The EGEE project: enabling e-Science

Mike Mineter
NeSC Edinburgh
mjm@nesc.ac.uk

EGEE is a project co-funded by the European Commission under contract INFSO-RI-508833

Acknowledgements

This presentation includes slides and information from :

- Fabrizio Gagliardi and Bob Jones (UK AHM 2004 talk)
- Roberto Barbera (Slides on applications)
- Other colleagues in EGEE

- Additional slides and preparation by Mike Mineter

Contents

- EGEE
  - *The vision - why EGEE got started*
  - *Where are we now?*
  - *Where are we going? (project goals)*
  - *How will we get there? (project structure and activities)*
- Current virtual organisations (i.e. application communities)
- Some other important questions about EGEE

Contents

Enabling Grids for E-science in Europe

- EGEE
  - *The vision - why EGEE got started*
  - *Where are we now?*
  - *Where are we going? (project goals)*
  - *How will we get there? (project structure and activities)*
- Current virtual organisations (i.e. application communities)
- Some other important questions about EGEE

Despite its name, EGEE has a scope much wider than Europe: it is an International project with partners world-wide

Contents

- EGEE
  - *The vision - why EGEE got started*
  - *Where are we now?*
  - *Where are we going? (project goals)*
  - *How will we get there? (project structure and activities)*
- Current virtual organisations (i.e. application communities)
- Some other important questions about EGEE

Despite its name EGEE has a scope much wider than e-Science. It is a an international project with partners world-wide, and is intended to also support non-scientific research and collaborations in industry, the public sector, …

Contents

Enabling Grids for E-science in Europe

- EGEE
  - *The vision - why EGEE got started*
  - *Where are we now?*
  - *Where are we going? (project goals)*
  - *How will we get there? (project structure and activities)*
- Current virtual organisations (i.e. application communities)
- Some other important questions about EGEE

However, “EGEE” *is* much better than “EGERIPSEWW”

Contents

- EGEE
  - *The vision - why EGEE got started*
  - *Where are we now?*
  - *Where are we going? (project goals)*
  - *How will we get there? (project structure and activities)*
- Current virtual organisations (i.e. application communities)
- Some other important questions about EGEE

Despite having very clear targets for March 2006, the goal of EGEE is to create an infrastructure that will be sustainable, far beyond the end of its initial phase of funding.

What is our vision of "the Grid"?

Enabling Grids for E-science in Europe

- The World Wide Web provides access to information that is stored in many millions of different geographical locations

- "The Grid" is an infrastructure which provides access to computing power and data distributed over the globe

  and supports collaboration within virtual organisations

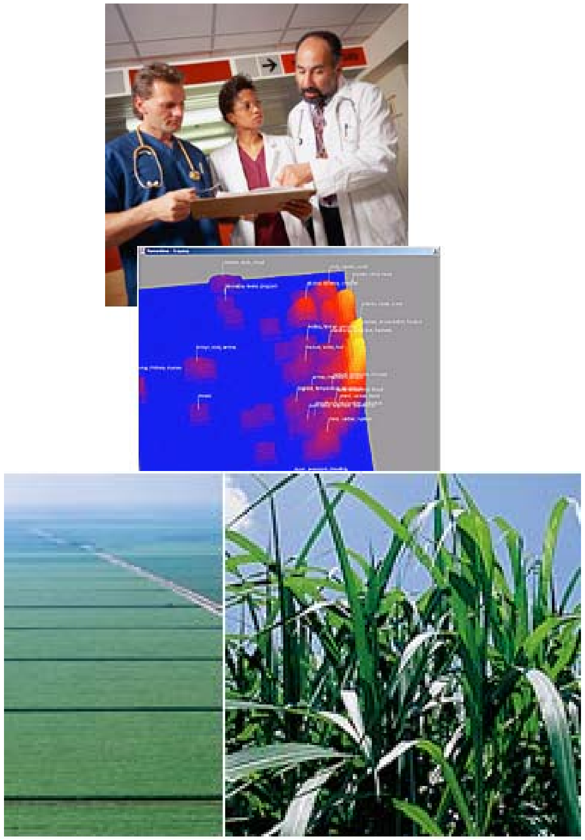
- The name Grid is chosen by analogy with the electric power grid

- Enabled by middleware: the "operating system of the Grid"

What is driving grid development?

Data and compute intensive sciences are next generation applications that have extreme needs but are likely to become mainstream in the next 5 years (Increasingly, also: collaboration intensive)

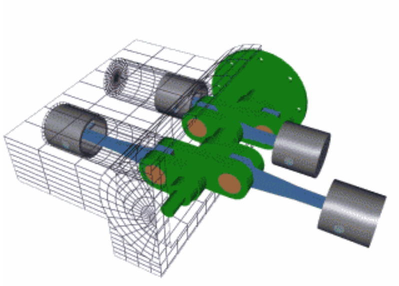
- Physics/Astronomy: data from different kinds of research instruments
- Medical/Healthcare: imaging, diagnosis and treatment
- Bioinformatics: study of the human genome and proteome to understand genetic diseases
- Nanotechnology: design of new materials from the molecular scale
- Engineering: design optimization, simulation, failure analysis and remote Instrument access and control
- Natural Resources and the Environment: weather forecasting, earth observation, modeling and prediction of complex systems: river floods and earthquake simulation



The Vision

- An international network of scientists will be able to model a new flood of the Danube in real time, using meteorological and geological data from several centers across Europe

- A team of engineering students will be able to run the latest 3D rendering programs from their laptops using the Grid.



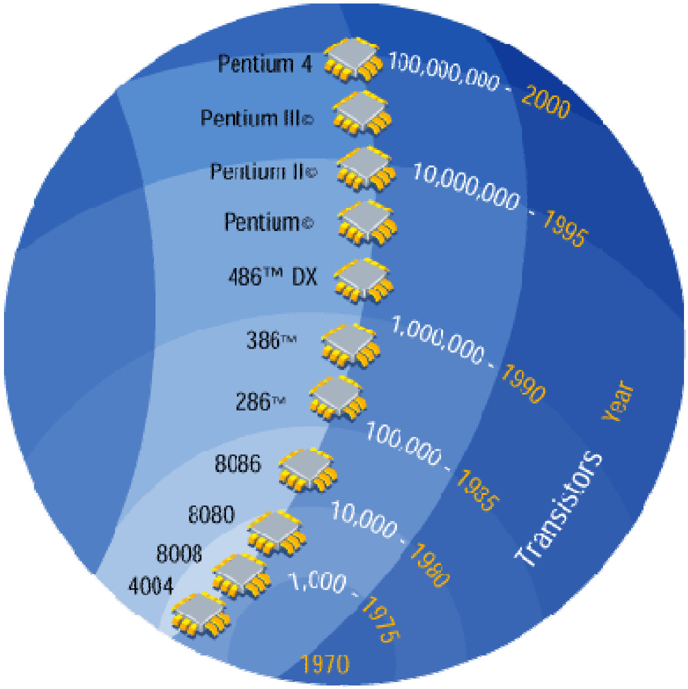
- A geneticist at a conference, inspired by a talk she hears, will be able to launch a complex bio-molecular simulation from her mobile phone

Access to a production quality GRID will change the way science and much else is done

The Grid: why now?

- Networking, commodity computing and distributed software tools are ripe for Grid technology
- Science more digital oriented and dominated by data
- Many public funded projects in the US and in the EU
- Also industrial and commercial Grids (for examples: www.cern.ch/gridcafe and www.gridstart.org)

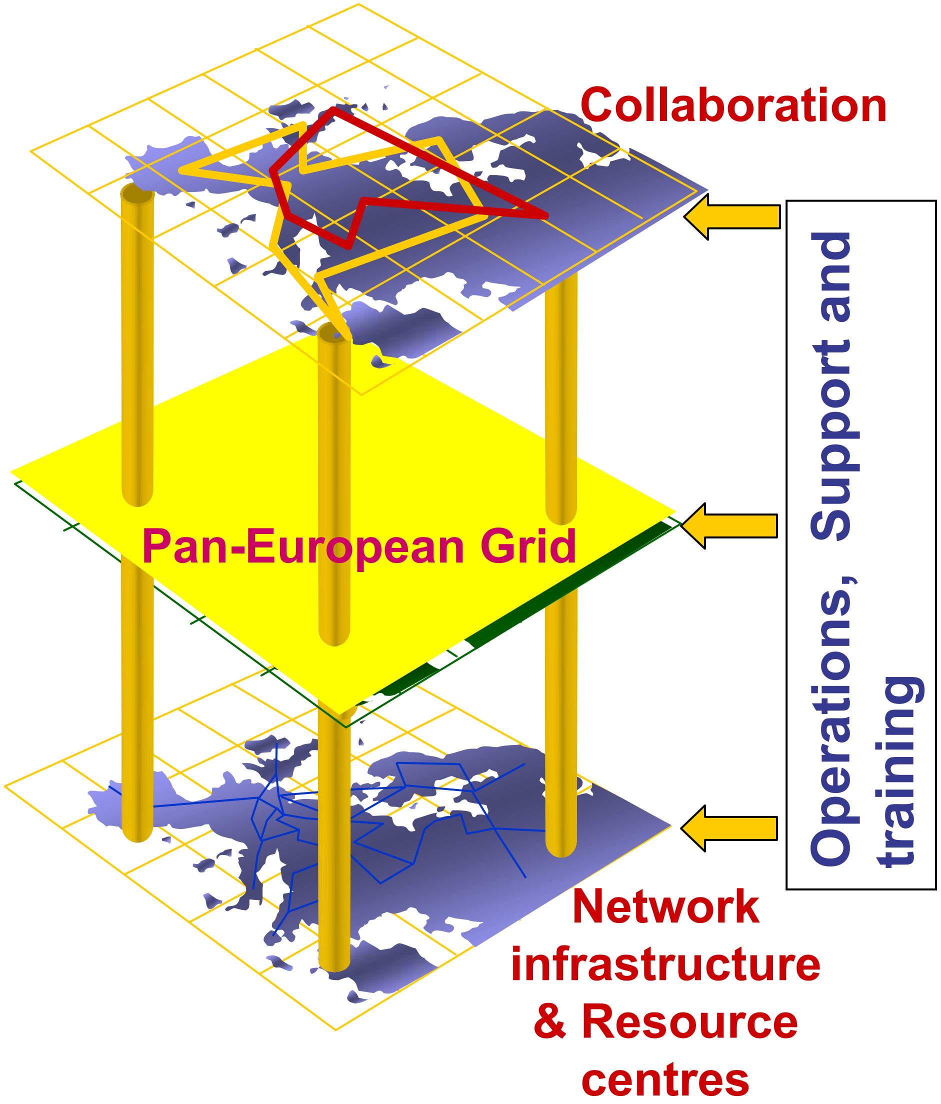
- Consequently: the EU has set the goal of creating e-Infrastructure



EGEE – towards e-infrastructure

Build a large-scale production grid service to:

- Underpin European science and technology
- Link with and build on national, regional and international initiatives
- Foster international cooperation both in the creation and the use of the e-infrastructure



Collaboration

Pan-European Grid

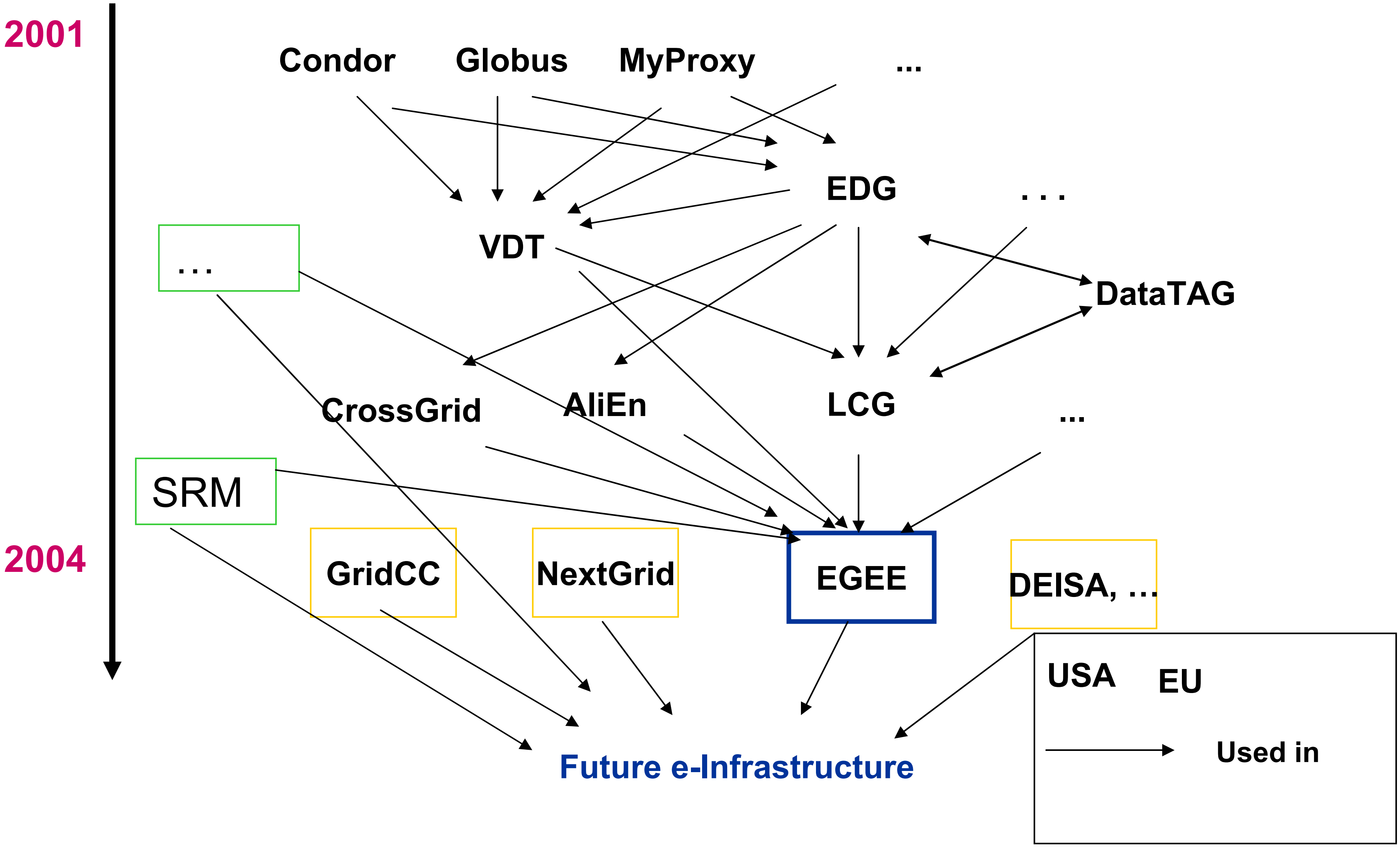
Network infrastructure & Resource centres

