performed in ESRIN Grid environment …

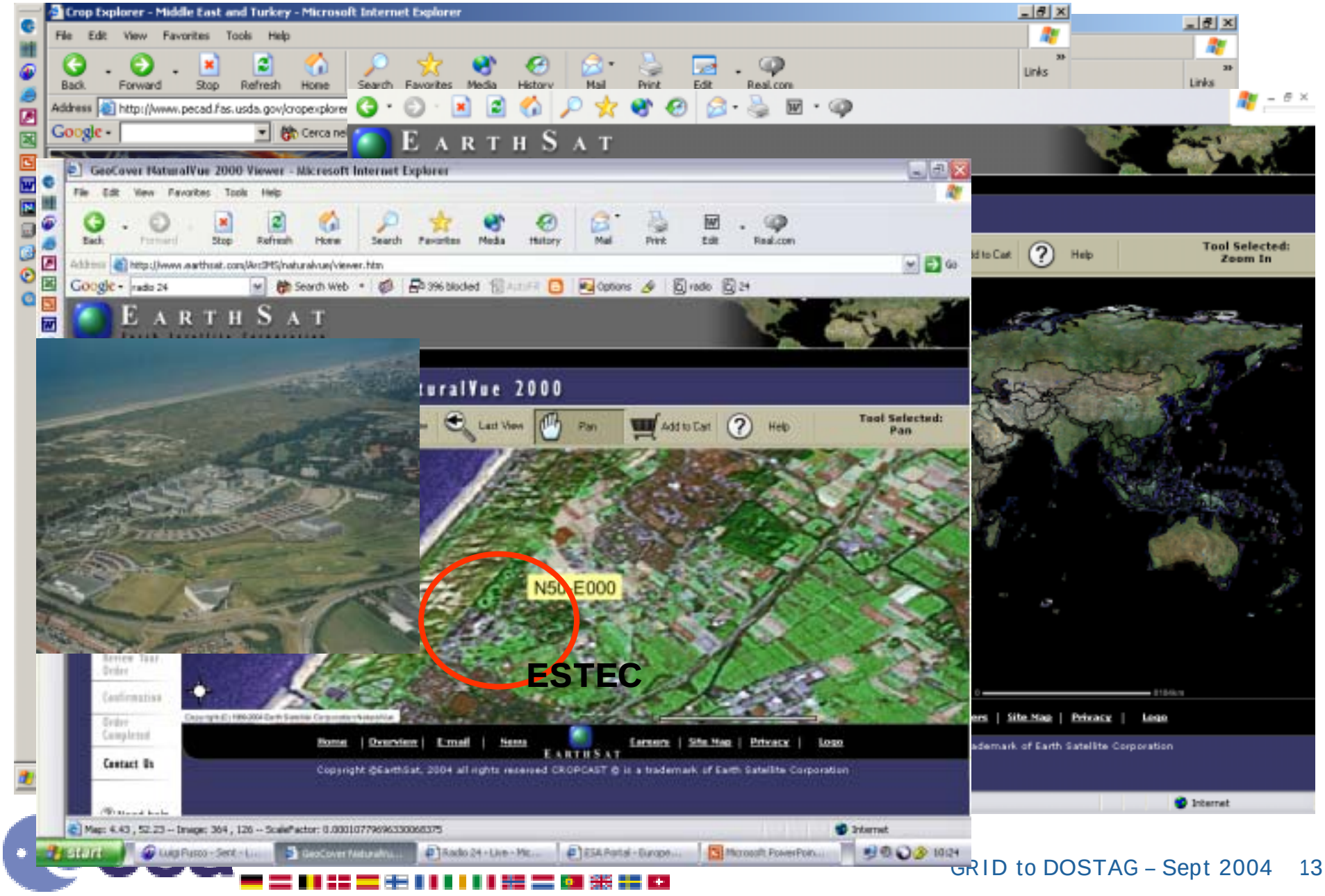
Grid technology: EO lesson learned

- **GRID is …**
  - **Accessing** network, data and resources, based on **well established protocols/standards**
  - **Secure** sharing of data and resources (via **certificates**)
  - **Better deployment of** under-exploited **resources**
  - An enabling technology for **new research approach and scientific collaborations** (e-infrastructure)
- GRID technologies are progressing rapidly in:
  - **Grid & Web service integration** (relatively mature, specially in EO environment)
- Way forward for EO
  - Europe takes the lead in e-science development initiative
  - **Move to "operational use" of GRID infrastructure** specially for data access, new science community services…
  - A possible environment to implement "Oxygen – **Open** and Operational" functionalities
  - **Need to confirm** controlled and **secure access by external users**

GRID integration with EO Web Services



Examples of emerging EO data services



Towards an EO Grid infrastructure

1. **GRID as model for the future European EO Ground Segment…**
   - Transfer to ESA EO operational environment of mature GRID capabilities

2. **… GRID as model for the Science User Segment**
   - Similar to other science community (e.g. High Energy Physics, Astrophysics, …)

1. EO Grid infrastructure Ground Segment vision (→2006)

- Transfer to operation of present GRID capabilities:
  - Support **reprocessing** requirements
    - Demonstrated already for L1 → L2 GOMOS processing, planned for GOMOS and MERIS reprocessing in 2005
  - Innovative solution for the data access and exchange
  - Increased exploitation of available resources
- GRID interconnection of ESA and other operational facilities (i.e. ESRIN, Stations, PACs, key user facilities…)
  - Reference infrastructure for new missions G/S studies
  - Enable new classes of large-scale, high-power applications (e.g. **processing** of large data volumes)
- Initial validation for GMES / GEO Architectures
- Some issues
  - High speed connectivity in place across facilities
  - Standard and new processor systems based on Grid compatible standard (e.g. Linux)

