The EGEE project and the gLite middleware

Marc-Elian Bégin
EGEE Technical Coordination and Dissemination
CERN IT

CERN Summer Students Grid Tutorial, CERN, 24 August 2005

www.eu-egee.org

Information Society

Overview

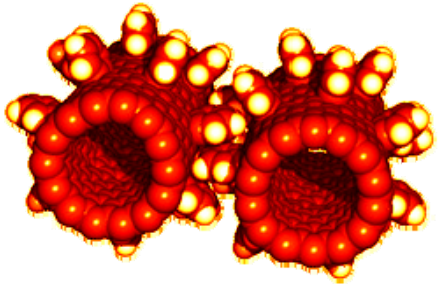
- EGEE - what is it and why is it needed?
- Middleware – current and future
- Operations
- Applications running in Production
- Summary

Overview

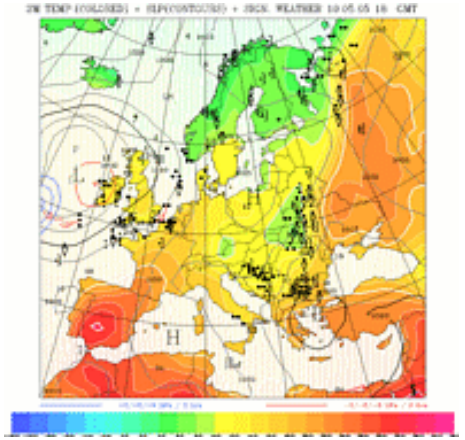
- EGEE - what is it and why is it needed?
- Middleware – current and future
- Operations
- Applications running in Production
- Summary

Computing intensive science

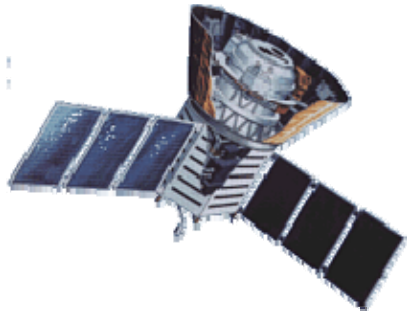
- **Science is becoming increasingly digital and needs to deal with increasing amounts of data**



- **Simulations get ever more detailed**
  - Nanotechnology – design of new materials from the molecular scale
  - Modelling and predicting complex systems (weather forecasting, river floods, earthquake)
  - Decoding the human genome



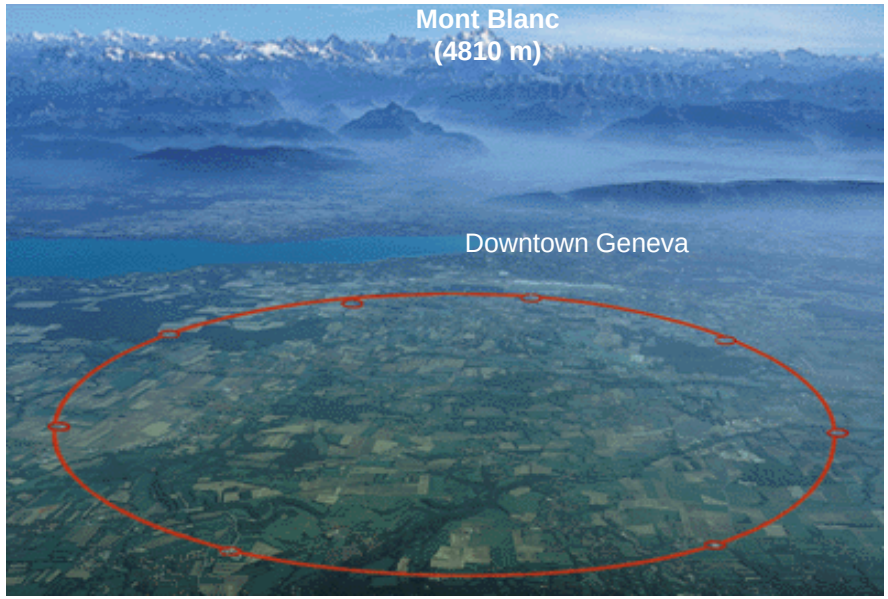
- **Experimental Science uses ever more sophisticated sensors to make precise measurements**
  - → Need high statistics
  - → Huge amounts of data
  - → Serves user communities around the world



Particle Physics

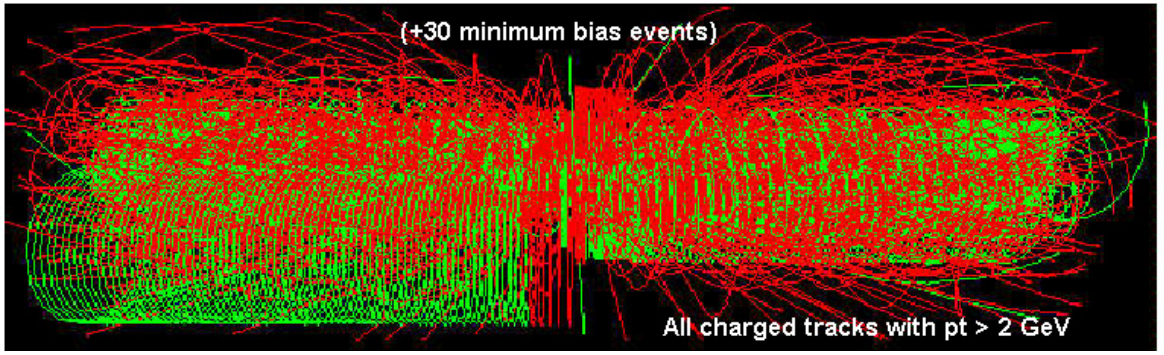
- **CERN: the world's largest particle physics laboratory**
- **Particle physics requires special tools to create and study new particles: accelerators and detectors**

- **Large Hadron Collider (LHC):**
  - One of the most powerful instruments ever built to investigate matter
  - four experiments: ALICE, ATLAS, CMS, LHCb
  - 27 km circumference tunnel
  - due to start up in 2007

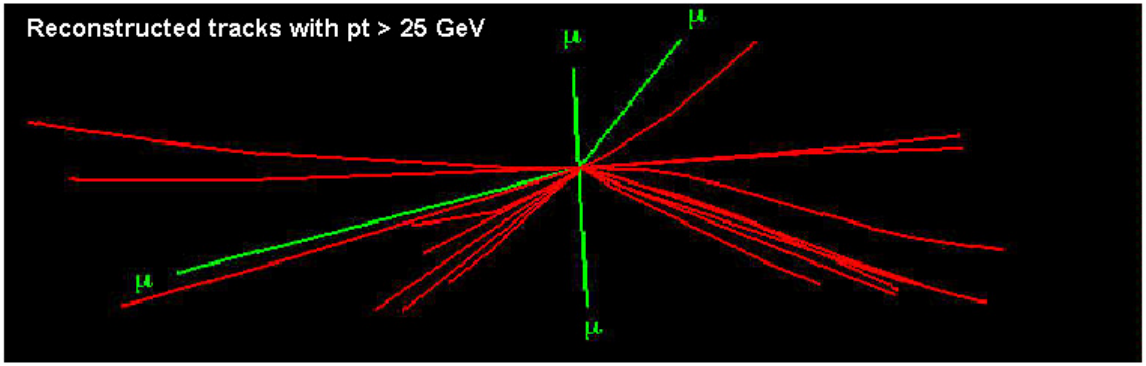


The LHC Data Challenge

Starting from this event



Looking for this "signature"