Operations, Support and training

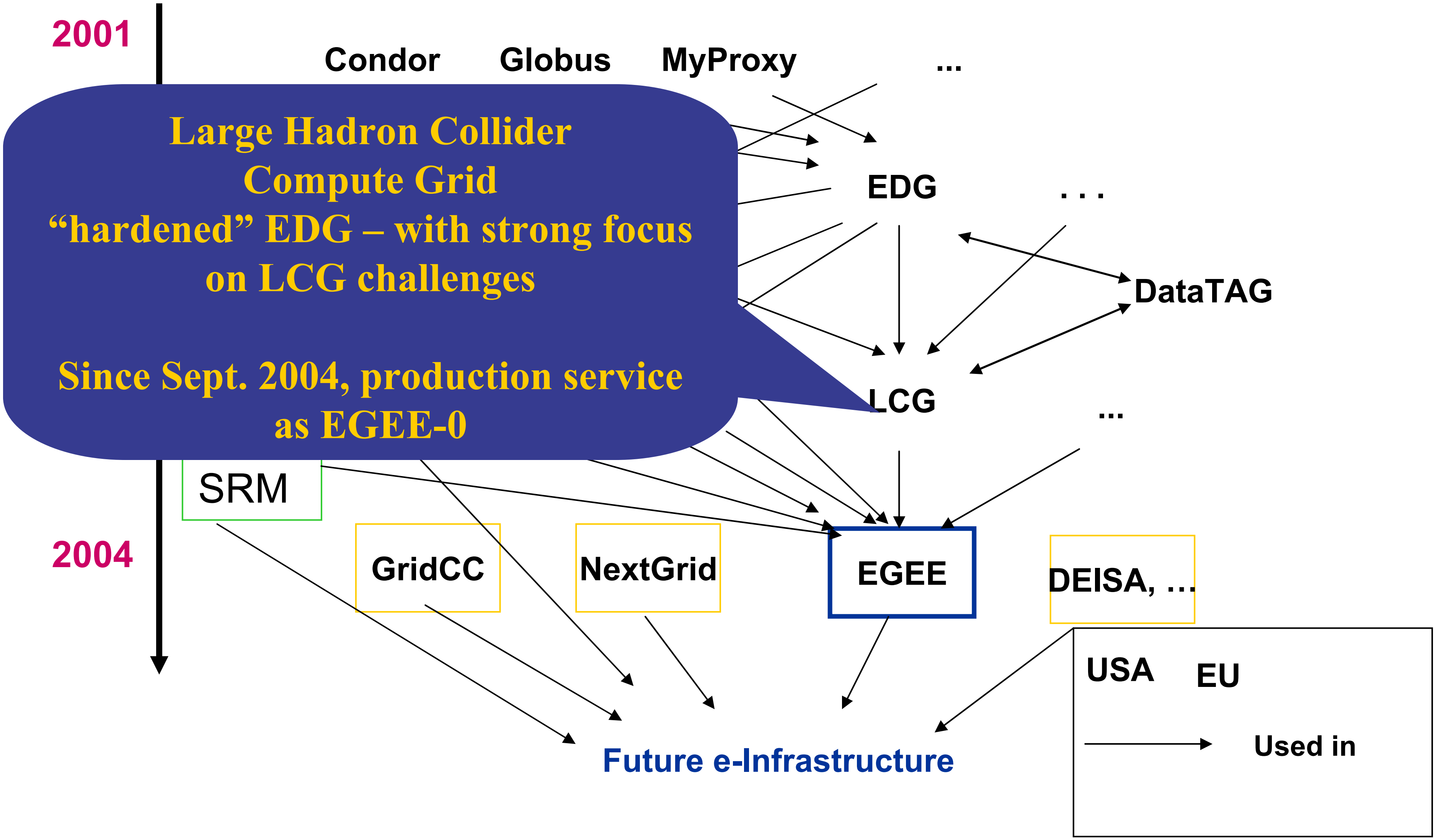
After the vision: where are we now?

- EGEE
  - *The vision - why EGEE got started*
  - ***Where are we now?***
  - *Where are we going? (project goals)*
  - *How will we get there? (project structure and activities)*
- Current virtual organisations
  (i.e. application communities)
- Some other important questions about EGEE

Part of the Grid "ecosystem"

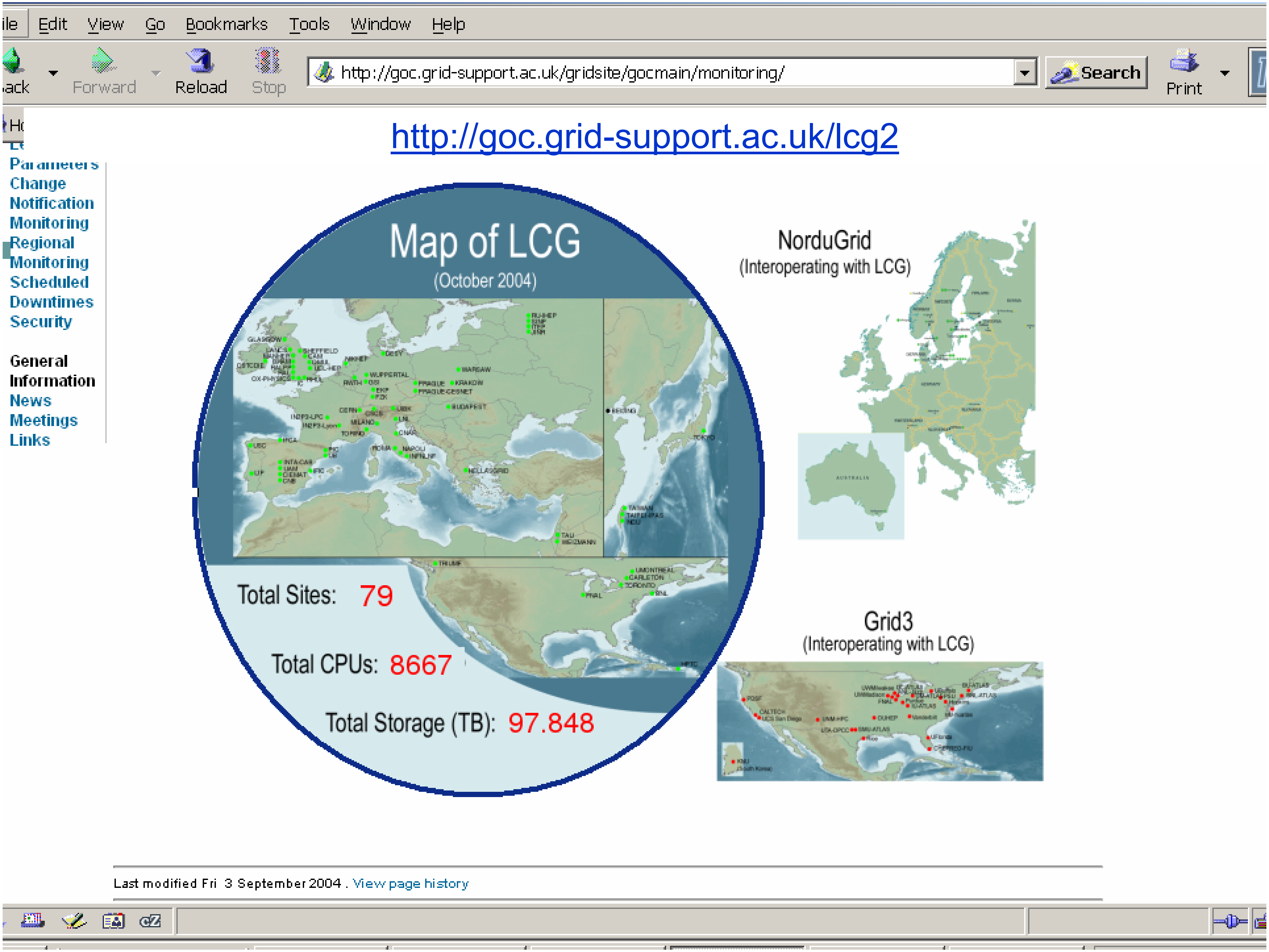


Part of the Grid "ecosystem"



Large Hadron Collider Compute Grid "hardened" EDG – with strong focus on LCG challenges

Since Sept. 2004, production service as EGEE-0



http://goc.grid-support.ac.uk/gridsite/gocmain/monitoring/

http://goc.grid-support.ac.uk/lcg2

Map of LCG

(October 2004)

Total Sites: 79

Total CPUs: 8667

Total Storage (TB): 97.848

NorduGrid

(Interoperating with LCG)

Grid3

(Interoperating with LCG)

Last modified Fri 3 September 2004 . View page history

CMS Data Challenge

Running with LCG-2 and CMS resources world-wide
(US Grid3 was a major component)

Pre-Challenge (Phase 1)

- After 8 months of continuous running:
  - 750,000 jobs
  - 3,500 KSI2000 months
  - 700,000 files
  - 80 TB of data

Data Challenge (Phase 2)

- 2,200 jobs/day (about 500 CPU's)
- Total 45,000 jobs

- 0.4 files/s registered to RLS
- Total 570,000 files registered to RLS

- 4 MB/s produced and distributed

Running the Production Service

Grid deployment has entered a new phase

- **Basic middleware is working**
  - responsible now for a small fraction of the problems
- **Outstanding performance/functionality issues**
  - RLS, RB / little modularity & lack of consistent interfaces …
  - some solutions are being developed but many cannot be addressed in current software/architecture - *set priorities for new middleware* (gLite)
- **Many operational issues**
  - mis-configuration, out of date mware, single points of failure, failover, mgmt interfaces …
  - resources unsuitable for applications needs (e.g. insufficient disk space)
  - slow response by sites to problems (holiday periods, security concerns)
  - new middleware will not help for many of these issues - grid partners must think ***Service***

The grid still does not appear as a single coherent facility

applications must adapt to the current service to gain maximum profit

but result has been very effective for LHCb - ~3000 concurrent jobs (August)

Grids for eInfrastructure…

Enabling Grids for E-science in Europe

- **What is missing?**
  - **Production-quality (stable, mature) Grid middleware**
  - **Production-quality operational support**
    - Grid Operation Centres, Helpdesks, etc.
  - **Multi-discipline grid-enabled application environment**
    - Now led by HEP, Bio-info
  - **Administrative and policy decision framework in order to share resources at pan-European scale (and beyond)**
    - Areas such as AAA (Authentication, Authorisation, Accounting)
    - End-to-end issues (Network related)
    - Funding Policies (Grid economics)
    - Resource Sharing Policies
    - Usage Policies

- EGEE project will tackle most of the above issues

Contents

- EGEE
  - *The vision - why EGEE got started*
  - *Where are we now?*
  - ***Where are we going? (project goals)***
  - *How will we get there? (project structure and activities)*
- Current virtual organisations (i.e. application communities)
- Some other important questions about EGEE

What will EGEE provide?