2. Earth Science infrastructure User segment Vision (→2006)

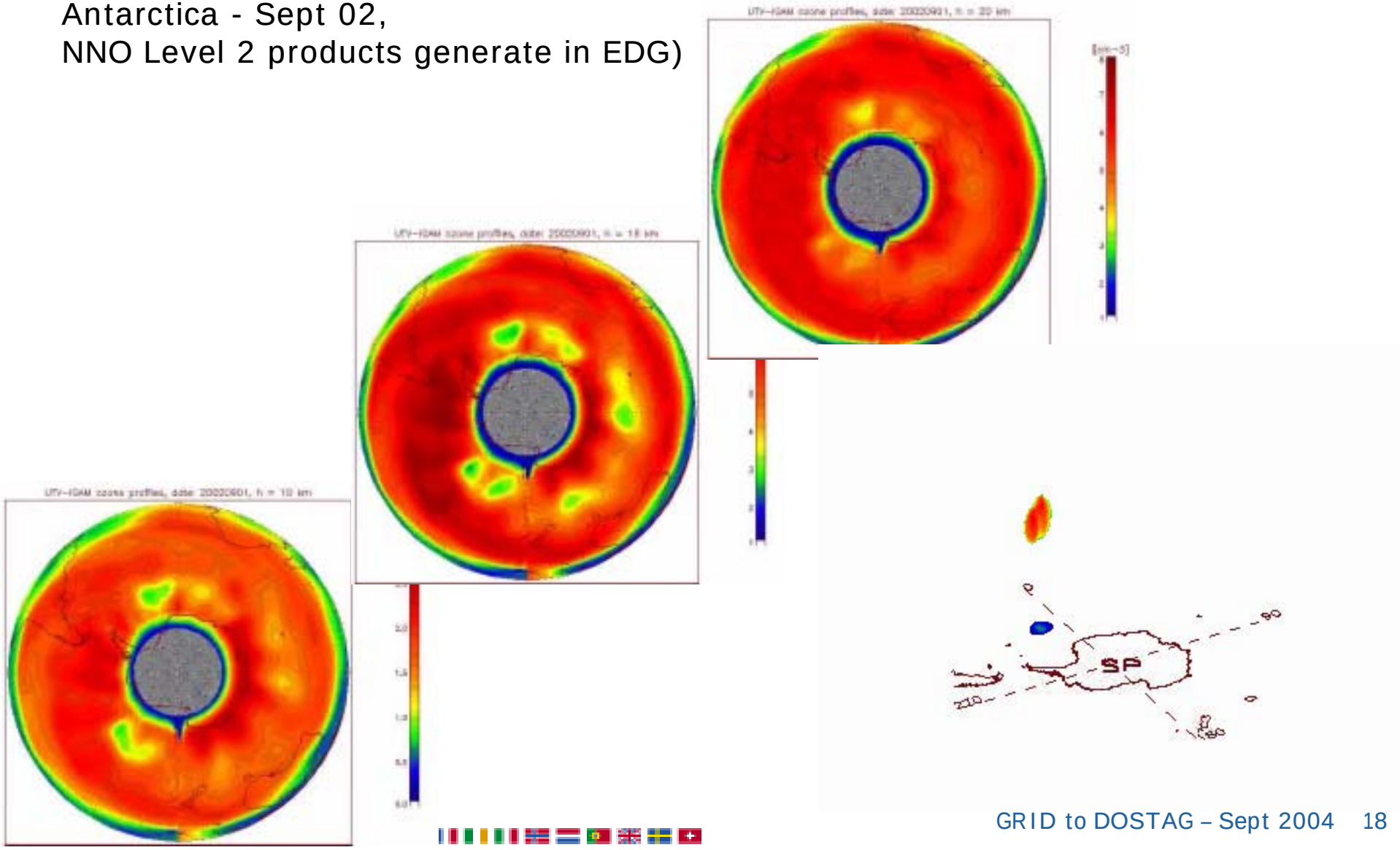
- Set up of the **EO Science User Segment**
  - Include **support to science** data processing / analysis based on Envisat tools (e.g. BEAM, to be extended to BEST, BEAT)
  - Routine demonstration of **science cooperation models** (e.g. CAL/VAL, new algorithm developments, integration of ground and space measurements…)
- **Move algorithms and analysis tools to data …**
  - **data assimilation** … integration in models…
  - Support large data sets science application projects…
- Some issues:
  - User access to data/services via secure connection (certificates)
  - Common standards across science users
  - Online data access for real time (and archive)