→ **Selectivity: 1 in 10^{13}**

(Like looking for a needle in 20 million haystacks)

Data handling

detectors

Filter

raw data

event simulation

Reprocessing

Analysis

interactive physics analysis

event summary data

processed data

The solution: the Grid

- **Integrating computing power and data storage capacities at major computer centres**

- **Providing users with seamless access to computing resources, 24/7, independent of geographic location**

→More effective and seamless collaboration of dispersed communities, both scientific and commercial

→Ability to run large-scale applications comprising thousands of computers, for wide range of applications

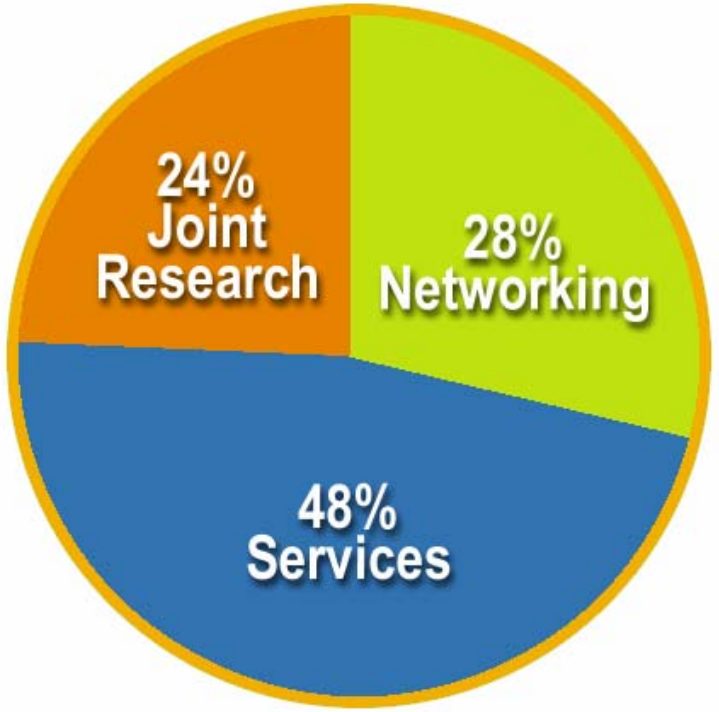
→The term “e-Science” has been coined to express these benefits

EGEE Goals

- **Goal of EGEE: develop a service grid infrastructure which is available to scientists 24 hours-a-day**
- **The project concentrates on:**
  - building a consistent, robust and secure Grid infrastructure that will attract additional computing resources and applications
  - continuously improve and maintain the middleware in order to deliver a reliable service to users
  - attracting new users from industry as well as science and ensure they receive the high standard of training and support they need
- **EGEE is about multi-science on the Grid !!!**

