Planning for EO Application Deployment

DataGrid WP9 EO Applications Parallel Session

5th Project Conference – 2-5 September, Budapest

Agenda

- Introduction
  - Summary of WP9 main objectives, tasks status, deliverables Status
  - Planning
- Update on activities by WP9 partners
- Discussion of Special Topics
  - Implementation of 9.4 Use Case
  - Data Packaging
  - User Interface tools
  - Planning for Demonstrations
  - Web-map GOME portal
  - Participation to CEOS GRID testbed
  - Dissemination and extension to other applications
  - Interaction with scientists

Summary of WP9 Objectives

- Feeding EO Requirements into the middleware development effort
  - Main vehicle is the Architecture Group
- Application of GRID techniques within EO applications
- Develop standard techniques, tools and interfaces for EO applications
- Testbed evaluation - demonstrating the benefits
- Coordination and synchronisation within WP9 and with other DataGrid work packages
- Scalability: Actions needed to achieve a full-scale EO Grid
- Dissemination of GRID Concepts & Techniques in EO / ES / Space

WP9 Tasks Summary

- Task 9.1 EO requirements definition
  - Completed
- Task 9.2 EO related middleware components (pm1 - pm18)
  - Completed
- Task 9.3 EO Platform Interface Development (pm7 - pm30)
  - On-going
- Task 9.4 EO Ozone/Climate application (pm13 - pm36)
  - Preparation of PM24 Demonstration on-going
- Task 9.5 Full scale Earth Science application (pm7 - pm18)
  - Nearly completed

WP9 Deliverables Summary

- D9.1 EO Requirements Specification
  - Delivered to the EU
- D9.2 Report on Grid Testbeds State-of-art
  - Delivered to the EU
- D9.6 Report on the EO application scaling study
  - Under PTB Review (2nd review cycle)
- D9.3 Demonstration (and report) of second software components release (task 9.4)
  - Due PM24 (December 2002) Ongoing
- D9.4 Report on EO application platform interface (task 9.3)
  - Due PM30 (June 2003)
- D9.5 Demonstration (and report) of EO application processing test-bed (task 9.4)
  - Due PM36 (December 2003)

High level Planning

- Recent Achievements Q5, Q6 (PM12-18) March-June 2002
  - Completion of TB1 Validation
  - Development of EO Infrastructure & Architecture
    - Input to EDG Architecture Group
    - Distributed Data Handling
  - EO Application Development
  - TB1 Site Installation
  - D9.6 EO Application Scaling Study
- Medium-term Planning Q7, Q8 (PM19-24) July-December 2002
  - WP9 Sites join the DataGrid
  - D9.3 Demonstration & Report
- Long-Term Planning Q9-Q12 (PM25-36) to December 2003
  - D9.4 EO Application Platform Interface
  - D9.5 Demonstration & Report

Update on individual achievements and plans (KNMI)

- EDG Testing and evaluation activities
  - Installing and testing Spitfire for meta database usage (done)
  - Filling meta database from the grid (planned)
  - Querying the database from the grid (planned)
- Adaptation of GOME processing and validation algorithms
  - Adjustments of Opera (ongoing)
- Development of EO Grid infrastructure
  - None foreseen
- Outreach, new users
  - Presentation on conferences, internal KNMI (ongoing)

Update on individual achievements and plans (IPSL)

- EDG Testing and evaluation activities
  - Continuation of Testbed1 testing validation (move to v1.2)
  - Implantation of prototype GUI
  - Test on the Grid of the second IPSL application : "destruction rate of ozone in the Polar vortex"
- Adaptation of GOME validation algorithms
  - Participation to task 9.4 GOME Application demonstration
  - Finalisation of the Lidar metadata catalogue by using MySQL SGBD
  - Test of Spitfire to extract information from the metadata catalogue
  - Test of the partitioning by latitude, longitude and date
    - One month of global data (800 000 files),
    - Study and implementation of the needed functionalities
- Outreach, new users
  - Proposition of GOMOS validation demonstration

Update on individual achievements and plans (ESA)

- EDG Testing and evaluation activities using TB1.1.4 and TB1.2
  - Extension of Grid Surfer to submit/monitor multiple jobs simultaneously
  - Grid Surfer function to perform data storage and replication tests
  - Design, prototype development of EO WEB portal
- GOME large scale processing
  - Preparations for 9.4 use case, AMS data loading on going
  - Metadata catalog creation and population integration in MUIS
  - NNO L2 Dataset production
- Development of EO Grid infrastructure
  - D9.6 scaling study completed, review of requirements vs EDG testbeds
  - Input to ATF and EDG MW development WPs
  - ESA installation (UI, CE, SE) move to TB1.2, operational interface with ENEA-GRID, AMS and MUIS
- Proposition of CEOS-GRID, involvement of more applications / users in WP9