- Simplified access *(access to all the operational resources the user needs)*

- On demand computing *(fast access to resources by allocating them efficiently)*

- Pervasive access *(accessible from any geographic location)*

- Large scale resources *(of a scale that no single computer centre can provide*)

- Sharing of software and data *(in a transparent way)*

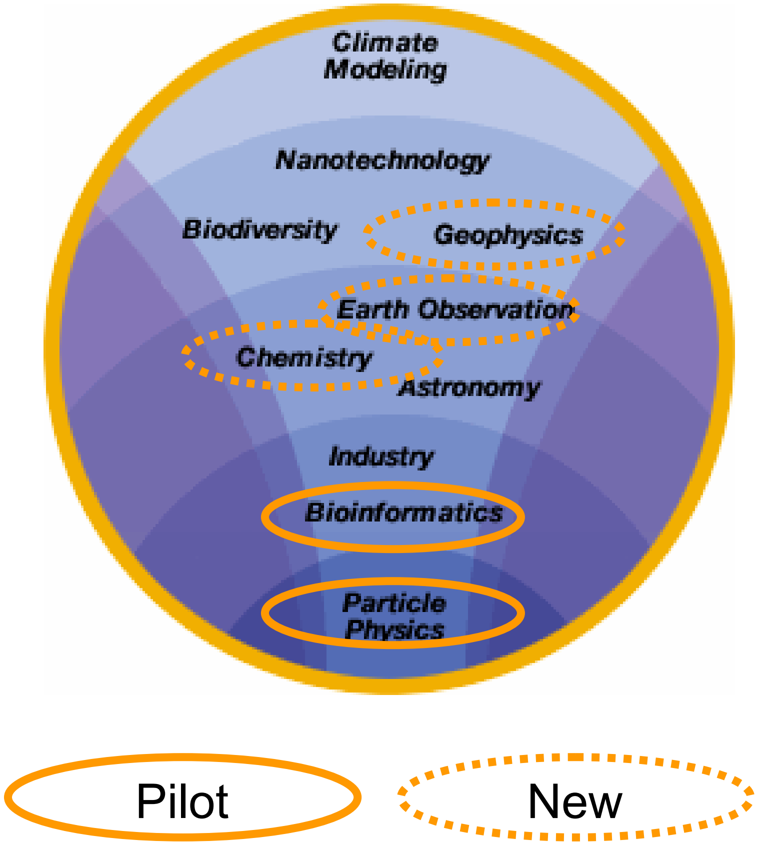
- Improved support *(use the expertise of all partners to offer in-depth support for all key applications)*

In 2 years EGEE will:

- **Establish production quality sustained Grid services**
  - 3000 users from at least 5 disciplines
  - over 8,000 CPU's, 50 sites
  - over 5 Petabytes (10^{15}) storage

- Demonstrate a viable general process to **bring other scientific communities on board**

- **Propose a second phase** in mid 2005 to take over EGEE in early 2006

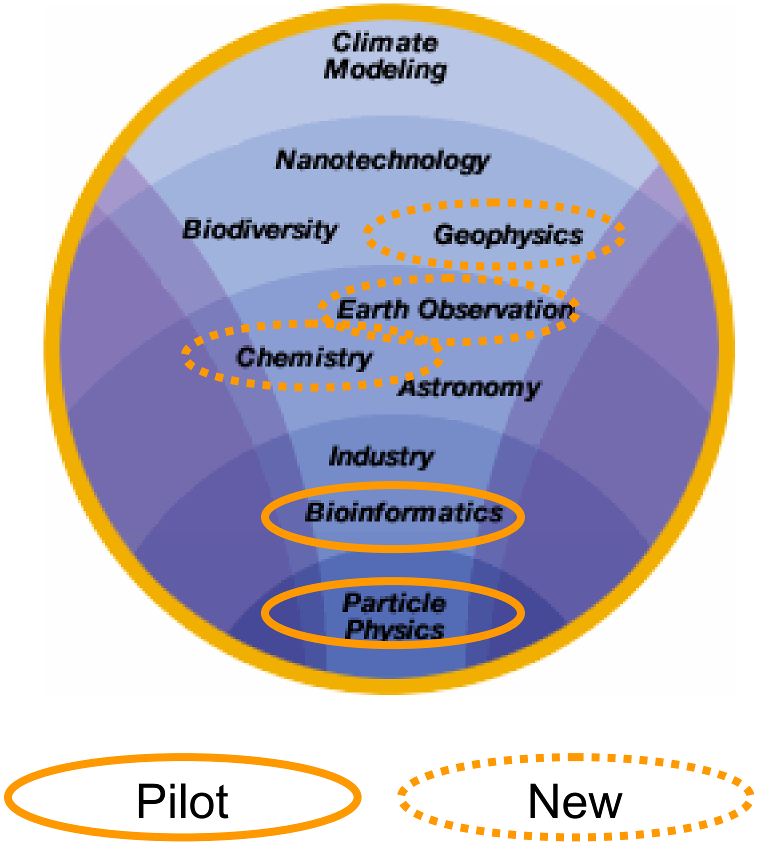


In 2 years EGEE will:

- **Establish production quality sustained Grid services**
  - Reliable and secure
  - 24 hr/day; 7 day/week
  - Sustained: ~20 years

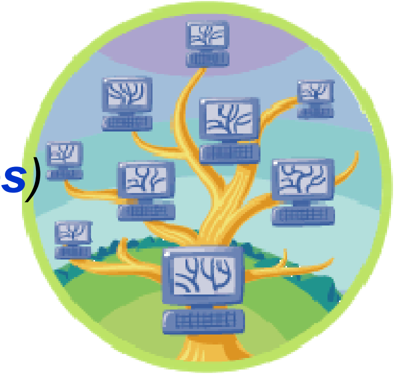
- Demonstrate a viable general process to **bring other scientific communities on board**

- **Propose a second phase** in mid 2005 to take over EGEE in early 2006



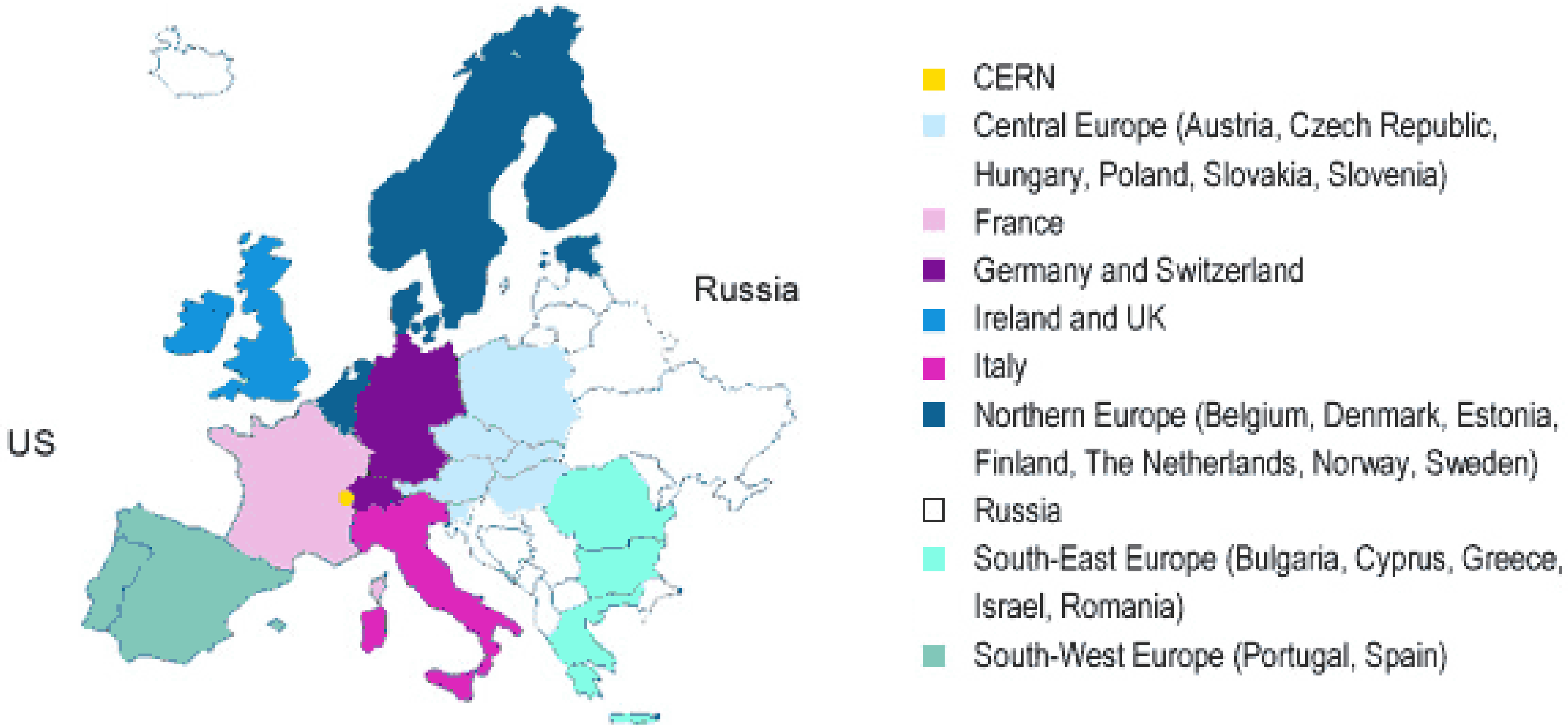
Contents

- EGEE
  - *The vision - why EGEE got started*
  - *Where are we now?*
  - *Where are we going? (project goals)*
  - ***How will we get there? (project structure and activities**)*
- Current virtual organisations (i.e. application communities)
- Some other important questions about EGEE



EGEE Figures & Organization

- Coordinator: European Organization for Nuclear Research - **CERN**
- **70** leading institutions in **27** countries, federated in regional Grids
- **32 M € EU** funding in 2004-2005 (twice from partners)



■ CERN

■ Central Europe (Austria, Czech Republic, Hungary, Poland, Slovakia, Slovenia)

■ France

■ Germany and Switzerland

■ Ireland and UK

■ Italy

■ Northern Europe (Belgium, Denmark, Estonia, Finland, The Netherlands, Norway, Sweden)

□ Russia

■ South-East Europe (Bulgaria, Cyprus, Greece, Israel, Romania)

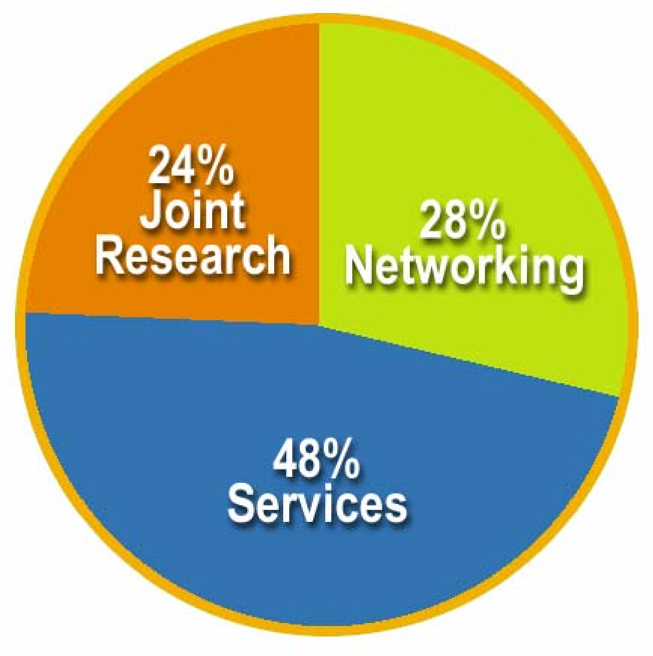
■ South-West Europe (Portugal, Spain)

EGEE Activities

Enabling Grids for E-science in Europe

32 Million Euros EU funding over 2 years starting 1st April 2004

- **48 % service activities** (Grid Operations, Support and Management, Network Resource Provision)
- **24 % middleware re-engineering** (Quality Assurance, Security, Network Services Development)
- **28 % networking** (Management, Dissemination and Outreach, User Training and Education, Application Identification and Support, Policy and International Cooperation)