EGEE Activities

- 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

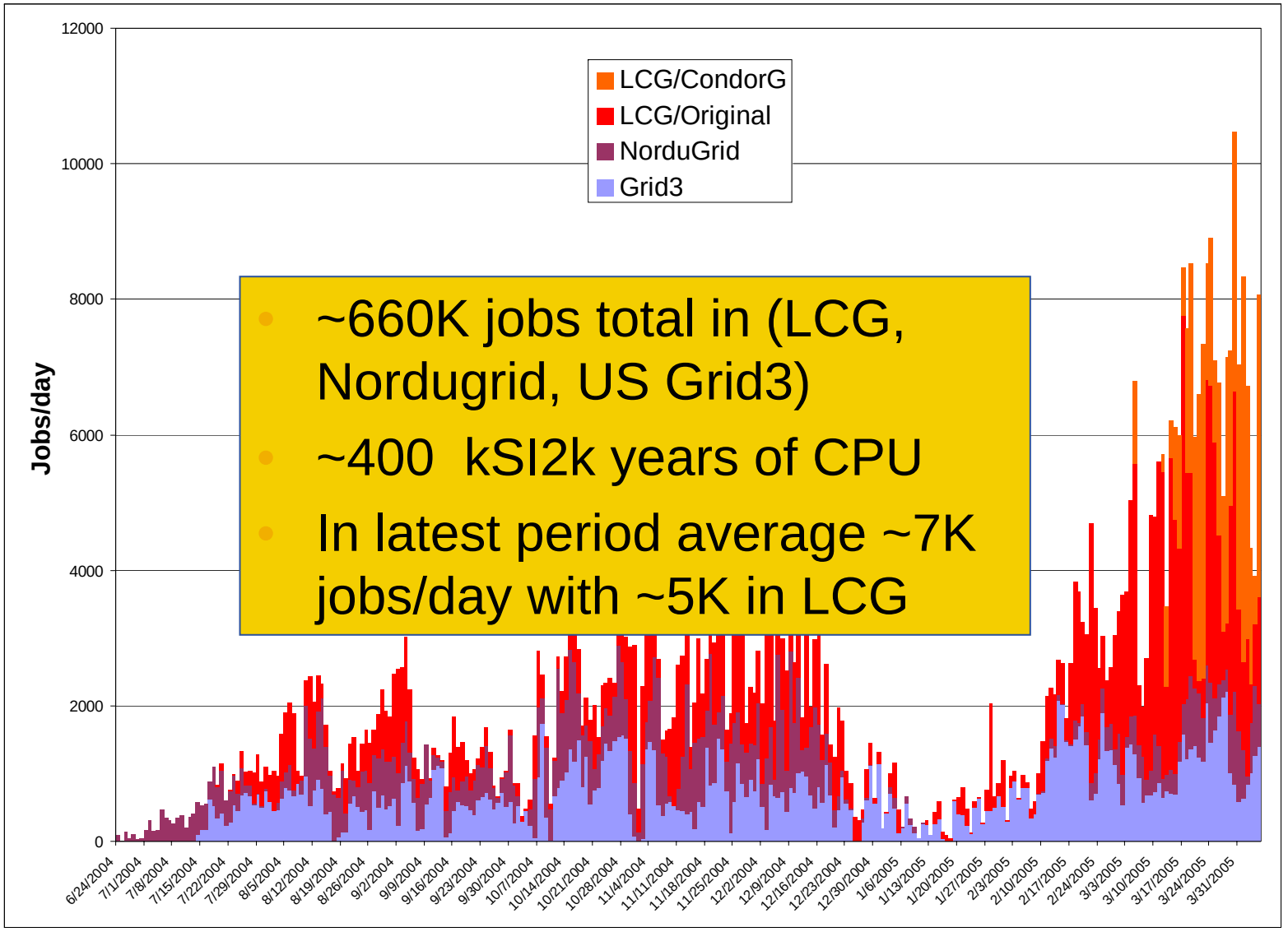
EGEE Infrastructure

Enabling Grids for E-sciencE

(HEP) Utilisation

12000

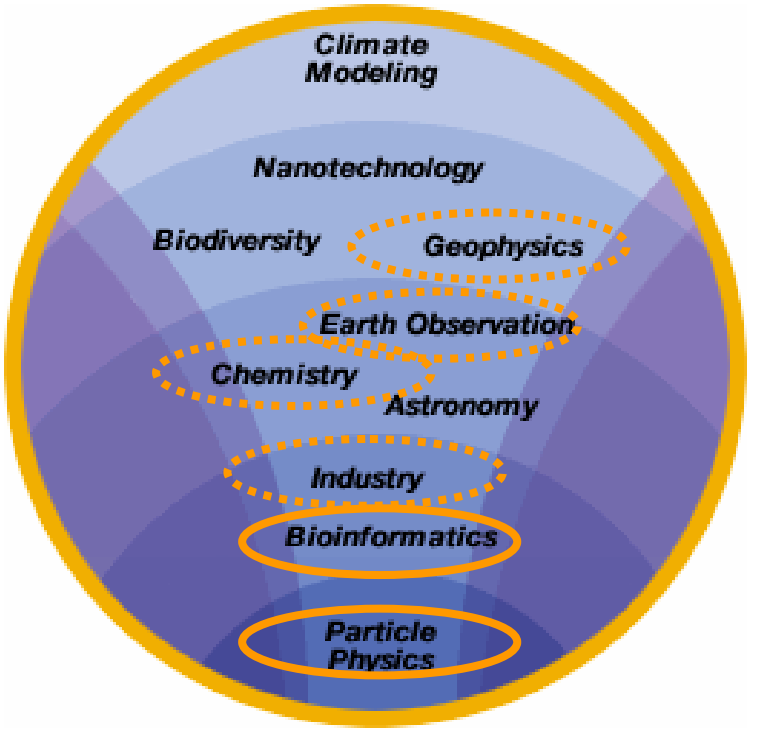
6000



- ~660K jobs total in (LCG, Nordugrid, US Grid3)
- ~400 kSI2k years of CPU
- In latest period average ~7K jobs/day with ~5K in LCG

Deployment of applications

- **Pilot applications**
  - High Energy Physics
  - Biomed applications
- **Generic applications – Deployment under way**
  - Computational Chemistry
  - Earth science research
  - EGEODE: first industrial application (Geophysics)
  - Astrophysics
- **With interest from**
  - Hydrology
  - Seismology
  - Stock market simulators
  - Digital video etc.
  - Industry (provider, user, supplier)



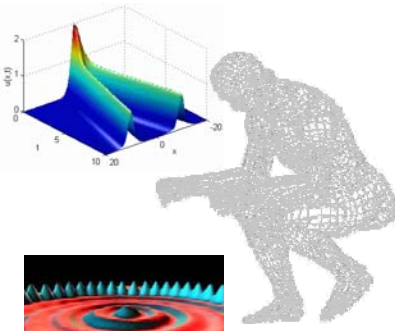
Why others have joined?

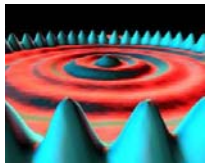
- **Transparent access to millions of files across different administration domains**
- **Low cost access to large computing resources**
  - Mobilise quickly very large amount of CPU on very prompt basis
- **Produce and store massive amount of data**
- **Develop applications using distributed complex workflows**
- **Eases distributed scientific collaborations**

- **… All this at a reduced cost!!**

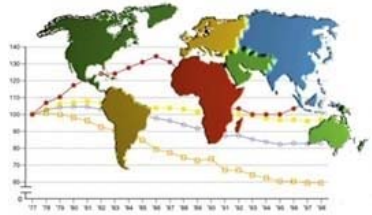
From Phase I to II

- **From 1st EGEE EU Review in February 2005:**
  - "The reviewers found the overall performance of the project very good."
  - "… remarkable achievement to set up this consortium, to realize appropriate structures to provide the necessary leadership, and to cope with changing requirements."

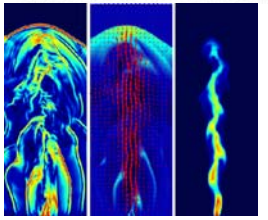


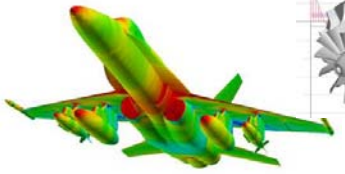


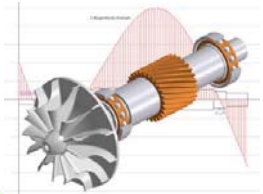
- **EGEE I**
  - Large scale deployment of EGEE infrastructure to deliver production level Grid services with selected number of applications



- **EGEE II**
  - Natural continuation of the project's first phase
  - Opening up to a larger user community
  - → increased multidisciplinary Grid infrastructure
  - → more involvement from Industry
  - Extending the Grid infrastructure world-wide
  - → increased international collaboration