IPSL resource status

- People involved
  - Monique Petitdidier (IPSL) : coordination and management 40 %
  - Cathy Boonne (IPSL) : application, system 20%
  - Zaharia Strachman (LPNHEP/IN2P3) : application, system 70 %
  - Christine Leroy and Luc Peltier : technical leader 100 %
  - Scientists 20 % : Sophie Godin (IPSL/SA), Slimane Bekki (IPSL/SA)
  - Staff of the Jussieu Campus computer centre
- Facilities
  - 3 PC (UI, CE) and a storage element (500 Go) implemented on a dedicated network
- Effort spent and planned
  - Unfunded : from january 2001 to June2002 → 18,90 pm
  - Short and medium term → 8 pm

KNMI resources status

- People involved
  - John van de Vegte, Wim Som de Cerff (both 50%)
- Effort spent and planned
  - According the technical annex

Partner resources status: ESA resources

Near future resources	Task no						P-M
	0	1	2	3	4	5	
L. Fusco - mgt	10%						2.7 pm/m funded internally
F Rossi - sys				25%	25%		
S. Casadio -science					10%		
P. Gonsalvez - app					100%		
G D'Acunzo - sys ad				50%	50%		
J. Linford - coord	20%			40%	40%		3.0 pm/m part funded by project
A. Terracina - sys			50%	50%			
M. Fulcoli - sys / app			30%	70%			
total	30%	compl	80%	235%	225%	compl	5.7 pm/m

Total project funding: 84 mm,
spent in y1: 27, available in y2: 33, available in y3: 24

ESA resources status (2 of 2)

- Facilities
  - 16-node Linux cluster with 800 GB Raid
  - Gigabit link to ENEA distributed processing facility
  - 1 Linux PC running RH 7.2 with high-res graphics
  - 2 Linux PCs running RH6.2
  - 1 Sun ultra machine
- Software and applications
  - RSI IDL Licenses
  - NNO algorithm, more to come
  - AMS, MUIS, …

Discussion of Special Topics

???Possible use of teleconf???

- Implementation of 9.4 Use Case: in view of end of year demonstration
- Data Packaging: alternative modes to compress data volumes and number of files
- User Interface tools and components: developing set of tools to easy EO user access
- Planning for demonstrations: who will do what and when till end of year (e.g. GRID surfer, ESTEC demo, …)
- Web-map GOME portal: development of Web-services for EO
- Participation to CEOS GRID testbed using DataTAG networking guidelines
- Dissemination and extension to other applications
- Interaction with scientists

Discussion of Special Topics: Implementation of 9.4 Use Case in view of end of year demonstration

- Adaptation of GOME processing and validation algorithms
- Use of Spitfire catalogues
- GOME L1 and L2 dataset replication on the DataGrid SEs
  - Interface to AMS

Discussion of Special Topics: alternative modes to partition data files

Many geophysical applications, especially validation, requests ozone profiles for given areas.

- Orbit approach
  - The retrieved Ozone profiles are stored by orbits
  - The difficulty is to access efficiently to the selected area
- File approach
  - Each file corresponds to a given pixel of an orbit
  - Advantage : efficient access to the selected area
  - Drawback : too many files and metadata records
- Longitude and latitude partitioning ; IPSL proposal
  - Each day, the individual files are partitioned into latitude and longitude "boxes"
  - Advantage : efficient access to the selected area and less number of files
  - Drawback : need to partition the produced files data