Example of initial in-house demonstrations

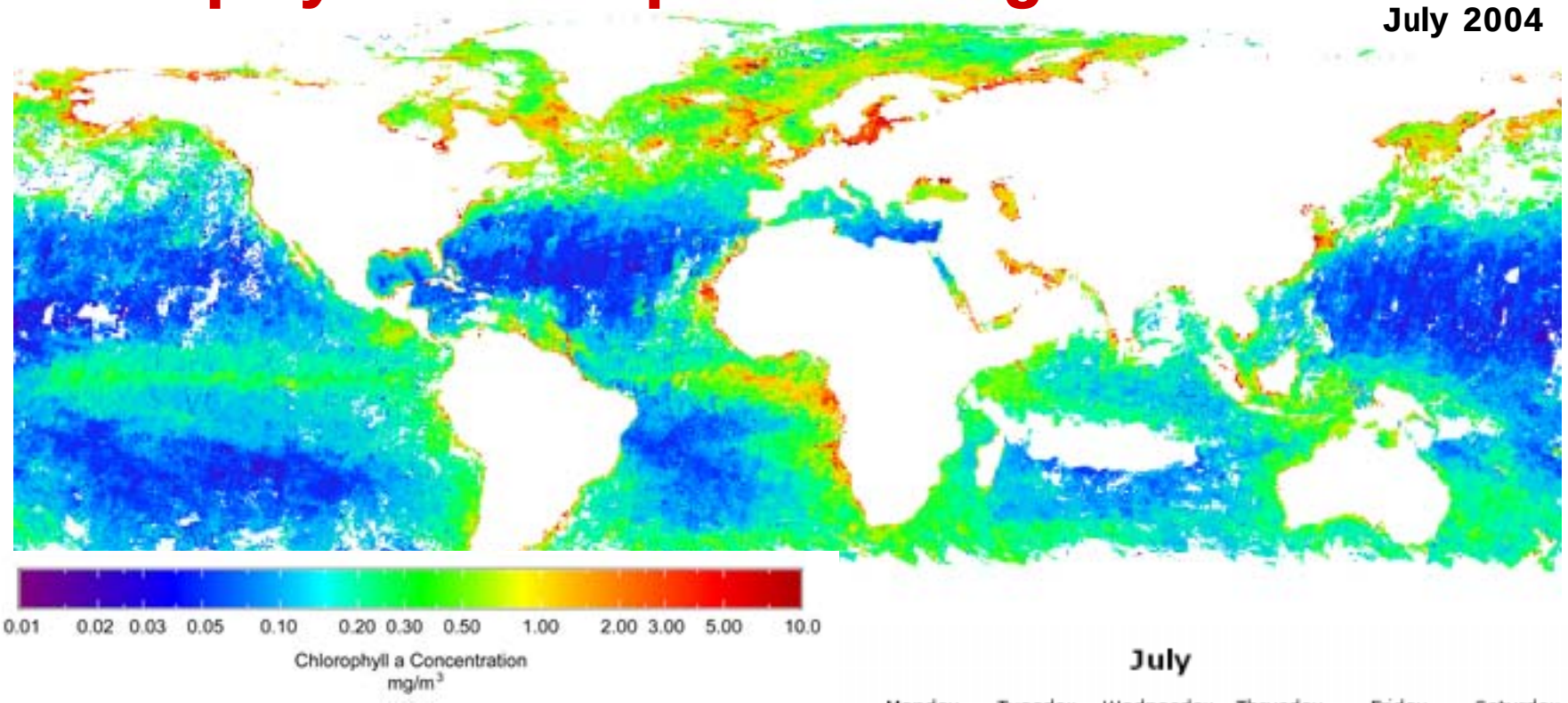
- Reprocessing of GOME from level 1 to level 2
  - Validation of GOME Ozone profiles vs LIDAR
  - Integration of Ozone profiles
- MERIS global products at 1 Km (BEAM tools)
  - Chlorophyll
  - Global products for PR/communication
  - MGVI (comparison of diffent algorithms)
- ASAR Global Mode over Antarctica
- ...

The science GOME user results

S. Casadio – ESA ESRIN
(GOME 3D Ozone volume over
Antarctica - Sept 02,
NNO Level 2 products generate in EDG)



Chlorophyll BEAM processing on GRID



Input Data:

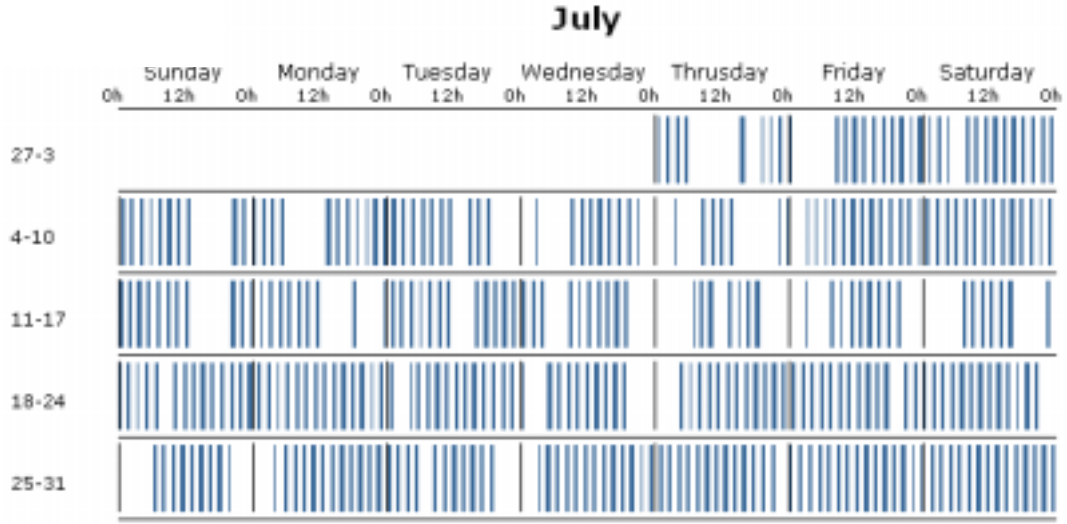
178.92 Gb in 1828 files

Output Files:

700 Kb (JPEG)

28 Mb (TIFF)

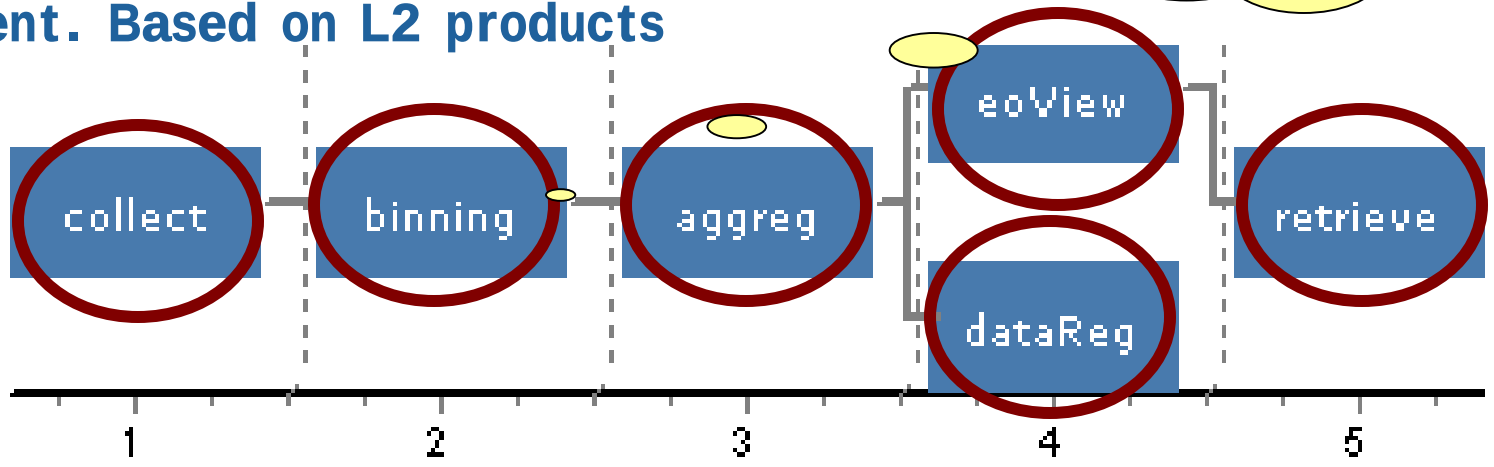
36 Mb (DIM-MAP)