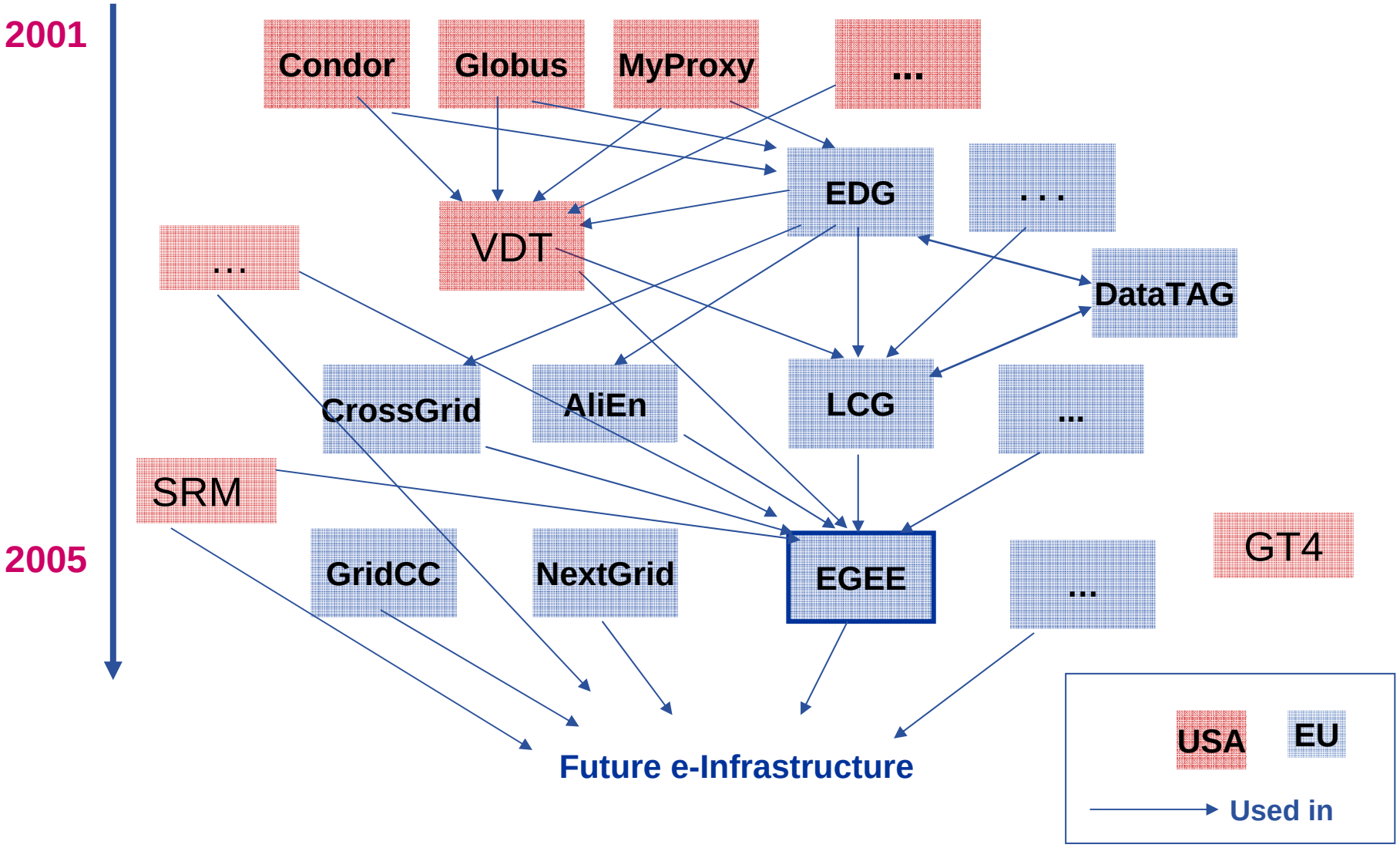


Overview

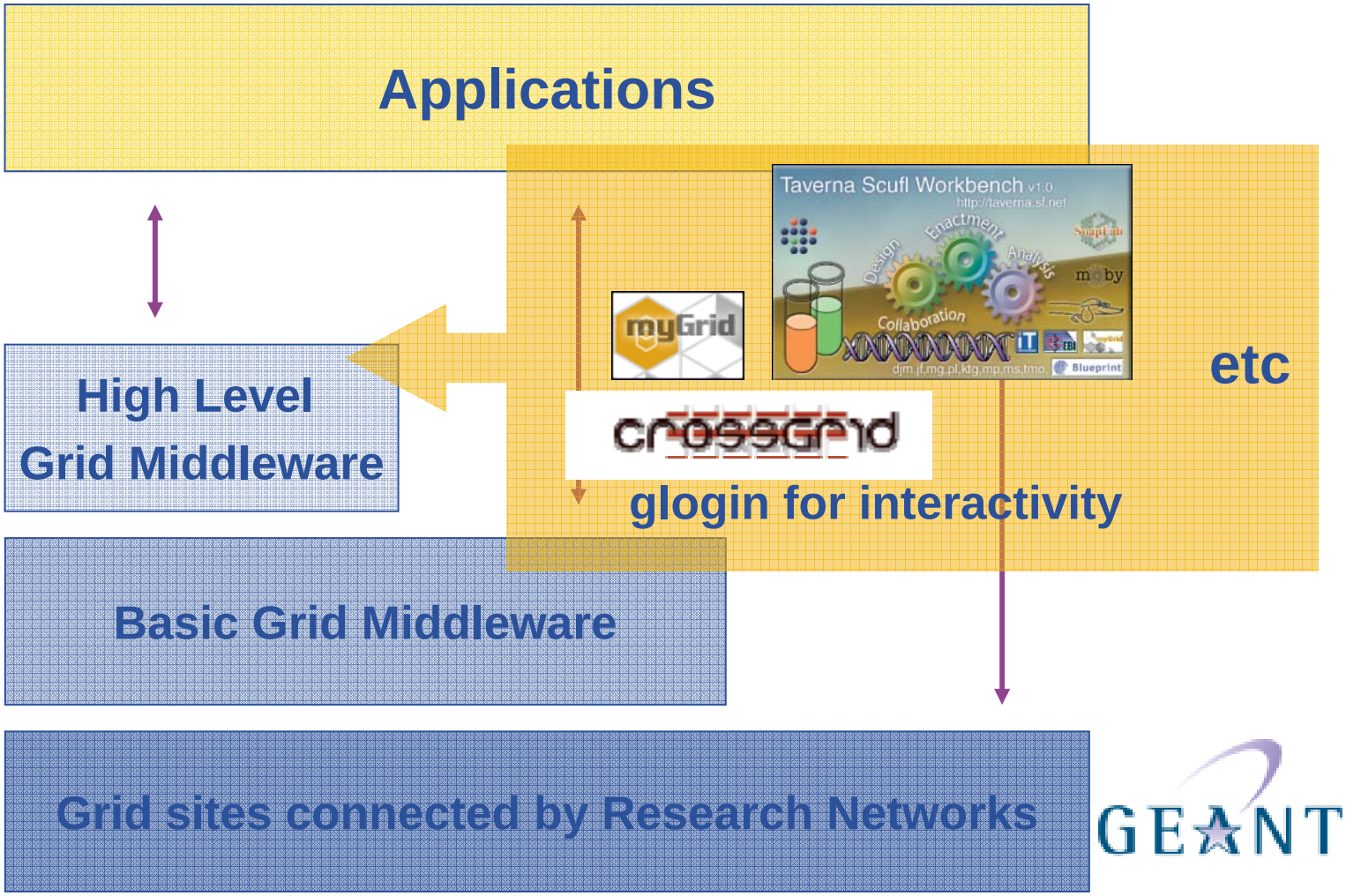
- EGEE - what is it and why is it needed?
- Middleware – current and future
- Operations
- Applications running in Production
- Summary

Where does grid middleware come from?

EGEE view of history



The Grid Stack: The Full Picture



gLite the next generation of middleware

New EGEE Middleware - gLite

- Intended to replace LCG-2
- Started with existing components from AliEN, EDG, VDT, etc.
- Aims to address LCG-2 shortcoming and advanced needs from applications
- Prototyping short development cycles for fast user feedback
- Secure Grid Services, with consistent security infrastructure (critical for many new users)

gLite
Lightweight Middleware for Grid Computing

New EGEE Middleware – gLite (2)

- Lightweight services
- Interoperability & Co-existence with deployed infrastructures
- Performance & Fault Tolerance
- Portable
- Service oriented architecture
- Site autonomy
- Open source license
- 1st release of gLite (v1.0) made end March'05
- Now testing gLite v1.3 in pre-production

New EGEE Middleware – gLite (3)