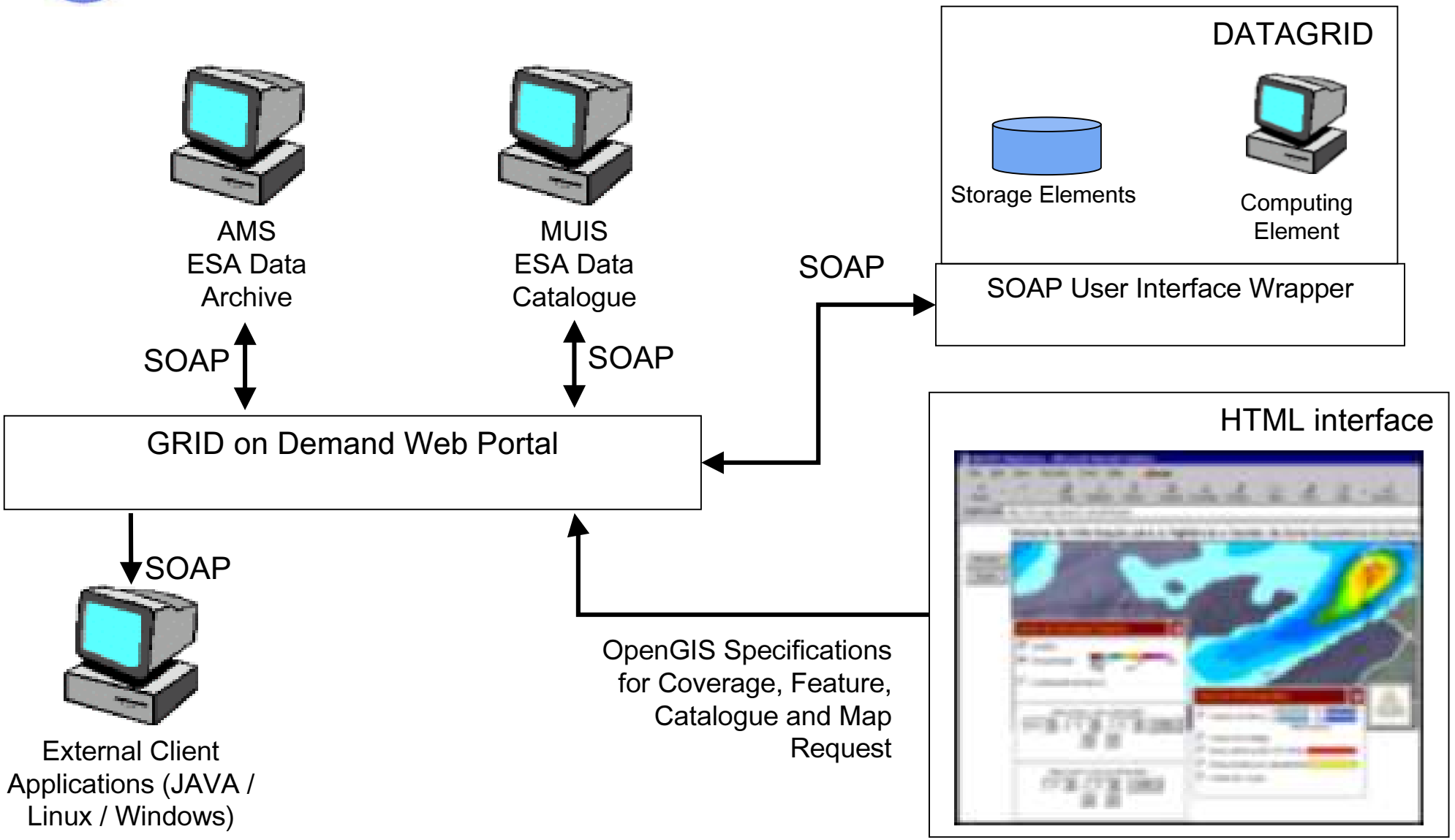
Discussion of Special Topics: User Interface tools and components

- Currently prototyping in ESA
  - EO Application interface components and scripts
  - Java GUI
  - EO Web / SOAP Services and GIS
  - Interface for transfering files to / from AMS

Discussion of Special Topics: Planning for demonstrations till end of year

- Re-run of 10 July demonstrations using 1.2 Testbed
- GRID surfer and WEB portal services
- ESTEC demo
- IST year conference
- 6th FWP launch

Discussion of Special Topics: Web-map GOME portal: development of Web-services for EO



Discussion of Special Topics: Participation to CEOS GRID testbed using DataTAG networking guidelines

- Status report
- Participation of US partners to DataGrid WP9

Discussion of Special Topics: Dissemination and extension to other applications

- ESA aims to enlarge the number of different application uses of the EDG Testbed
  - Dutch Space OMI simulation (workflow): funded by ESA as separate contract. OMI is new instrument (funded by NL) to fly on NASA platform
  - SSI / Alenia / ACS tbd (number crunching): funded by ESA as separate contract. Interest in demonstration of possible use of QUADRIX very high performing MPI
  - LAMMA hi-res weather forecast with integration of EO high resolution data. Extend cooperation in other EC funded project
  - University College London interested in demonstration SAR interferometry processing on GRID environment

Discussion of Special Topics: Interaction With EO Scientists

- The DataGrid testbed offers the opportunity for EO scientists to perform large-scale processing
- However it is agreed that initial results of the EO scientific algorithms using the Grid will be considered only for evaluation of the Grid performance