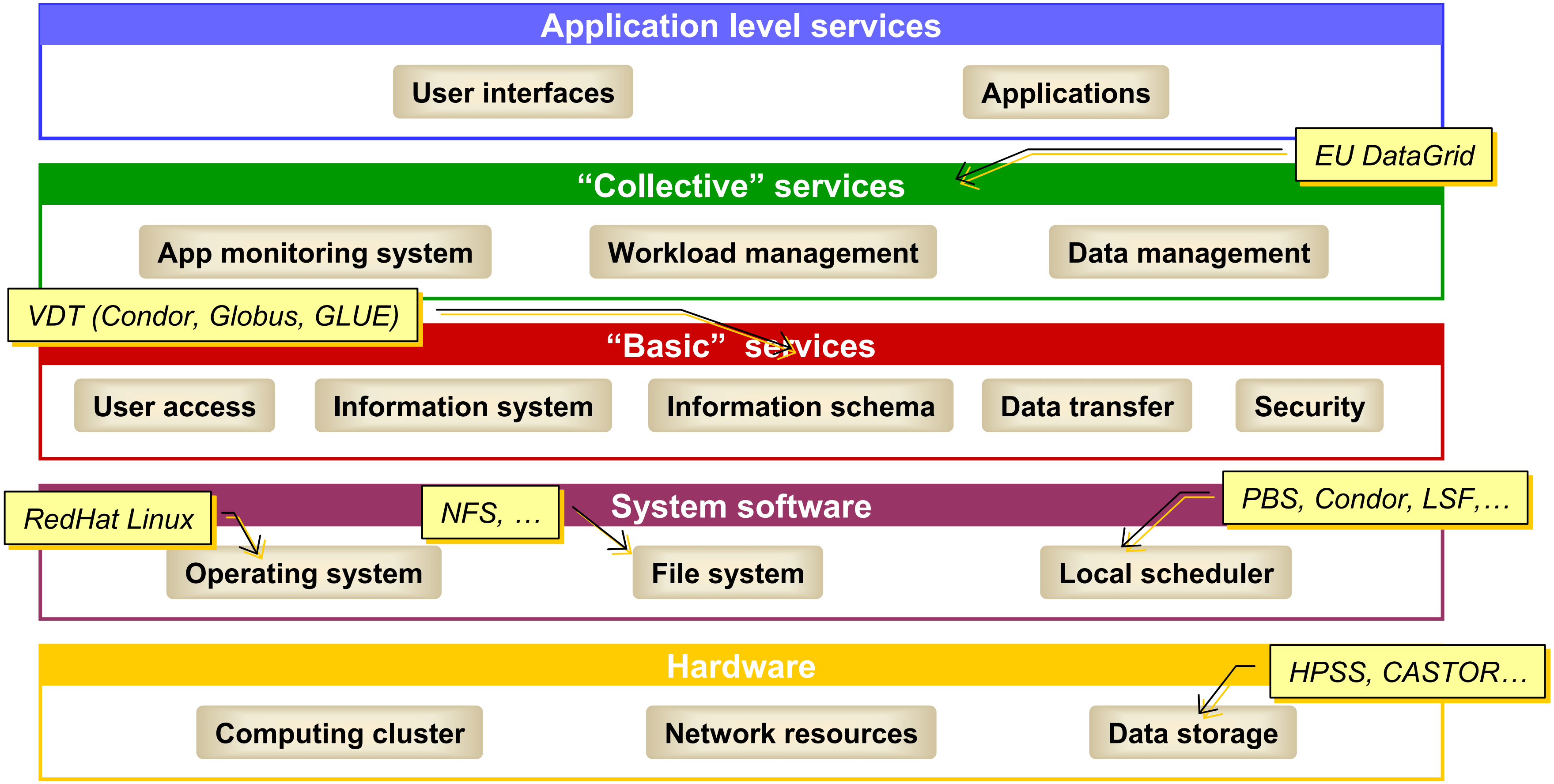
Emphasis in EGEE is on operating a production grid and supporting the end-users

Contents

- EGEE
  - *The vision - why EGEE got started*
  - *Where are we now?*
  - *Where are we going? (project goals)*

  - ***Project structure and activities***
    - **Middleware – *current and future***
    - Operations – *providing a production service*
    - Networking – *enabling multiple effective VO's*
- Current virtual organisations (i.e. application communities)
- Some other important questions about EGEE

Current production mware: LCG-2



Application level services

User interfaces

Applications

EU DataGrid

“Collective” services

App monitoring system

Workload management

Data management

VDT (Condor, Globus, GLUE)

“Basic” services

User access

Information system

Information schema

Data transfer

Security

RedHat Linux

NFS, …

System software

PBS, Condor, LSF,…

Operating system

File system

Local scheduler

Hardware

HPSS, CASTOR…

Computing cluster

Network resources

Data storage

gLite

- “gLite” - the new EGEE middleware (under test)
- Service oriented - components that are :
  - Loosely coupled (by messages)
  - Accessible across network; modular and self-contained; clean modes of failure
  - So can change implementation without changing interfaces
  - Can be developed in anticipation of new uses
- … and are based on standards.
  Opens EGEE to:
  - New middleware (plethora of tools now available)
  - Heterogeneous resources (storage, computation…)
  - Interact with other Grids (international, regional, national and thematic)

Contents

Enabling Grids for E-science in Europe

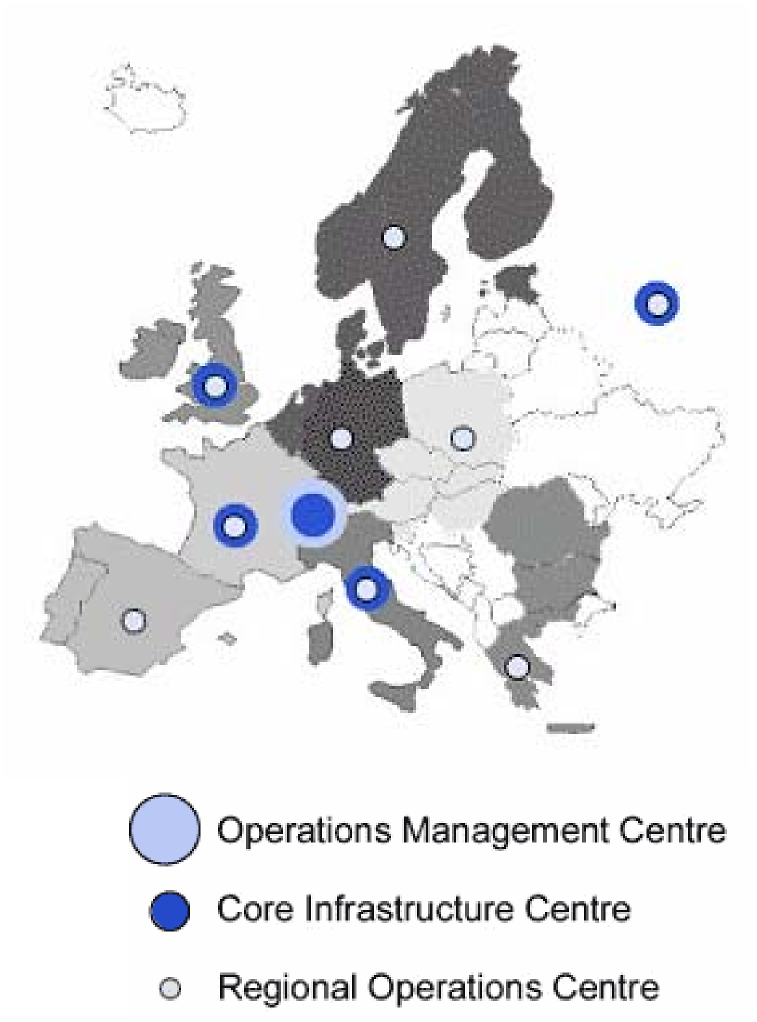
- EGEE
  - *The vision - why EGEE got started*
  - *Where are we now?*
  - *Where are we going? (project goals)*

  - ***Project structure and activities***
    - Middleware – current and future
    - **Operations – *providing a production service***
    - Networking – *enabling multiple effective VO's*

- Current virtual organisations (i.e. application communities)
- Some other important questions about EGEE

EGEE Service Activities

- Create, operate, support and manage a production quality infrastructure
- Offered services:
  - Middleware deployment and installation
  - Software and documentation repository
  - Grid monitoring and problem tracking
  - Bug reporting and knowledge database
  - VO services
  - Grid management services


○ Operations Management Centre
● Core Infrastructure Centre
○ Regional Operations Centre

Contents

- EGEE
  - *The vision - why EGEE got started*
  - *Where are we now?*
  - *Where are we going? (project goals)*

  - ***Project structure and activities***
    - Middleware – current and future
    - Operations – providing a production service
    - **(Human) Networking – *enabling multiple effective VO's***

- Current virtual organisations
  (i.e. application communities)
- Some other important questions about EGEE

Bringing new applications to the grid

1. ***Outreach events*** inform people about the grid / EGEE
2. Application experts discuss ***specific characteristics*** with the users
3. ***Migrate application*** to EGEE infrastructure with the help of EGEE experts
4. ***Initial deployment*** for testing purposes
5. Production usage - user community contributes computing resources for heavy production demands - "***Canadian dinner party***"

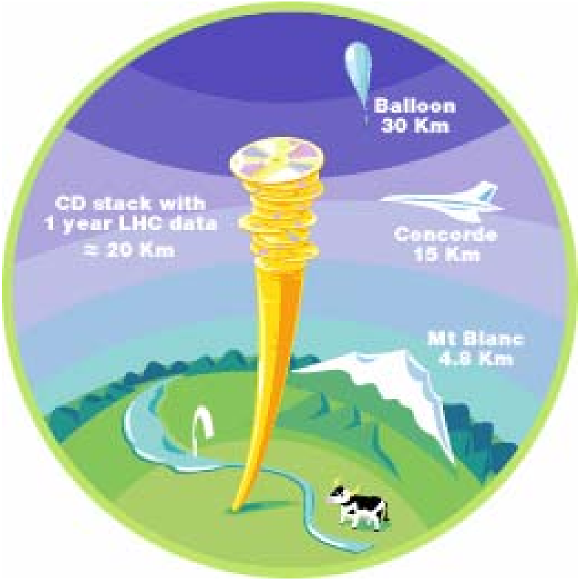
…. Supported by training and regional operations as well as by applications experts

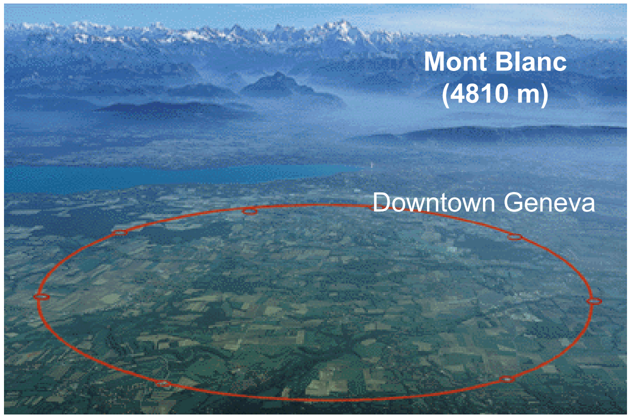
Contents

- EGEE
  - *The vision - why EGEE got started*
  - *Where are we now?*
  - *Who is "we"?*
  - *Where are we going? (project goals)*
  - *How will we get there? (project structure and activities)*
- **Current virtual organisations (i.e. application communities)**
- Some other important questions about EGEE

EGEE pilot application: Large Hadron Collider

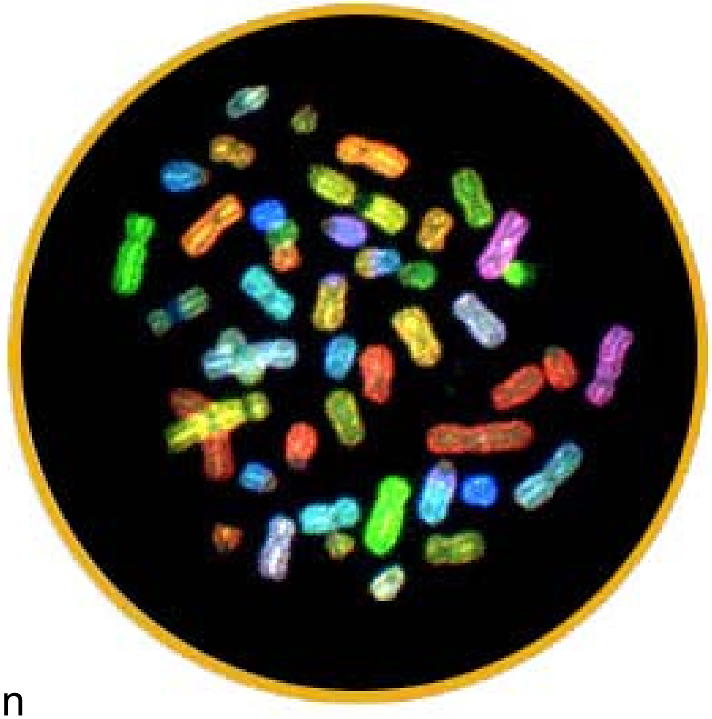
- Data Challenge:
  - 10 Petabytes/year of data !!!
  - 20 million CDs each year!
- Simulation, reconstruction, analysis:
  - LHC data handling requires computing power equivalent to ~100,000 of today's fastest PC processors!
- Operational challenges
  - Reliable and scalable through project lifetime of decades