Chlorophyll processing on GRID

BEAM code is embedded here

Task chaining in the "Science Grid on Demand" environment. Based on L2 products



COLLECT FUNCTION

Searches and retrieves the requested files from the storage elements

AGGREGATION FUNCTION
Assembles the resulting files from the binning processes

DATA REGISTRY FUNCTION
Registries the data files and the corresponding metadata in the defined storage element

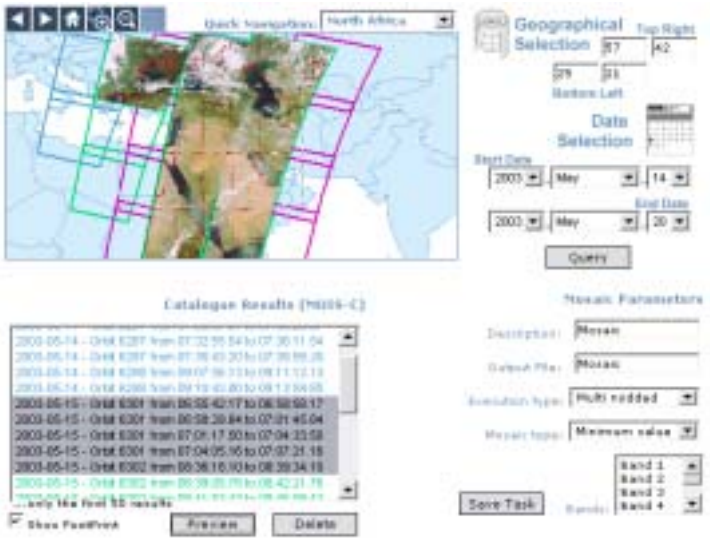
BINNING FUNCTION
Projects the Envisat files to a given datum and projection and according to the requested geographical bounding box

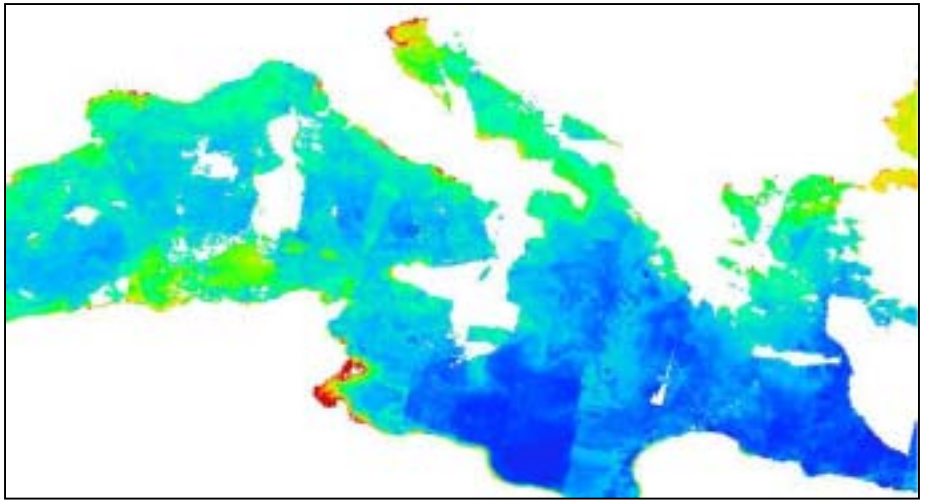
Earth Observation Visualization Function
Generic functions for image composition and format conversion (JPEG/TIFF/…)

RETRIVE FUNCTION
Generic function for file transfer (FTP, SFTP, gridFTP,…)