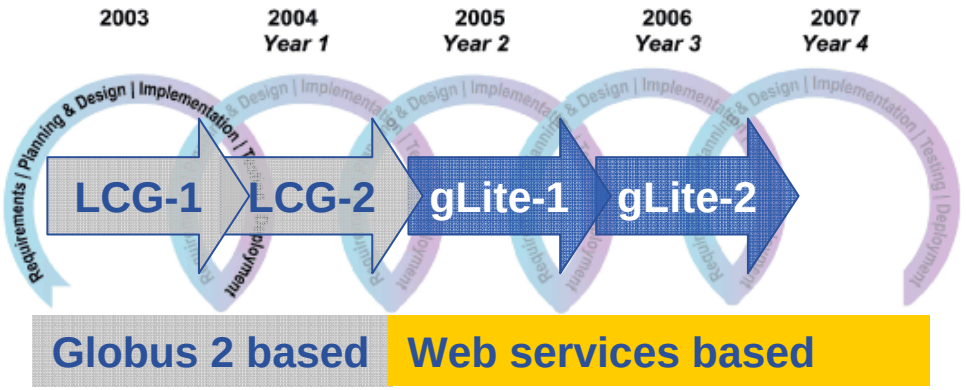
- **Service Oriented Architecture - components that are :**
  - "Loosely Coupled" (by messages and only via interfaces), but "Highly Cohesive" (each method needed and consistent)
  - Accessible across network; modular and self-contained; clean modes of failure; consistent logging
  - Implementations can be interchanged (contract only depending on public interfaces)
- **… and based on standards. Opens EGEE to:**
  - New middleware (plethora of tools now available)
  - Heterogeneous resources (storage, computation…)
  - Interact with other Grids (international, regional and national)
  - Looking at: WSRF and OGSA
  - Releasing services with plain Web-Services interfaces (WS-I compliant)

New EGEE Middleware – gLite (4)

- **Lightweight (existing) services**
  - Easily and quickly deployable
  - Use existing services where possible as basis for re-engineering
- **Resilience and Fault Tolerance**
  - Parts of the Grid can fail without bringing the entire infrastructure down
  - Services able to swap service end-points to continue to work
- **Co-existence with deployed infrastructure**
  - Reduce requirements on site components
  - Co-existence (and convergence) with LCG-2 and Grid3

gLite middleware Release Process

- **Development**
  - Workload Management
  - Information Systems
  - Security
  - Data Management
- **Integration**
- **Testing**

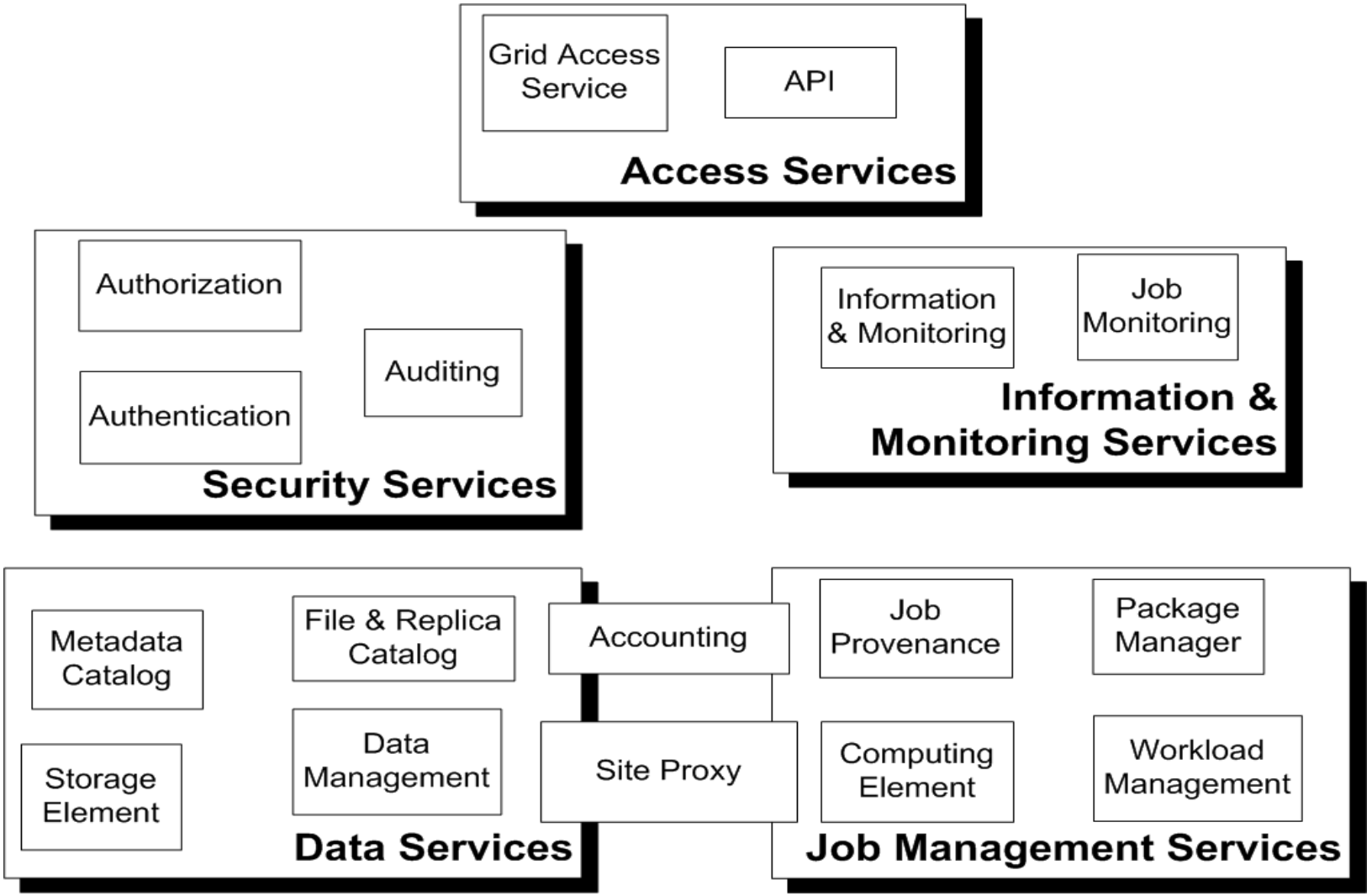


- **gLite offers a complete data management solution in a distributed environment building on existing technology**

gLite the next generation of middleware

(*a few more details*)

gLite Services



gLite Services in Release 1
Software stack and origin (simplified)

- **Computing Element**
  - Gatekeeper *(Globus)*
  - Condor-C *(Condor)*
  - CE Monitor *(EGEE)*
  - Local batch system *(PBS, LSF, Condor)*
- **Workload Management**
  - WMS *(EDG)*
  - Logging and bookkeeping *(EDG)*
  - Condor-C *(Condor)*
- **Information and Monitoring**
  - R-GMA *(EDG)*
- **Storage Element**
  - glite-I/O *(AliEn)*
  - File Transfer Service *(EGEE)*
  - GridFTP *(Globus)*
  - SRM: Castor *(CERN)*, dCache *(FNAL, DESY)*, other SRMs
- **Catalog**
  - File/Replica & Metadata Catalogs *(EGEE)*
- **Security**
  - GSI *(Globus)*
  - VOMS *(DataTAG/EDG)*
  - Authentication for C and Java based (web) services *(EDG)*