EGEE pilot application: BioMedical

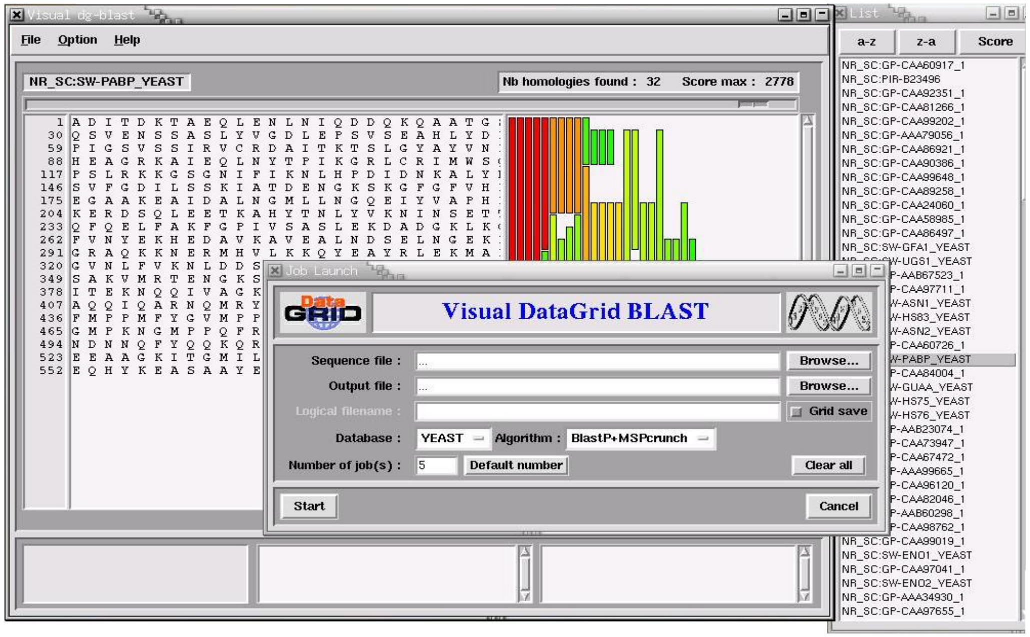
- BioMedical
  - Bioinformatics (gene/proteome databases distributions)
  - Medical applications (screening, epidemiology, image databases distribution, etc.)
  - Interactive application (human supervision or simulation)
  - Security/privacy constraints
    - Heterogeneous data formats - Frequent data updates - Complex data sets - Long term archiving
- BioMed applications deployed
  - **GATE -** Geant4 Application for Tomographic Emission
  - **GPS@ -** genomic web portal
  - **CDSS -** Clinical Decision Support System



http://egee-na4.ct.infn.it/biomed/applications.html

BLAST – comparing DNA or protein sequences

- BLAST is the first step for analysing new sequences: to compare DNA or protein sequences to other ones stored in personal or public databases.
- Ideal as a grid application – trivial to parallelise as independent concurrent jobs on one or more CEs.
  - Requires resources to store databases and run algorithms
  - Large user community



NR_SC:SW-PABP_YEAST

Nb homologies found : 32 Score max : 2778

```
  1 A D I T D K T A E Q L E N L N I Q D D Q K Q A A T G
 30 Q S V E N S S A S L Y V G D L E P S V S E A H L Y D
 59 P I G S V S S I R V C R D A I T K T S L G Y A Y V N
 88 H E A G R K A I E Q L N Y T P I K G R L C R I M W S
117 P S L R K K G S G N I F I K N L H P D I D N K A L Y
146 S V F G D I L S S K I A T D E N G K S K G F G F V H
175 E G A A K E A I D A L N G M L L N G Q E I Y V A P H
204 K E R D S Q L E E T K A H Y T N L Y V K N I N S E T
233 Q F Q E L F A K F G P I V S A S L E K D A D G K L K
262 F V N Y E K H E D A V K A V E A L N D S E L N G E K
291 G R A Q K K N E R M H V L K K Q Y E A Y R L E K M A
320 G V N L F V K N L D D S
349 S A K V M R T E N G K S
378 I T E K N Q Q I V A G K
407 A Q Q I Q A R N Q M R Y
436 F M P P M F Y G V M P P
465 G M P K N G M P P Q F R
494 N D N N Q F Y Q Q K Q R
523 E E A A G K I T G M I L
552 E Q H Y K E A S A A Y E
```

Sequence file : ...

Output file : ...

Logical filename :

☐ Grid save

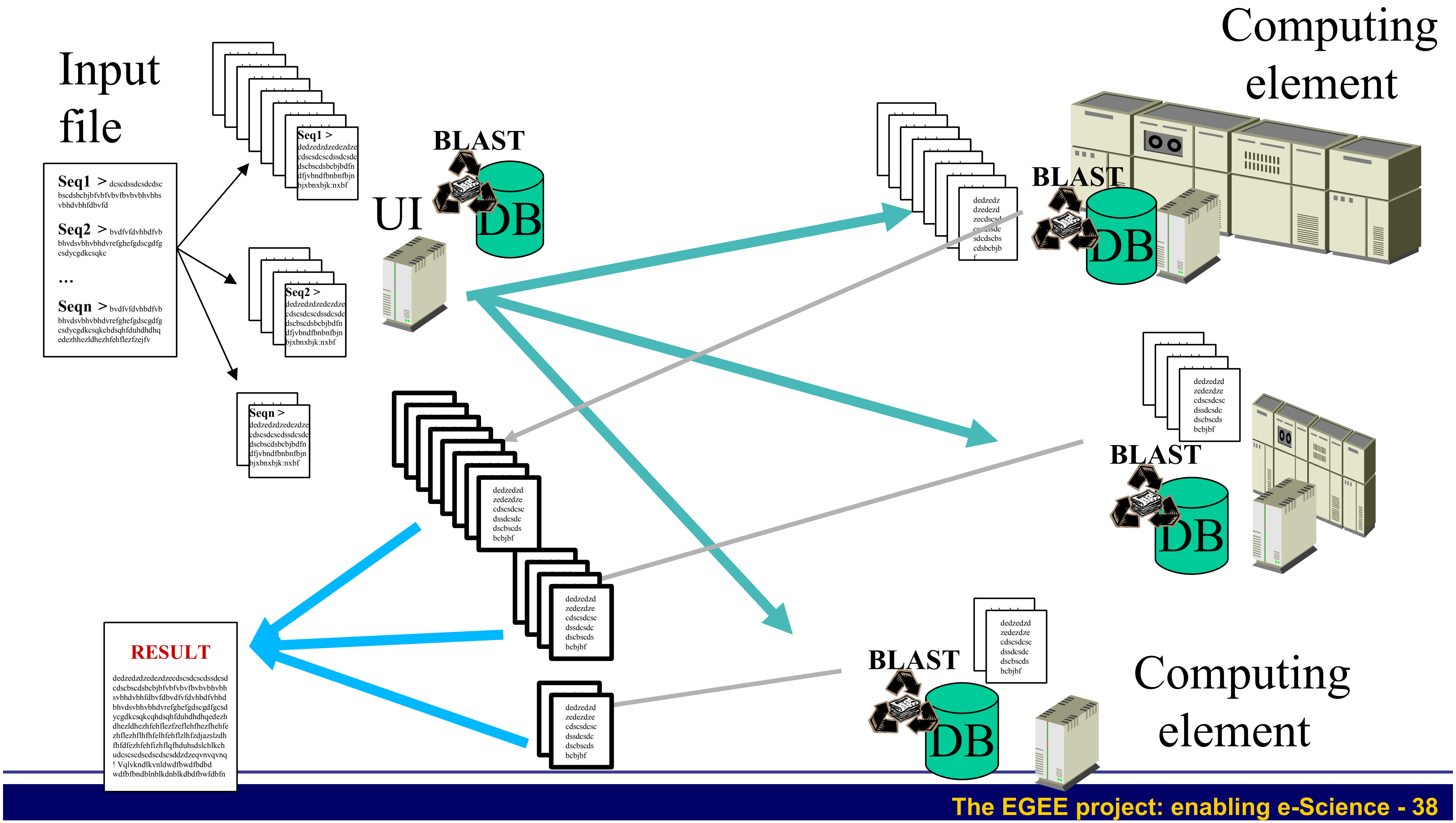
Database : YEAST

Algorithm : BlastP+MSPcrunch

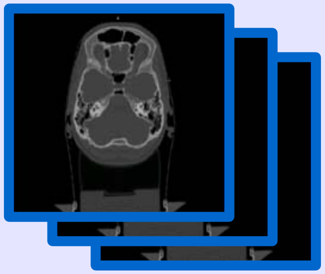
Number of job(s) : 5

```
NR_SC:GP-CAA60917_1
NR_SC:PIR-B23496
NR_SC:GP-CAA92351_1
NR_SC:GP-CAA81266_1
NR_SC:GP-CAA99202_1
NR_SC:GP-AAA79056_1
NR_SC:GP-CAA86921_1
NR_SC:GP-CAA90386_1
NR_SC:GP-CAA99648_1
NR_SC:GP-CAA89258_1
NR_SC:GP-CAA24060_1
NR_SC:GP-CAA58985_1
NR_SC:GP-CAA86497_1
NR_SC:SW-GFA1_YEAST
W-UGS1_YEAST
P-AAB67523_1
P-CAA97711_1
W-ASN1_YEAST
W-HS83_YEAST
W-ASN2_YEAST
P-CAA60726_1
W-PABP_YEAST
P-CAA84004_1
W-GUAA_YEAST
W-HS75_YEAST
W-HS76_YEAST
P-AAB23074_1
P-CAA73947_1
P-CAA67472_1
P-AAA99665_1
P-CAA96120_1
P-CAA82046_1
P-AAB60298_1
P-CAA98762_1
NR_SC:GP-CAA99019_1
NR_SC:SW-ENO1_YEAST
NR_SC:GP-CAA97041_1
NR_SC:SW-ENO2_YEAST
NR_SC:GP-AAA34930_1
NR_SC:GP-CAA97655_1
```

BLAST gridification


Input file
Seq1 > dcscdssdcsdcdsc
bscdsbcbjbfvbfvbvfbvbvbhvbhs
vbhdvbhfdbvfd
Seq2 > bvdfvfdvhbdfvb
bhvdsvbhvbhdvrefghefgdscgdfg
csdycgdkcsqkc
…
Seqn > bvdfvfdvhbdfvb
bhvdsvbhvbhdvrefghefgdscgdfg
csdycgdkcsqkchdsqhfduhdhdhq
edezhhezldhezhfehflezfzejfv
Seq1 >
Seq2 >
Seqn >
BLAST
UI
DB
Computing element
BLAST
DB
BLAST
DB
BLAST
DB
Computing element
RESULT
dedzedzdzedezdzecdscsdcscdssdcsd
cdscbscdsbcbjbfvbfvbvfbvbvbhvbh
svbhdvbhfdbvfdbvdfvfdvhbdfvbhd
bhvdsvbhvbhdvrefghefgdscgdfgcsd
ycgdkcsqkcqhdsqhfduhdhdhqedezh
dhezldhezhfehflezfzeflehfhezfhehfe
zhflezhflhfhfelhfehflzlhfzdjazslzdh
fhfdfezhfehfizhflqfhduhsdslchlkch
udcscscdscdscdscsddzdzeqvnvqvnq
! Vqlvkndlkvnldwdfbwdfbdbd
wdfbfbndblnblkdnblkdbdfbwfdbfn