Chlorophyll processing on GRID

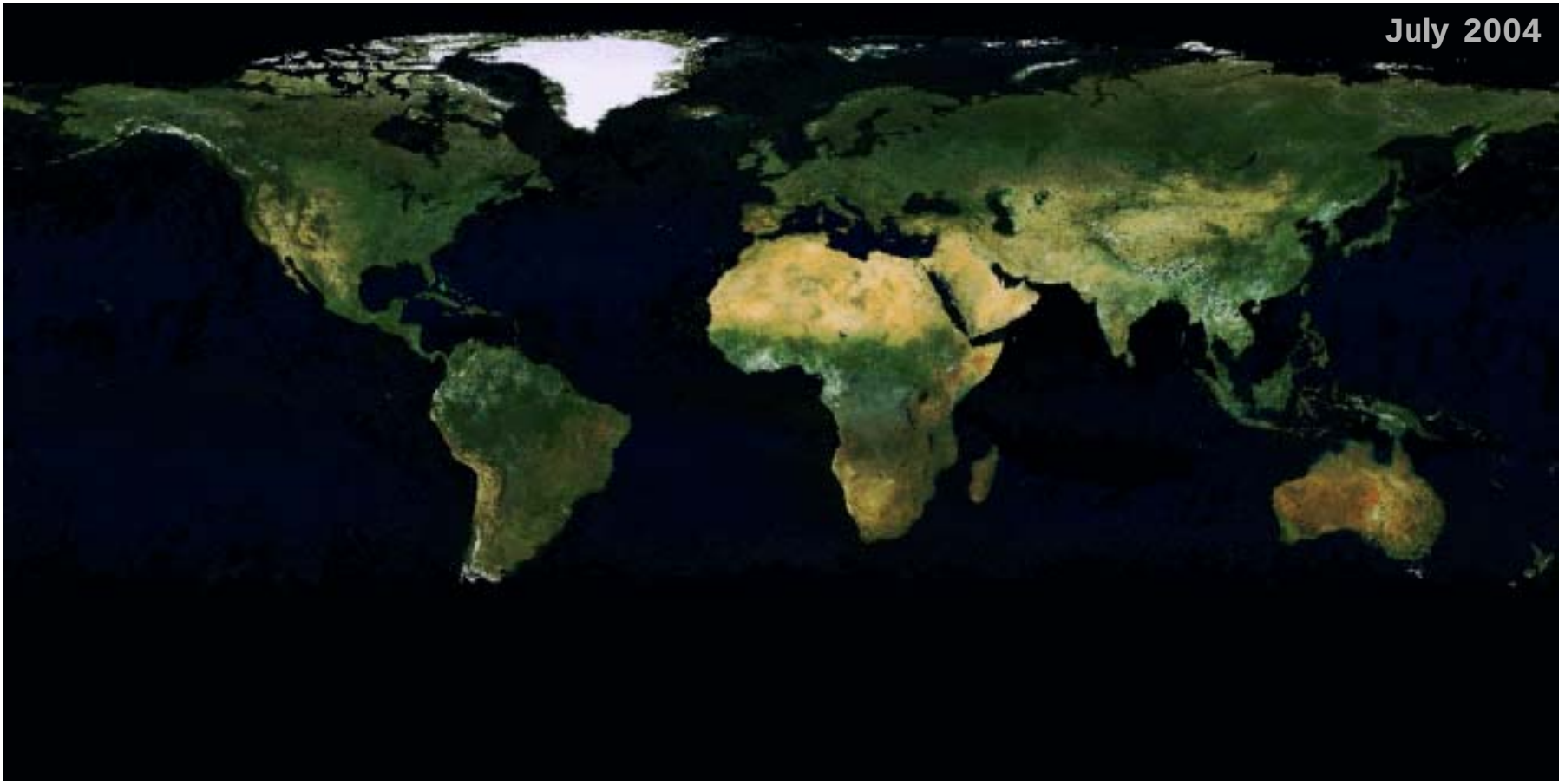
Web Based Interface Control



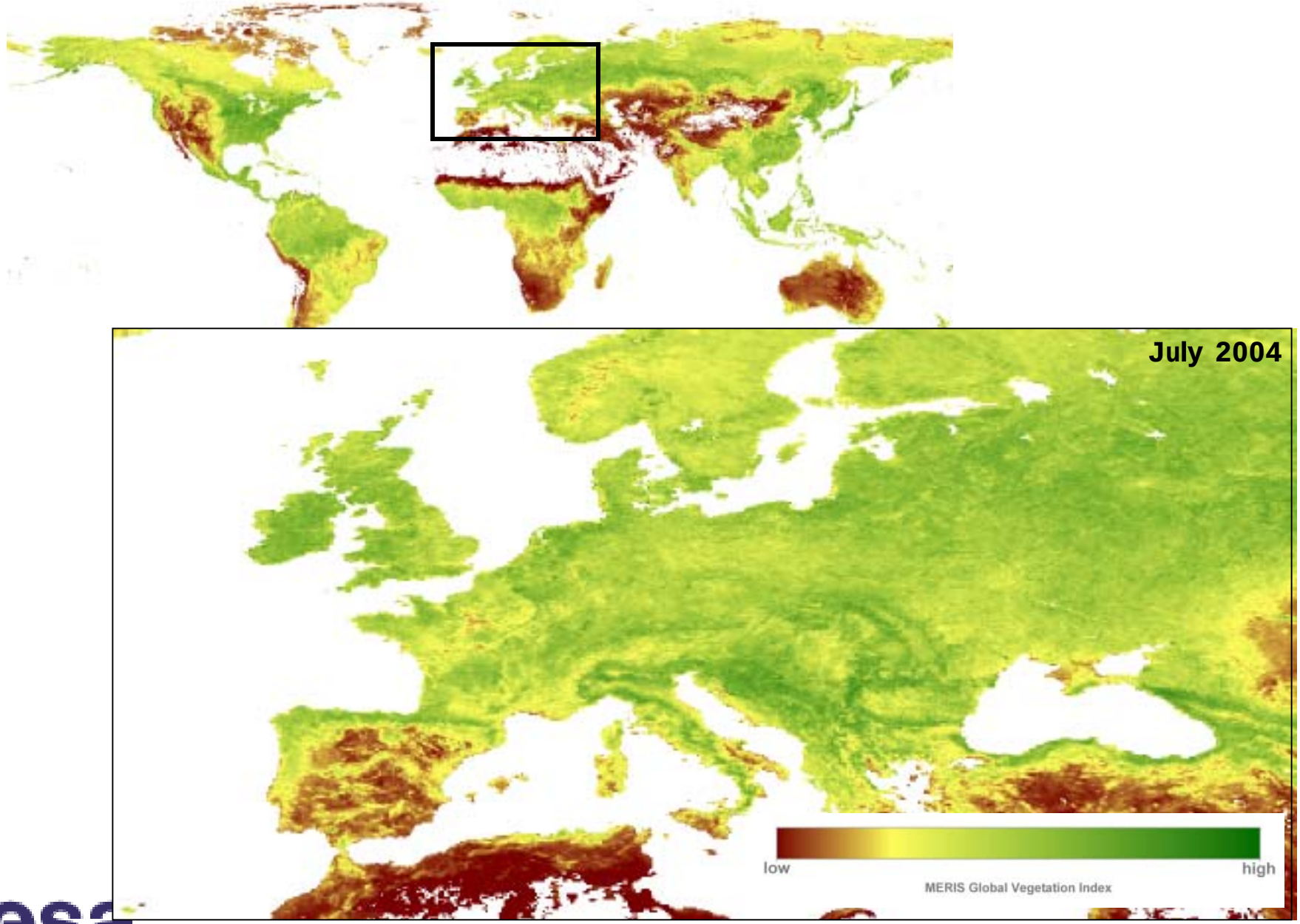


MERIS 1km mosaic for PR/Communication

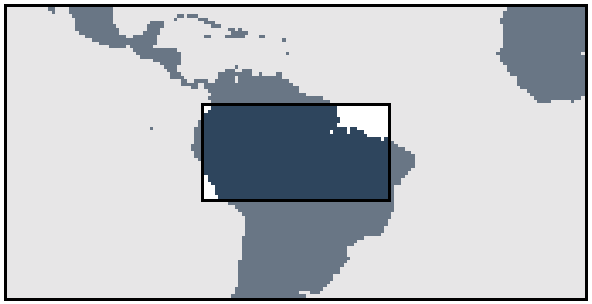
Based on BEAM mosaiking tools, L2 standard products.