Now doing rigorous scalability and performance tests on pre-production service

Workload management

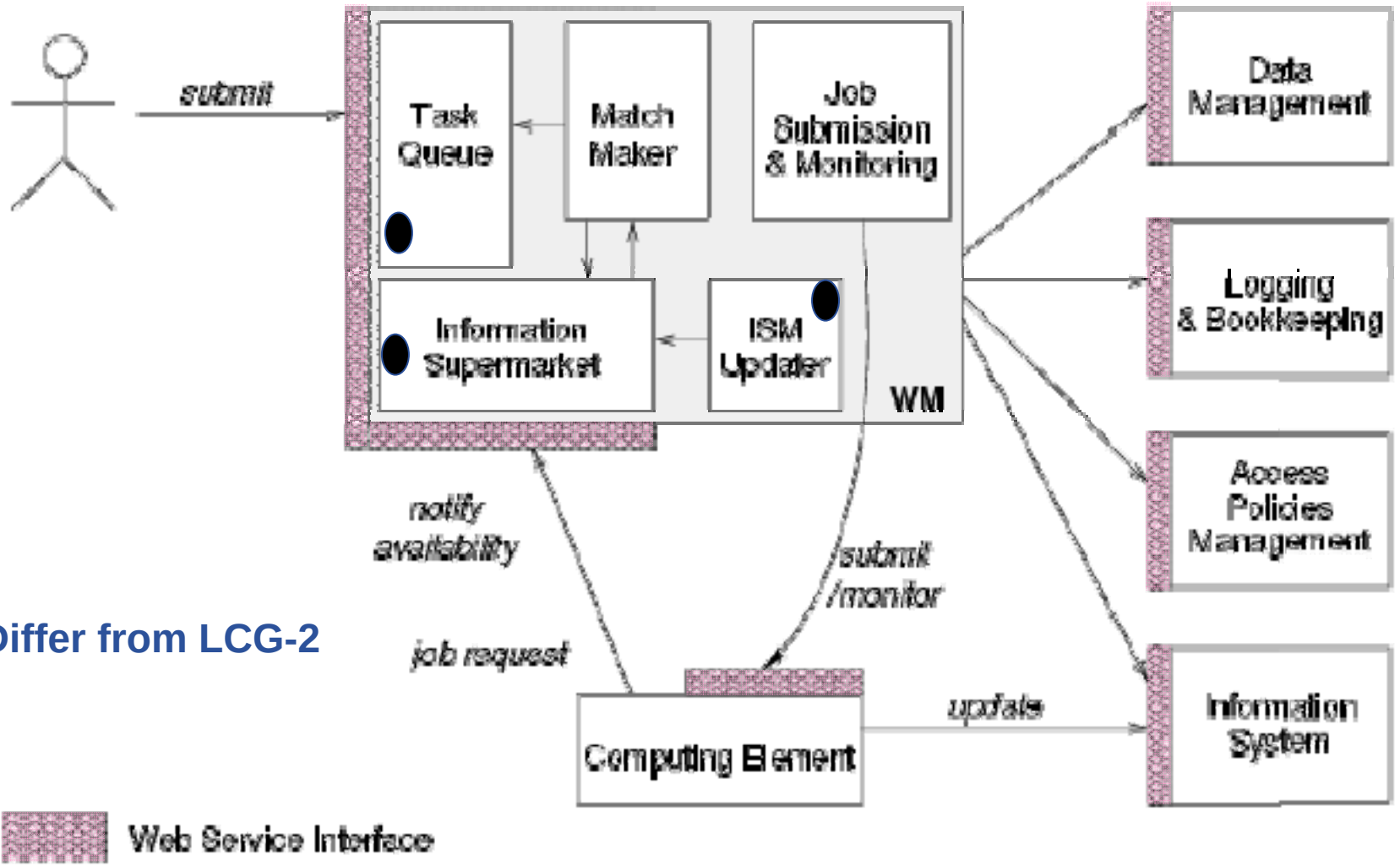
(getting a job to run)

Workload Management System

- **Workload Management System (WMS) comprises a set of Grid middleware components responsible for distribution and management of tasks across Grid resources**
  - applications are conveniently, efficiently and effectively executed.

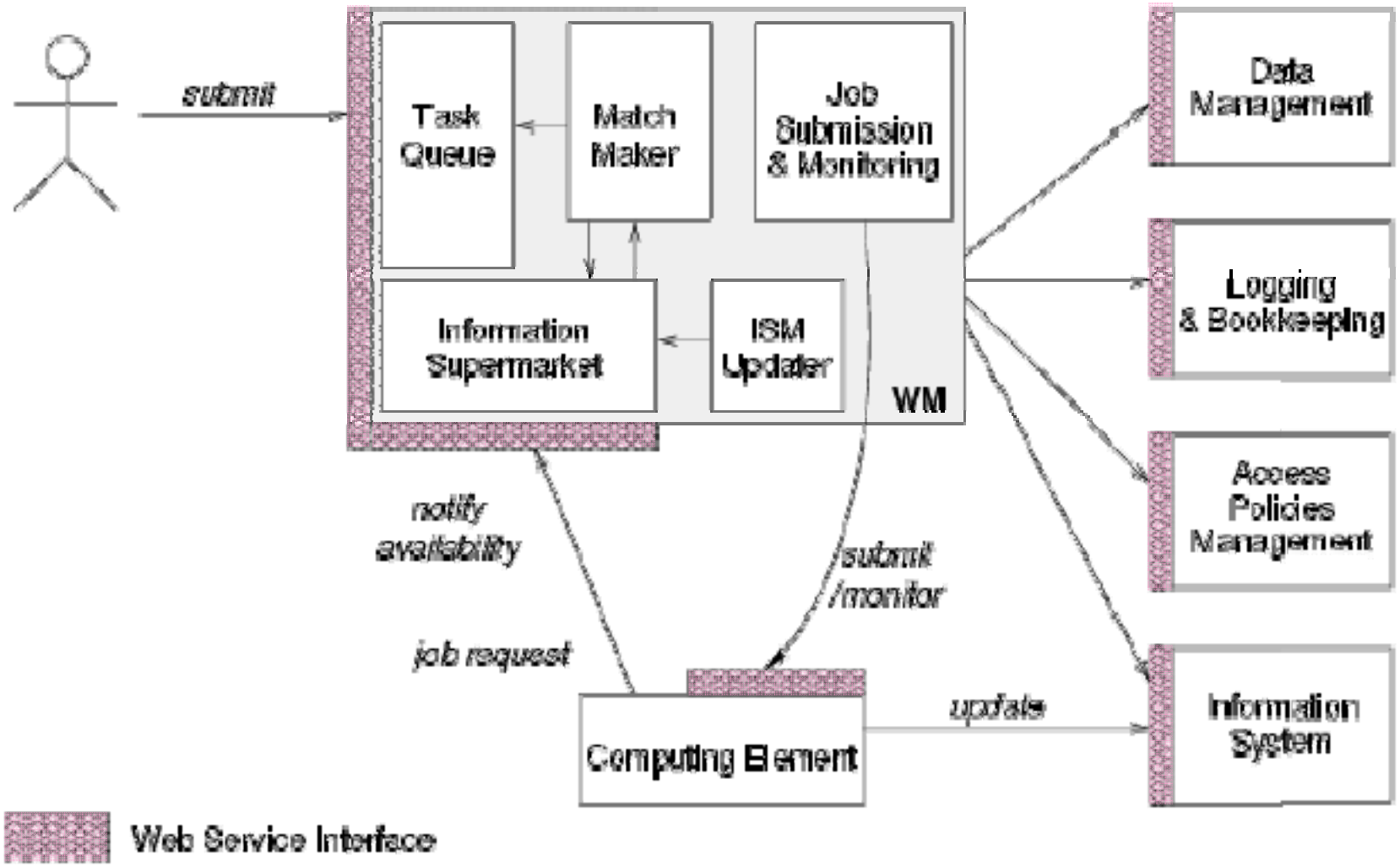
- **Comparable services from other grid projects are, among others, the EDG WMS, Condor and the Eurogrid-Unicore resource broker.**