Monte carlo simulation for radiotherapy planning



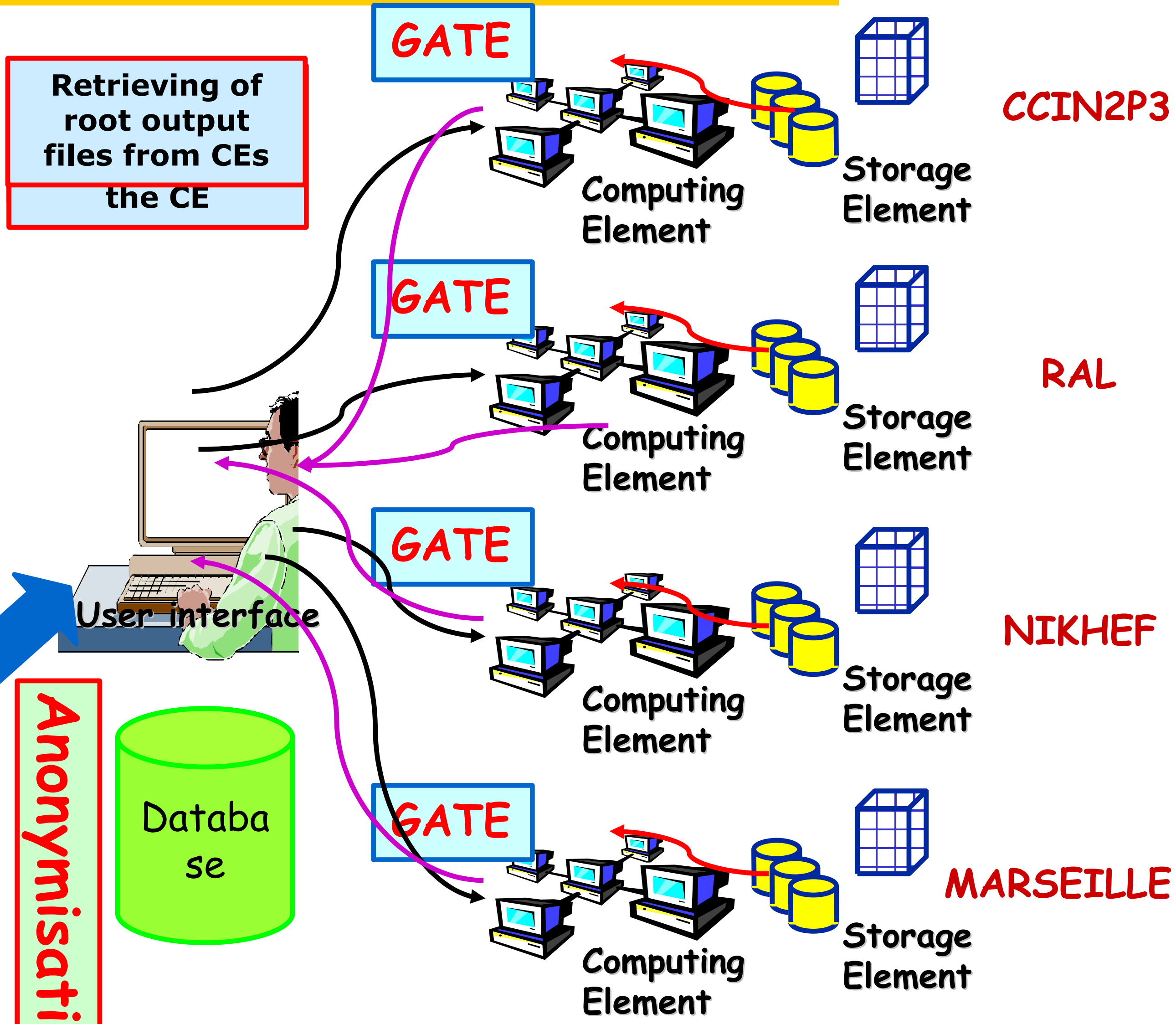
Scanner slices: DICOM format

Concatenation

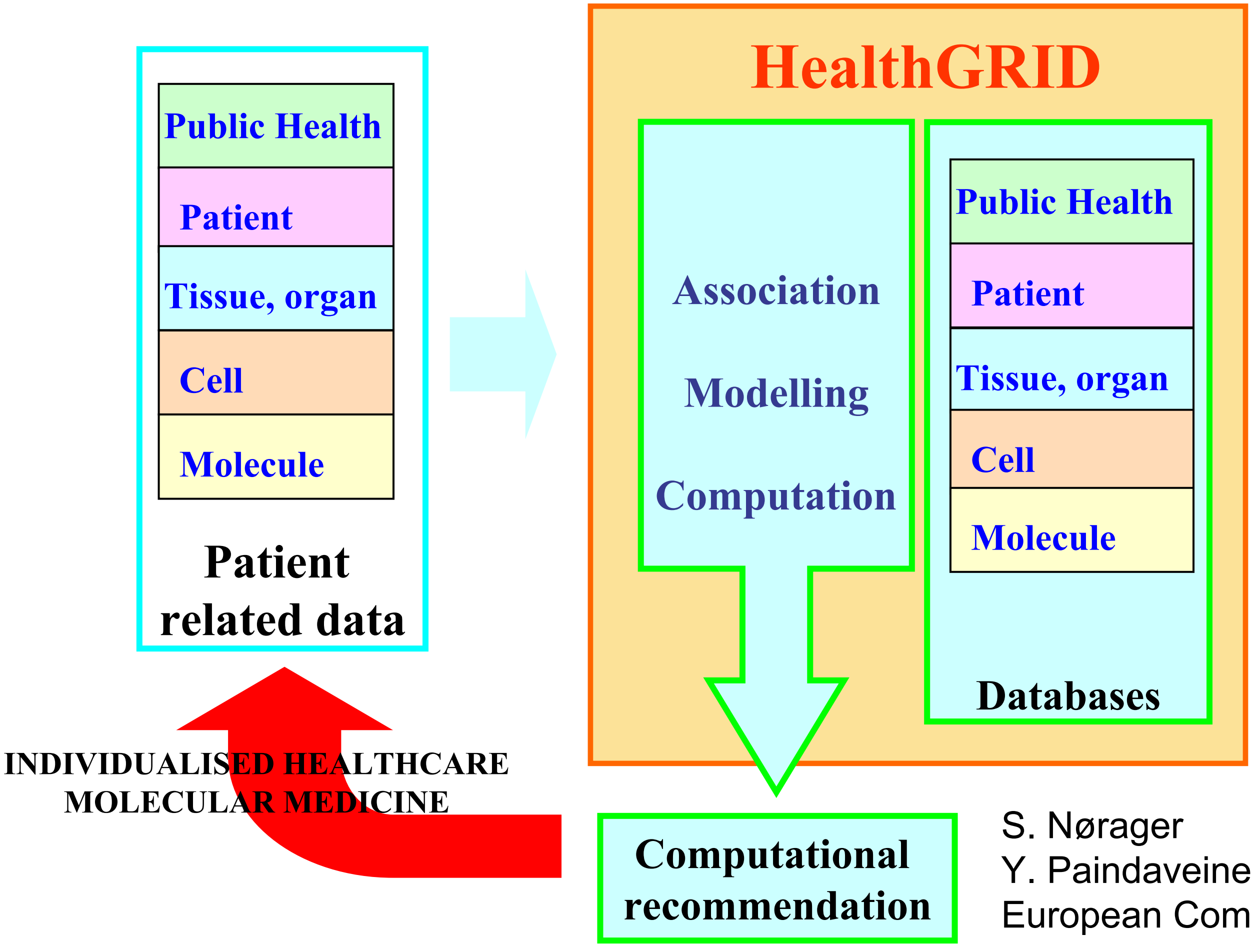
Image: text file

Binary file: Image.raw Size 19M

Anonymisation



A look at the future: the HealthGrid vision



Public Health

Patient

Tissue, organ

Cell

Molecule

Patient related data

HealthGRID

Association

Modelling

Computation

Public Health

Patient

Tissue, organ

Cell

Molecule

Databases

INDIVIDUALISED HEALTHCARE
MOLECULAR MEDICINE

Computational recommendation

In this context "Health" does not involve only clinical practice but covers the whole range of information from molecular level (genetic and proteomic information) over cells and tissues, to the individual and finally the population level (social healthcare).

S. Nørager
Y. Paindaveine
European Commission
DG-INFSO

Earth Sciences in EGEE

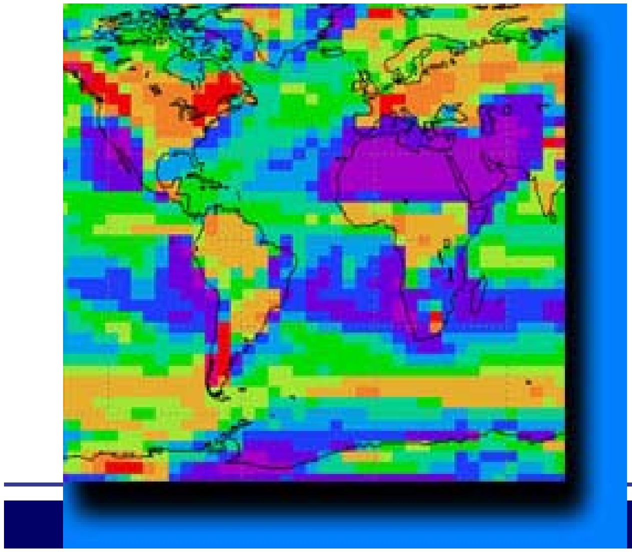
- Research
  - Earth observations by satellite
    - (ESA(IT), KNMI(NL), IPSL(FR), UTV(IT), RIVM(NL),SRON(NL))
  - Climate :
    - DKRZ(GE),IPSL(FR)
  - Solid Earth Physics:
    - IPGP (FR)
  - Hydrology:
    - Neuchâtel University (CH)
- Industry
  - CGG : Geophysics Company (FR)

Climate Applications in EGEE

Model: Atmosphere, Ocean, Hydrology, Atmospheric and Marine chemistry….

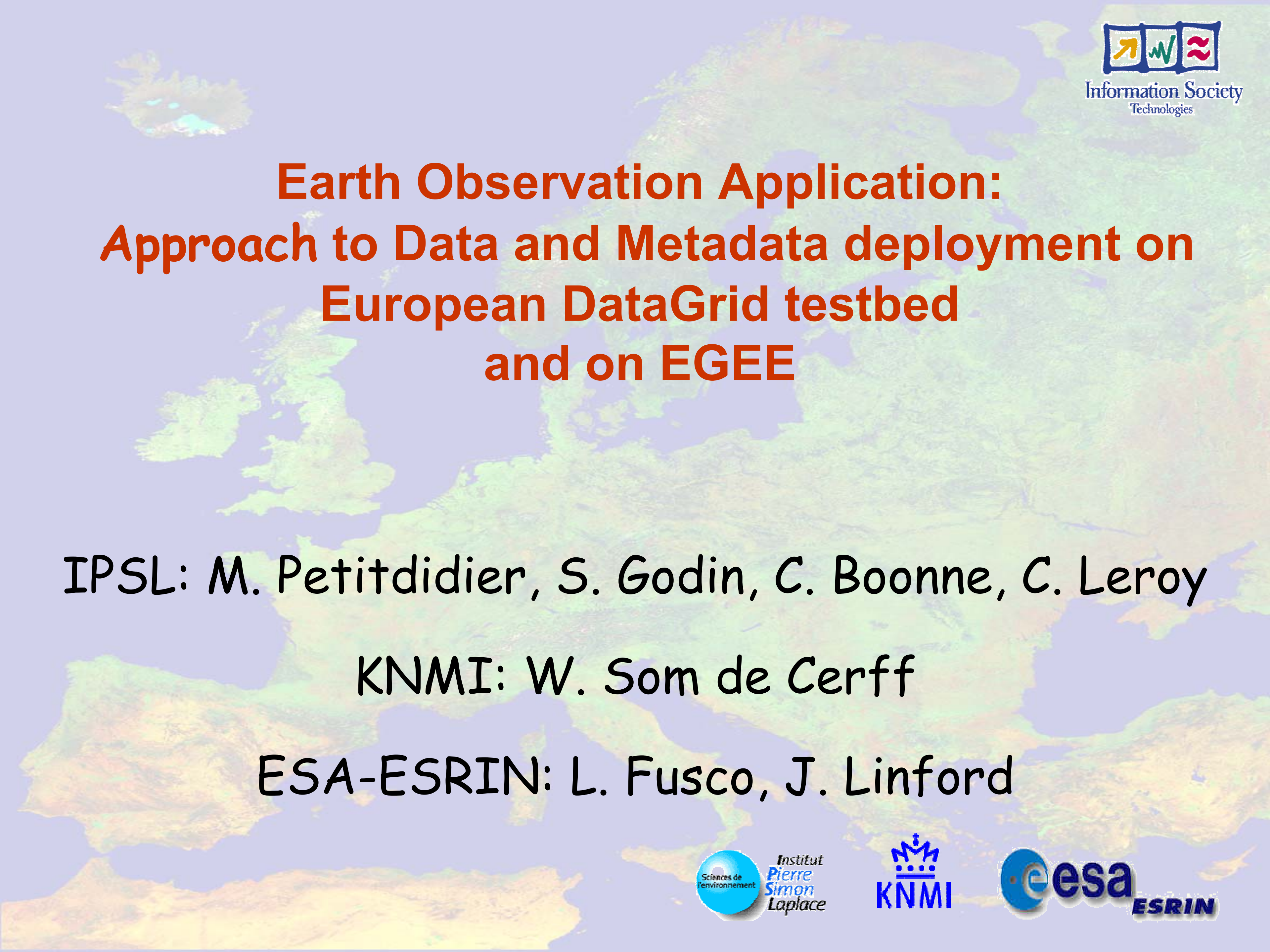
Goal: Comparison of model outputs from different runs and/or institutes

- Large volume of data (TB) from different model outputs, and experimental data
- Run made on supercomputer
  => Link the EGEE infrastruture with supercomputer Grids (DEISA)



EXAMPLE: For the IPCC Assessment reports many experiment are performed with different models (different spatial resolution, different time-step, different "physics" ..) and various sites.

The generated data need to be compared in a comprehensive and "unified" way.