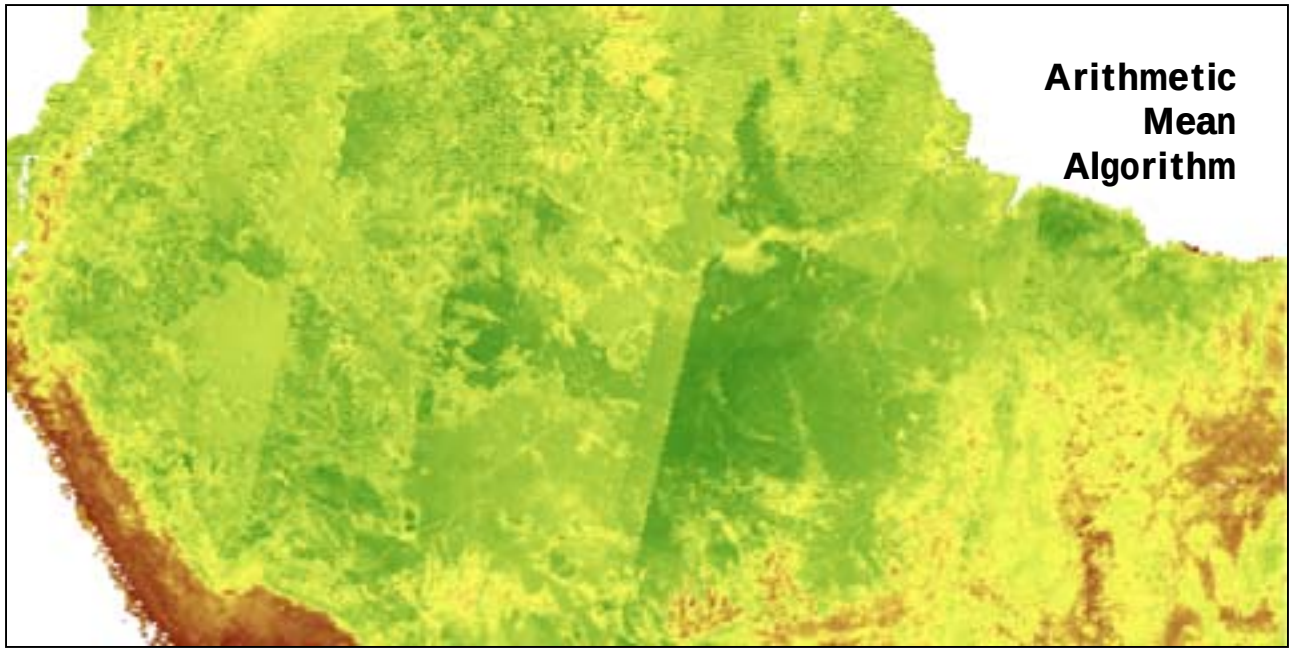


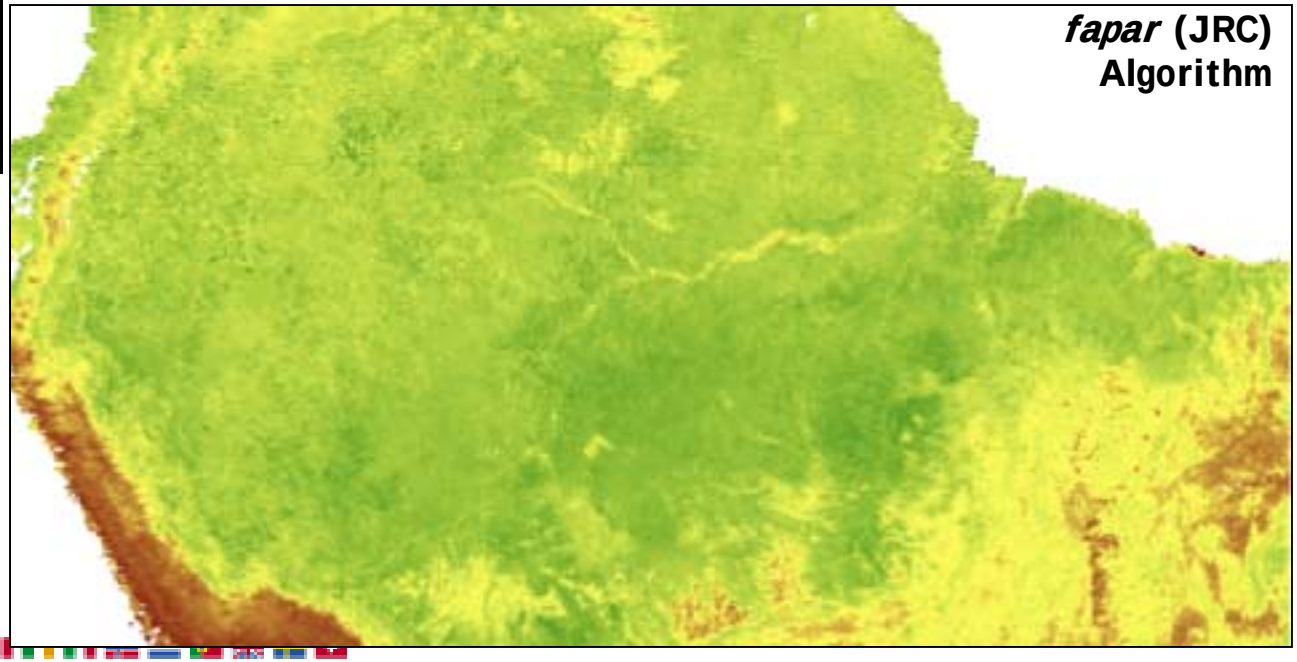
MERIS MGVI global product @ 1Km



MERIS MGVI comparison of algorithms