Workload Management System (2)



● - Differ from LCG-2

Workload Management System (2)

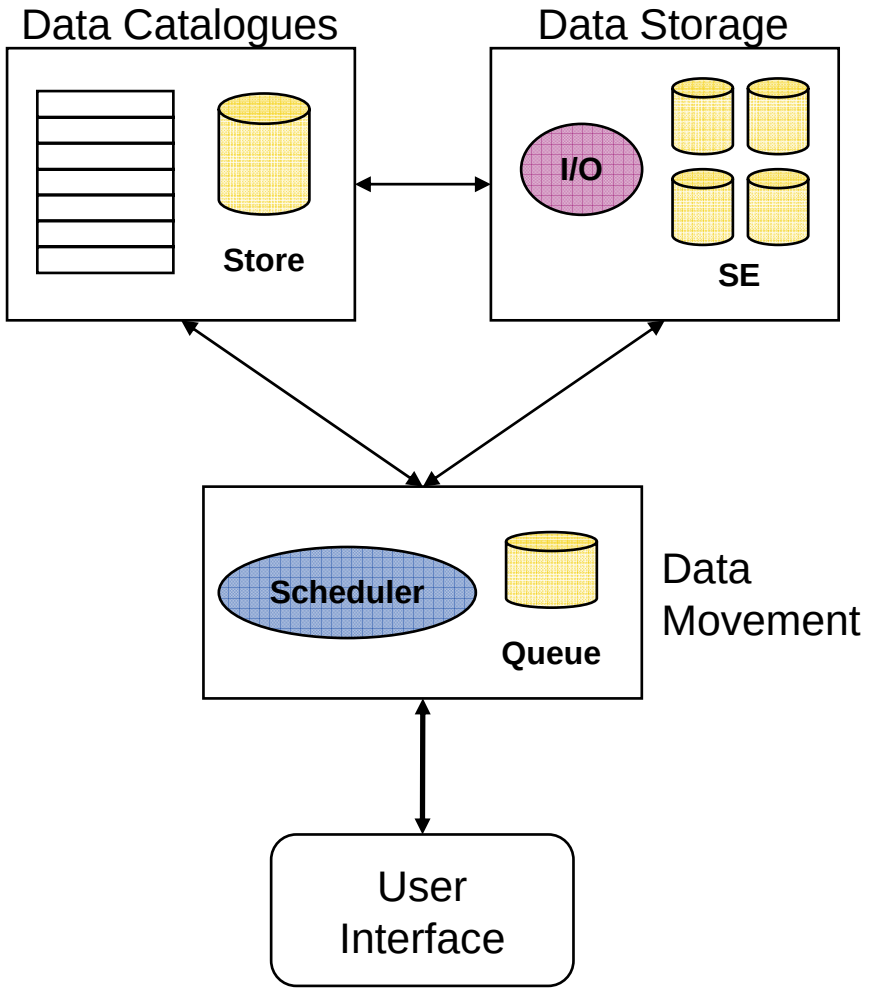


File management

(managing your data)

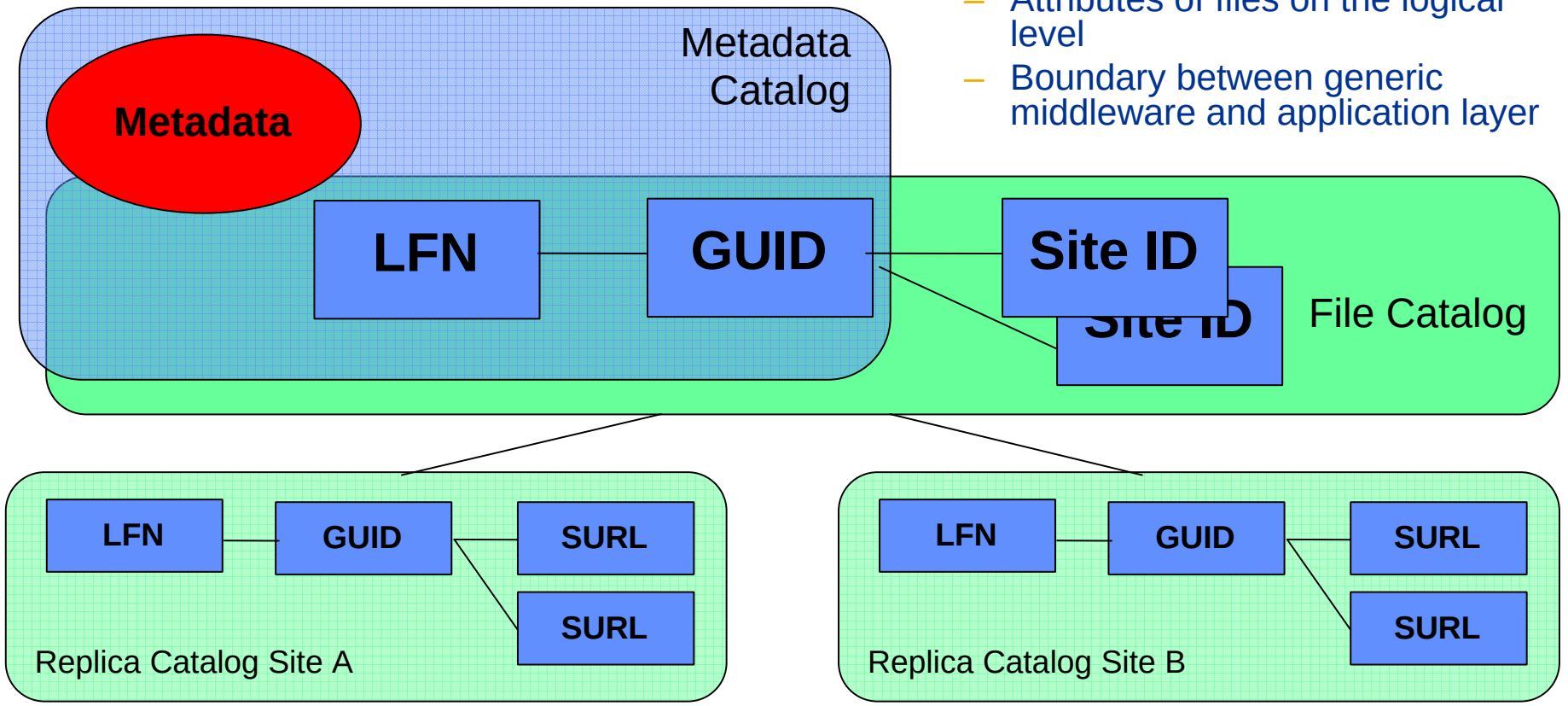
Data Management Services

- **Storage Element**
  - Storage Resource Manager
  - POSIX-I/O
  - Access protocols
- **Catalogues**
  - File Catalogue
  - Replica Catalogue
  - File Authorization Service
  - Metadata Catalogue
- **File Transfer**
  - Data Scheduler
    (not implemented yet)
  - File Transfer Service
  - File Placement Services
- **User Interface**