Earth Observation Application: Approach to Data and Metadata deployment on European DataGrid testbed and on EGEE

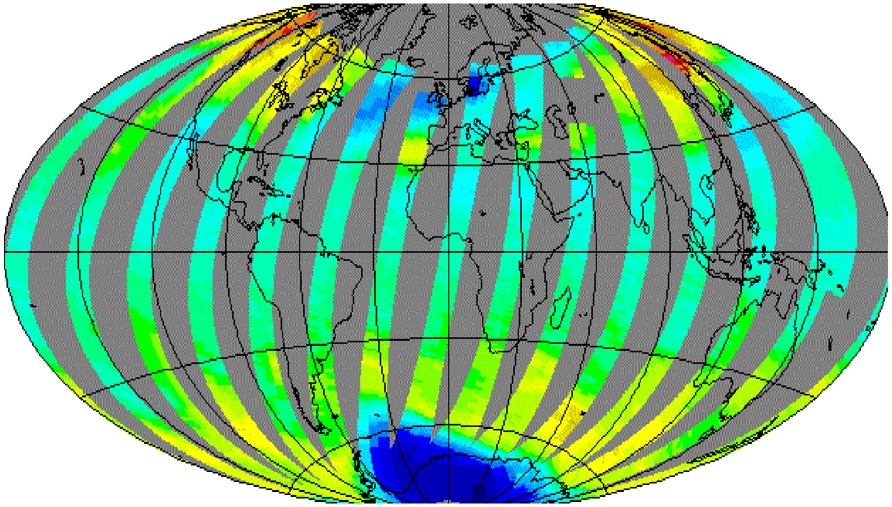
IPSL: M. Petitdidier, S. Godin, C. Boonne, C. Leroy

KNMI: W. Som de Cerff

ESA-ESRIN: L. Fusco, J. Linford

Earth Observation: Ozone

- Building on European Datagrid experience
- To produce and store the Ozone profiles or columns
  - Enhance availability
- To extend the processing capabilities
  - Validation against other data
  - Mid-latitude ozone studies
  - ...
- To facilitate collaboration
  - Including with emerging large scale European projects



GOME instrument
(~75 GB - ~5000 orbits/y)

~28000 profiles/day

Resources added to EGEE

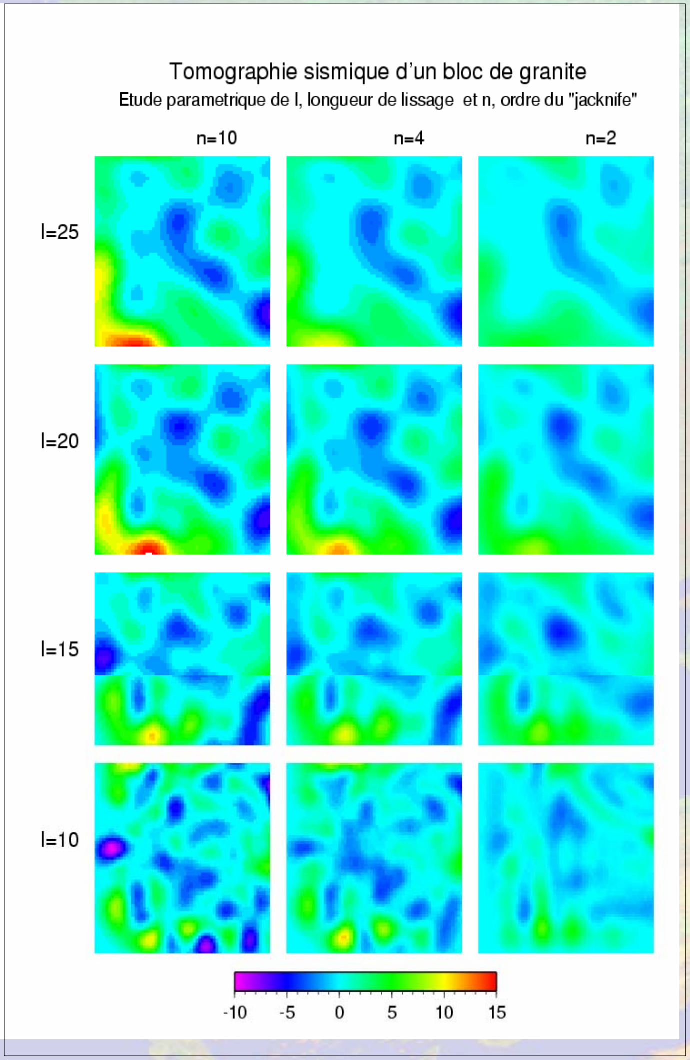
Starting point:

- ESA: UI, CE (15 nodes), SE (1.4 TB)
- IPSL+IPGP at Paris University Computer Center : 4PC, SE (500Gb), UI
- IPGP: UI
- DKRZ: UI, CE (2nodes), SE up to several TB as a function of the application
- KNMI: UI + possibility to use VO NIKHEF and Sara facilities for the Research ES

As new applications are ported new resources will be added

Solid Earth Physics Application

- Objectives : demonstration to drive the community, and production of scientific results
  - GPS data: final goal: workflow with data storage, processing, analysis and visualisation
  - Synthetic seismograms
    - Numerous data and computations, access to databases
  - Earth Core dynamo
- Strategy
  - Demonstrate the secure and restricted access to database
  - Propose tests inside EU project like SPICE
  - Obtain scientific results to constitute databases and propose to the concerned community access via the Grid



Geophysics Applications

Seismic processing Generic Platform:

- Based on Geocluster, an industrial application – to be a starter of the core member VO.

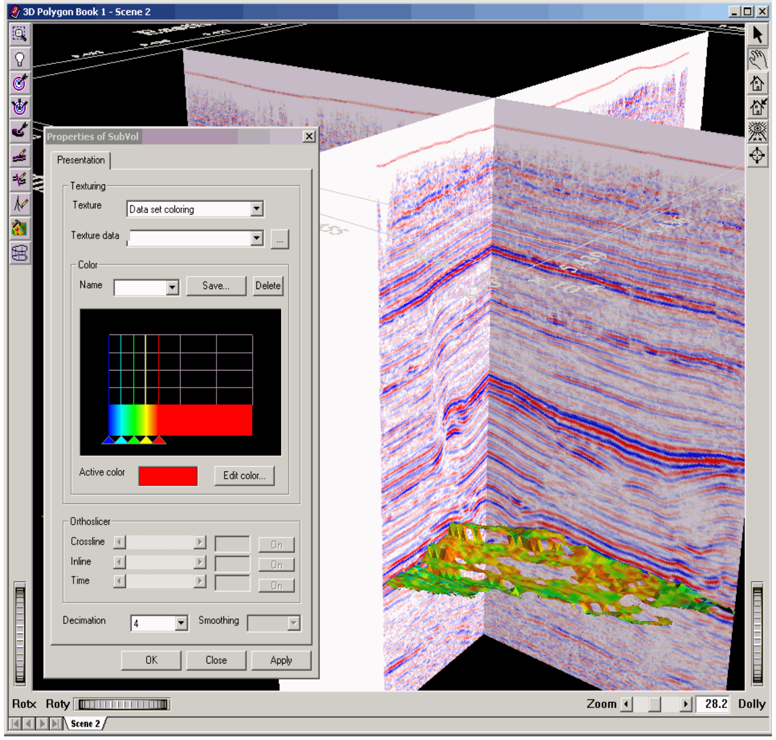
- Include several standard tools for signal processing, simulation and inversion.

- Opened: any user can write new algorithms in new modules (shared or not)

- Free for academic research

-Controlled by license keys (opportunity to explore license issue at a grid level)

- initial partners F, CH, UK, Russia, Norway



Texture: Data set coloring

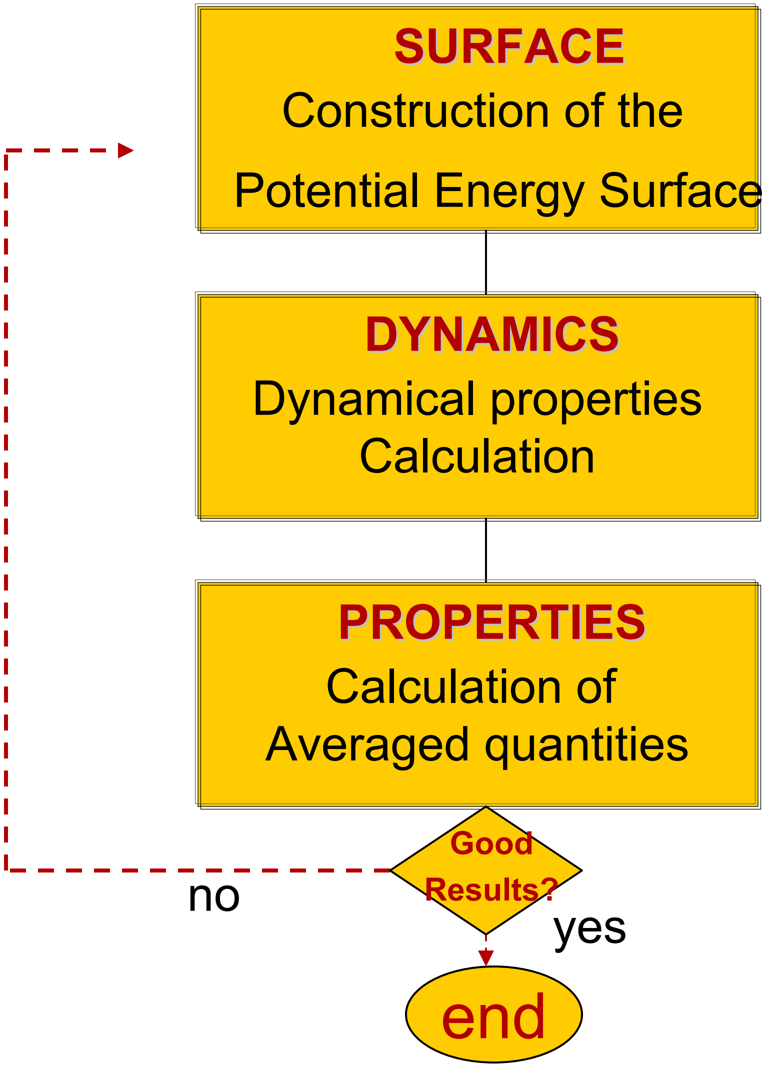
Texture data:

Active color: 🟥

Decimation: 4

28.2

Computational Chemistry: molecular simulator



SURFACE

Construction of the

Potential Energy Surface

DYNAMICS

Dynamical properties Calculation

PROPERTIES

Calculation of Averaged quantities

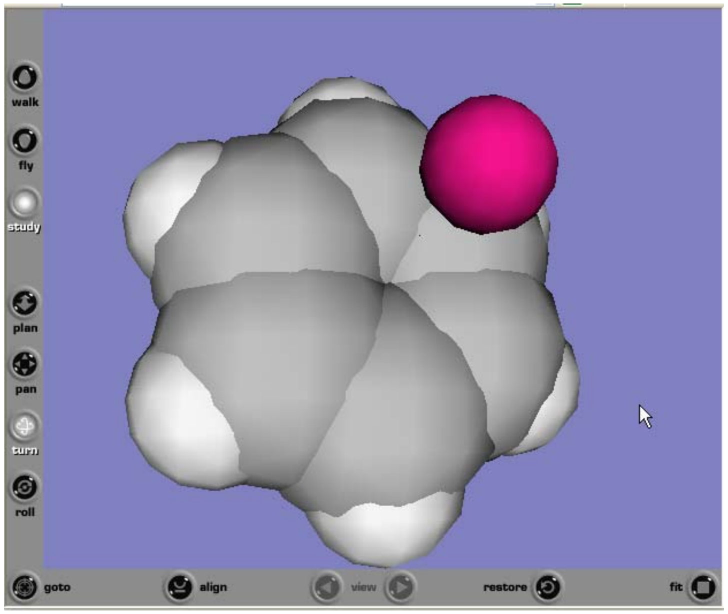
Good Results?

no

yes

end

Ar - Benzene



Critical Features of the Individual Programs

- AB INITIO METHODS (molpro, gamess, adc, gaussian,) resource requests are proportional to N^3 (**N** is the number of electrons) and to M^D (M is the number of grid points per dimension D) for CPU and disc demand.
- EMPIRICAL FORCE FIELDS (Venus, dl_poly, …) resource requests are proportional to P! (P is the number of atoms)

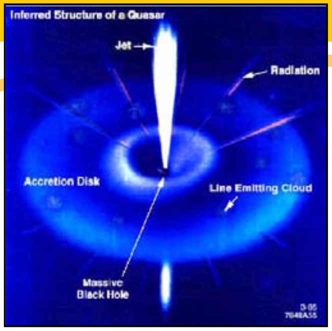
- DYNAMICS (APH3D, TIMEDEP, …) these programs use as input the output of the previous module most critical dependence is on the total angular momentum **J** value that can increase up to several hundred units and the size of the matrices depend on 2J+1
- KINETICS PROGRAMS use dynamics results for integrating relevant time dependent applications

The **MAGIC** telescope

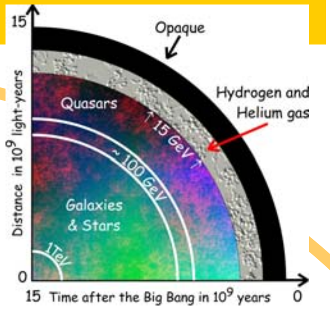
- **Largest Imaging Air Cherenkov Telescope**
  (17 m mirror dish)
- Located on Canary Island
  La Palma (@ 2200 m asl)
- Lowest **energy threshold** ever obtained with a Cherenkov telescope
- Aim: detect **γ–ray sources** in the unexplored energy range:
  30 *(10)*-> 300 GeV

The MAGIC Physics Program

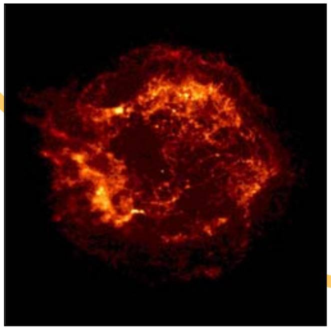
- Pulsars



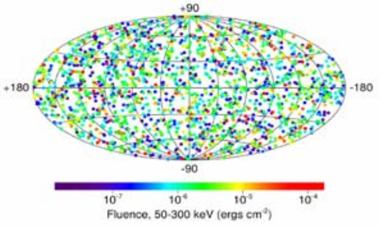
- AGNs



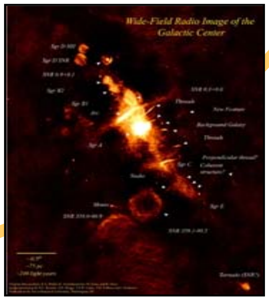
- Cosmological γ-Ray Horizon
- Origin of Cosmic Rays
- Tests of Quantum Gravity effects