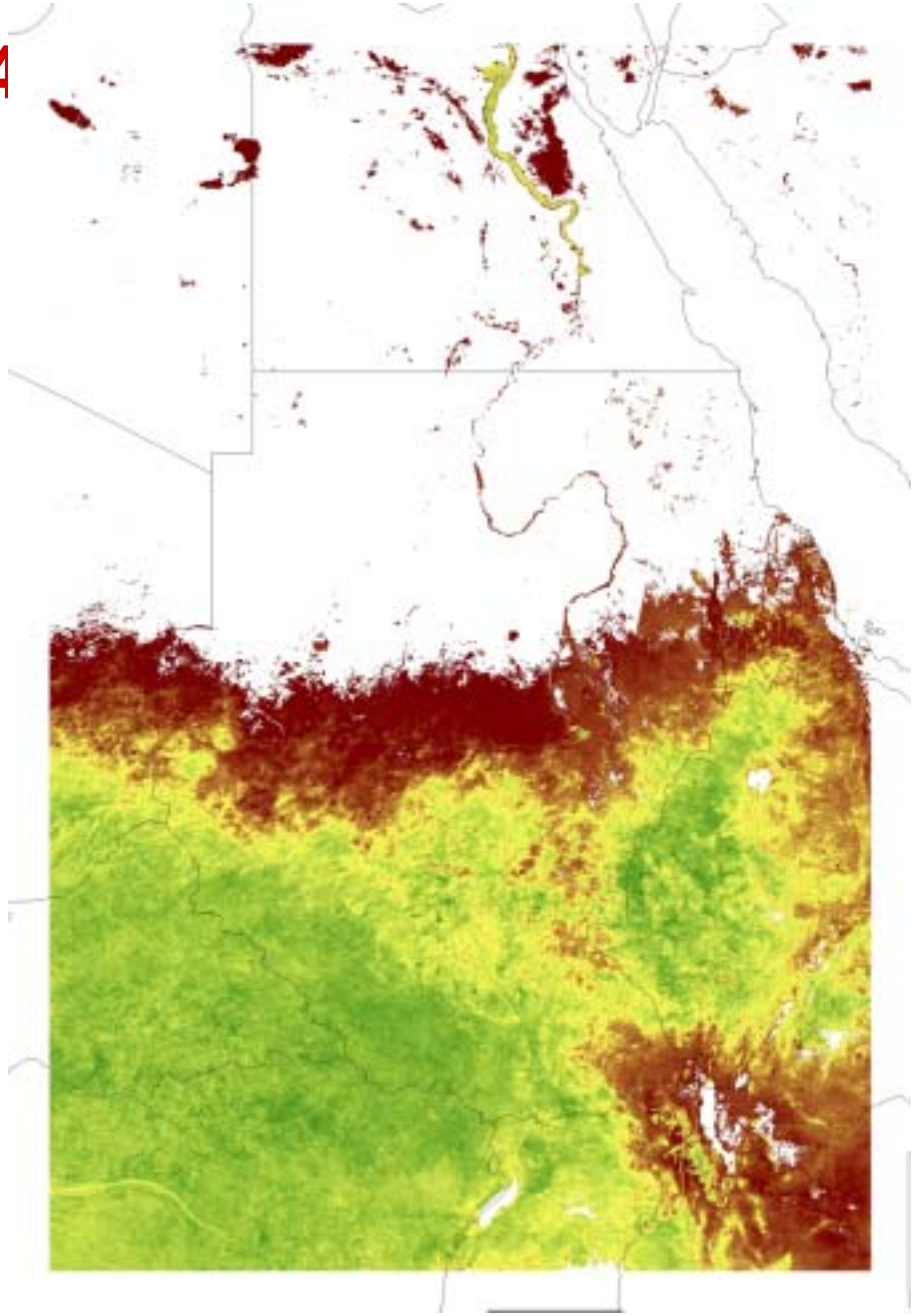


MERIS MGVI – Aug 04

SUDAN

DARFUR CRISIS

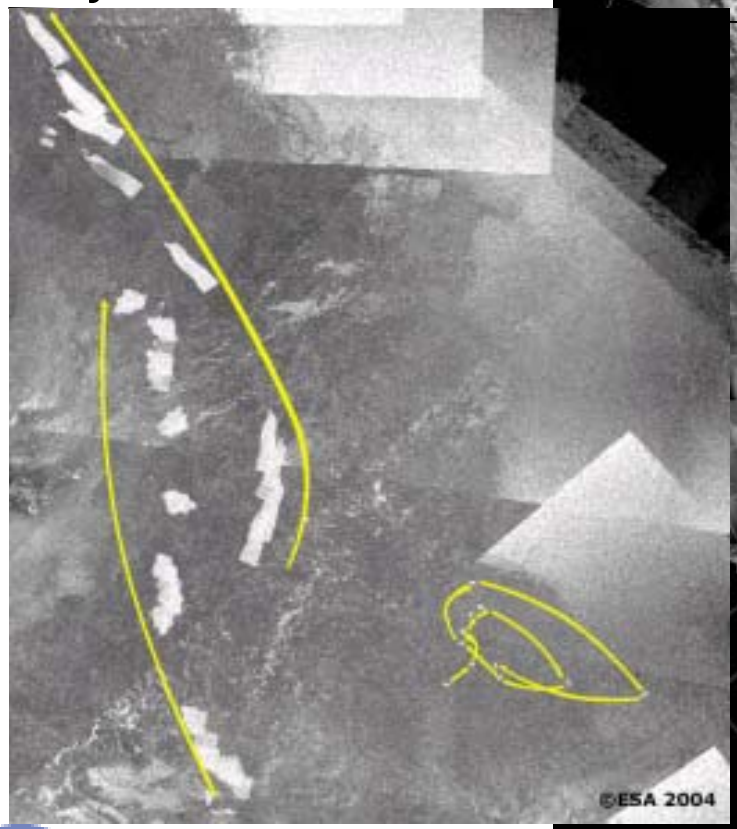
Vegetation Change Analysis for Humanitarian Crisis Management