- SNRs



- GRBs



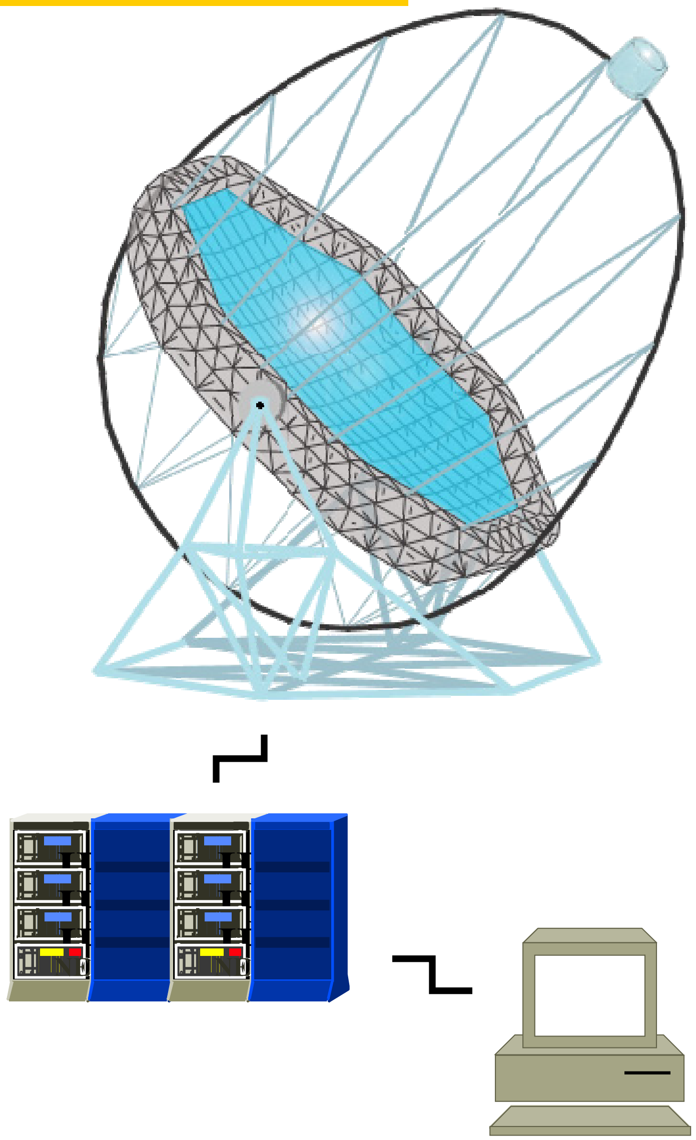
- Cold Dark Matter

Data Acquisition Rate & Storage

- Event Size:
  - 577 PM x 1 Byte x 30 samples
    ⇒ ~ 20 kByte/event

- Data Acquisition Rate:
  - 500 Hz typical trigger rate
    ⇒ ~ 10 MByte/sec

- Data Storage Requirements:
  - ~ 1000 h / year
    useful moonless observation time
    ⇒ ~ **36 TByte/year**



MAGIC Summary

Enabling Grids for E-science in Europe

MAGIC:

- is a new generation gamma ray Cherenkov telescope
- has large discovery potential both in astrophysics and fundamental physics
- just started data taking
- has large computing requirements
  - > 100 CPU
  - > 50 TB / year
- is well suited to join and test GRID technology with
  16 participating institutions over all Europe (and beyond)
  some with strong links to mayor GRID sites (Bologna, Barcelona)

Contents

- EGEE
  - *The vision - why EGEE got started*
  - *Where are we now?*
  - *Who is “we”?*
  - *Where are we going? (project goals)*
  - *How will we get there? (project structure and activities)*
- Current virtual organisations (i.e. application communities)
- **Some other important questions about EGEE**

Who else can benefit from EGEE?

- EGEE Generic Applications Advisory Panel:
  - For new applications

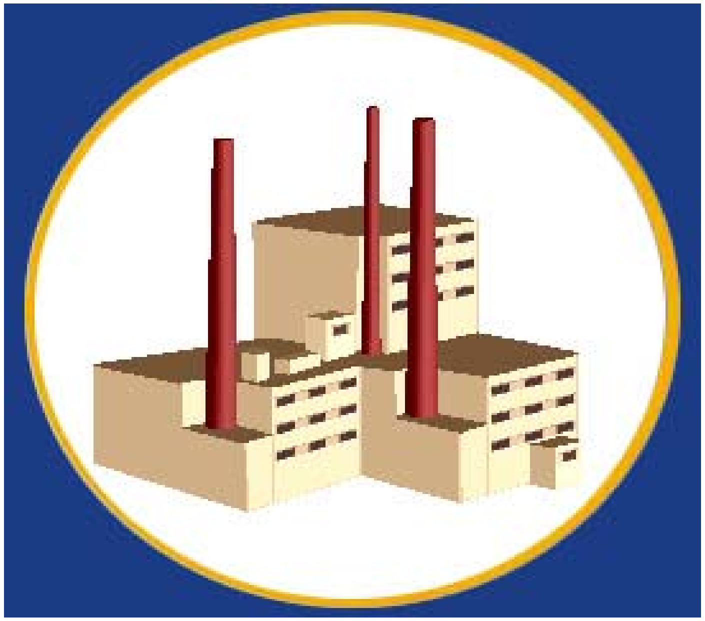
- EU projects: MammoGrid, Diligent, SEE-GRID …

- Expression of interest: Planck/Gaia (astroparticle), SimDat (drug discovery)

http://agenda.cern.ch/age?a042351

Links to industry?

- EGEE Industry Forum
  - raise awareness of the project in industry to encourage industrial participation in the project
  - foster direct contact of the project partners with industry
  - ensure that the project can benefit from practical experience of industrial applications



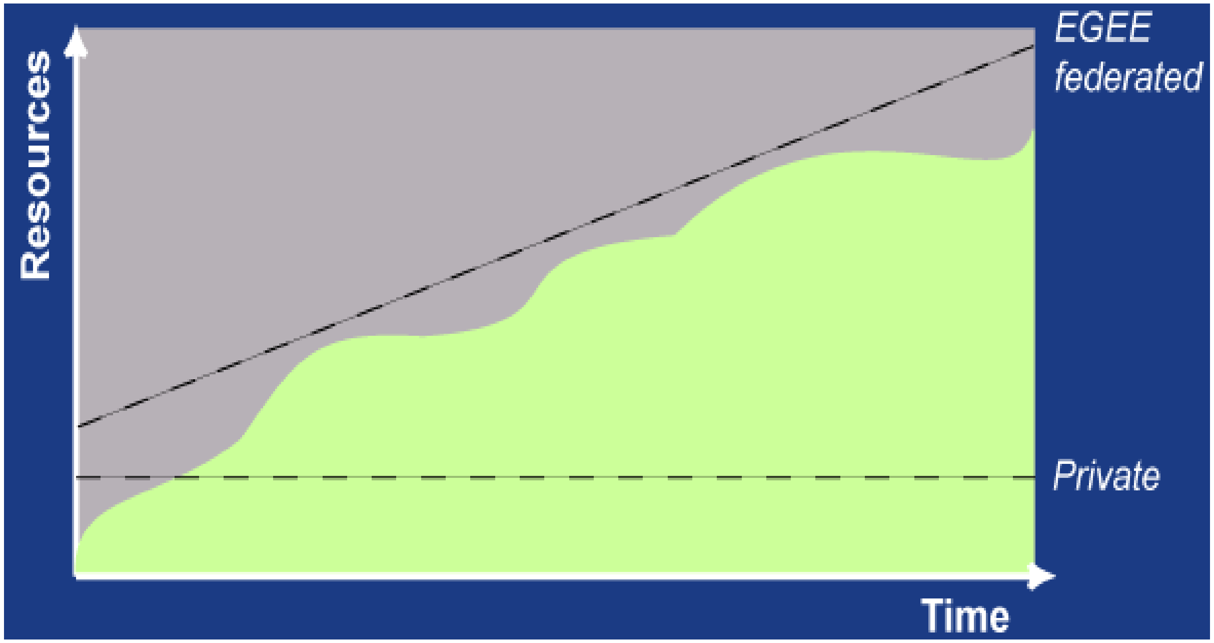
- For more info:

  http://public.eu-egee.org/industry/

Private or Federated Resources?

For applications that must operate in a closed environment, EGEE middleware can be downloaded and installed on closed infrastructures

- Approach being used by MammoGrid



EGEE sites are administered/owned by different organisations

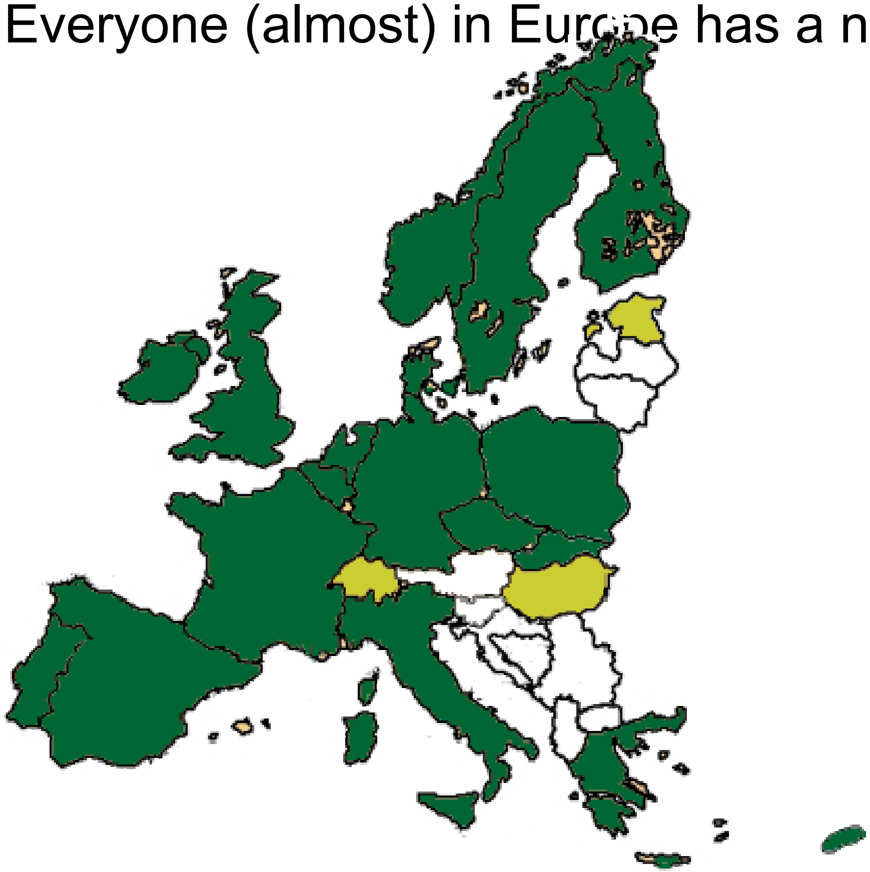
- Sites have ultimate control over how their resources are used
- Limiting the demands of your application will make it acceptable to more sites and hence make more resources available to you

Intellectual Property

- The existing EGEE grid middleware (LCG-2) is distributed under an Open Source License developed by EU DataGrid
  - Derived from modified BSD - no restriction on usage (academic or commercial) beyond acknowledgement
  - Same approach for new middleware (gLite)
- Application software maintains its own licensing scheme
  - Sites must obtain appropriate licenses before installation

How to access EGEE (III)

- Where to go for an accredited certificate?
- Everyone (almost) in Europe has a national CA



- Green: CA Accredited
- Yellow: being discussed

Other Accredited CAs:

- DoEGrids (US)
- GridCanada
- ASCCG (Taiwan)
- ArmeSFO (Armenia)
- CERN
- Russia (*HEP*)
- FNAL Service CA (US)
- Israel
- Pakistan

EGEE Plans for the coming year

Enabling Grids for E-science in Europe

42 deliverables in 1st year

- **November**

  2nd EGEE conference (Den Hague) in common with DEISA, SEE-GRID, DILIGENT etc.

- **December**

  Application migration reports

- **February 2005**

  1st EU review

- **March 2005**

  Large-scale deployment of gLite software

  Annual report

To read more about EGEE…

Enabling Grids for E-science in Europe

- Explore the web site!
- www.eu-egee.org

Summary

- **EGEE** is the first attempt to build a worldwide Grid infrastructure for data intensive applications from **many scientific domains**
- A **large-scale production grid service** is already deployed and being used for HEP and BioMed applications with new applications being ported
- Resources & user groups will **rapidly expand** during the project
- A process is in place for **migrating new applications** to the EGEE infrastructure
- A **training programme** has started with events already held
- Prototype “*next generation*” middleware is being tested (**gLite**)
- Plans for a **follow-on project** are being discussed

Further Information

EGEE www.eu-egee.org
LCG lcg.web.cern.ch/LCG/

NeSC www.nesc.ac.uk

The Grid Cafe www.gridcafe.org