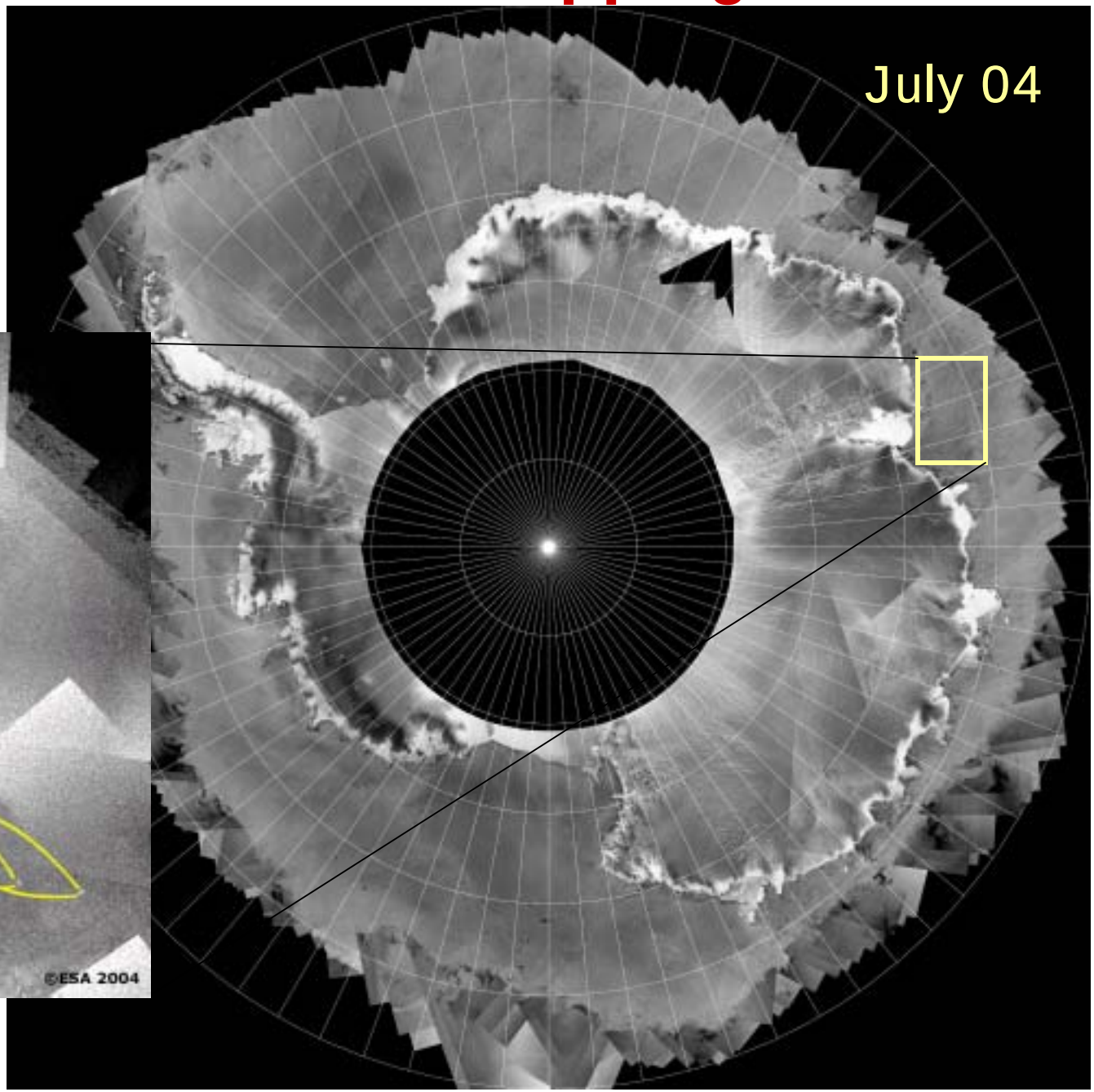


MERIS ASAR Global mode mapping

Extension of Sea Ice Pack

Icebergs displacement Apr-May 04





Near future plans: "Earth Science Grid on Demand"

- **Goal 1: bring the user algorithms and tools to the data**
  - Provide protected environment for hosting user specific modules
  - Support validation of new processing algorithms, new geophysical level 3 products …
  - Some test cases identified within the ENVISAT AOs with user involvement
    - MERIS data assimilation for chlorophyll forecast
    - Objective analysis (SST and ocean colour)
    - Alternative algorithm validation for Case 2 water
    - …
- **Goal 2: extend approach to wide science user community**
  - Planned for next year – new AO?

Near future plans: "Earth Science e-collaboration"

- **e-collaboration in Earth Science:**
  - Integrating GRID and other technologies (e.g. wireless, sensor web…)
  - Sharing data sources, results, models …
  - Building thematic Virtual Organisation

- **A dedicated GSP project is evaluating the potential of e-collaboration in EO (in 3 specific domains):**
  - Ozone Cal/Val
  - GMES Open Services
  - Forest and Rural services

- **Complementarity with EC IST strategic line "collaboration@work"**

Conclusion

- GRID is an emerging technology which can really help EO and Earth Science community
- EC has world leadership in developing the technology and supporting research communities e-infrastructure
- Convergence between GRID and Web Services puts EO in good position for immediate utilisation of the technology
- Few Earth Science groups have already shown interests in ongoing activities