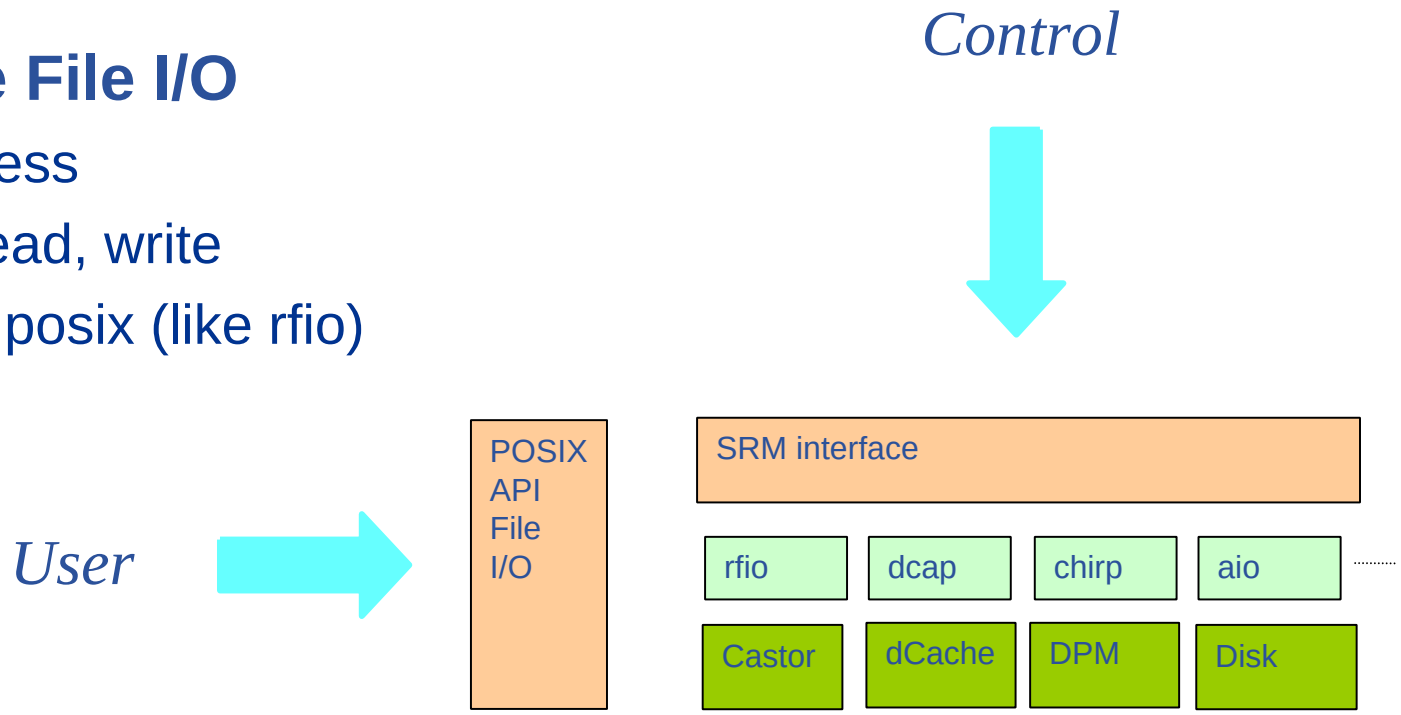
gLite Catalogs

- **File Catalog**
  - Filesystem-like view on logical file names
  - Keeps track of sites where data is stored
  - Conflict resolution
- **Replica Catalog**
  - Keeps information at a site
- **(Meta Data Catalog)**
  - Attributes of files on the logical level
  - Boundary between generic middleware and application layer



Storage Element Interfaces

- **SRM interface**
  - Management and control
  - SRM (with possible evolution)
- **Posix-like File I/O**
  - File Access
  - Open, read, write
  - Not real posix (like rfio)



File Transfer/Placement Service

- **File Transfer Service**
  - Reliable and secure file transfer service
  - From SURL to SURL
- **File Placement Service**
  - Ability to specify requirement for data (LFN or GUID)
  - Automatic trigger of transfer, before the job lands on the WN

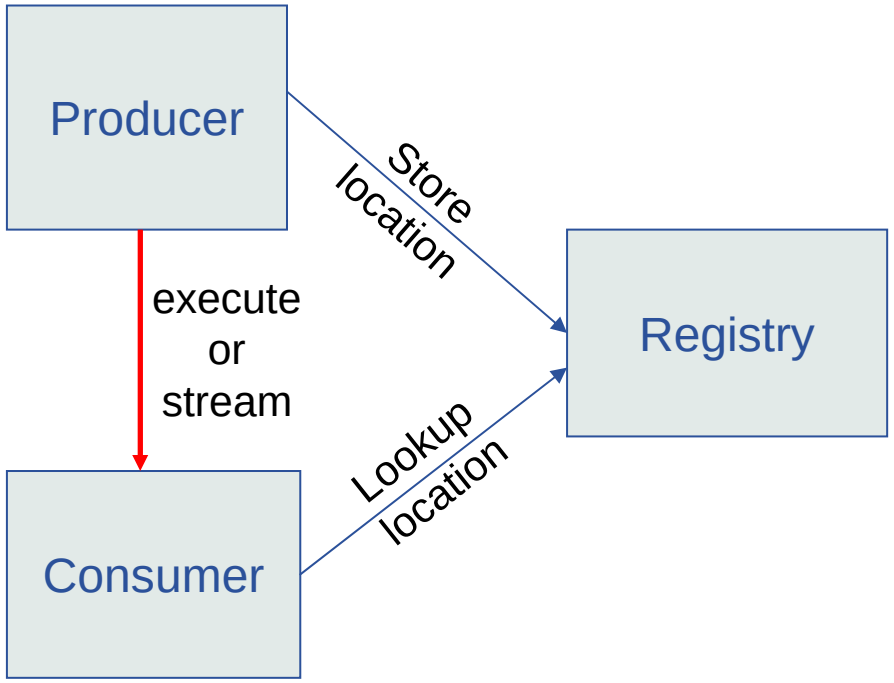
Information System

(finding and using resources)

gLite Information System – R-GMA

- **R-GMA is based on the Grid Monitoring Architecture (GMA) from the Grid Global Forum (GGF)**
- **Simple Consumer-Producer model that models the information infrastructure of a Grid as**
  - consumers (that request information)
  - producers (that provide information) and,
  - a central registry
- **Mediates the communication between producers and consumers**
- **Provide replication of data**
- **Offers a global view of the information as if each Virtual Organization had one large relational database**
- **Producers contact the registry to announce their intention to publish data**
- **R-GMA adds a standard query language (a subset of SQL) to the GMA model,**

gLite Information System – R-GMA



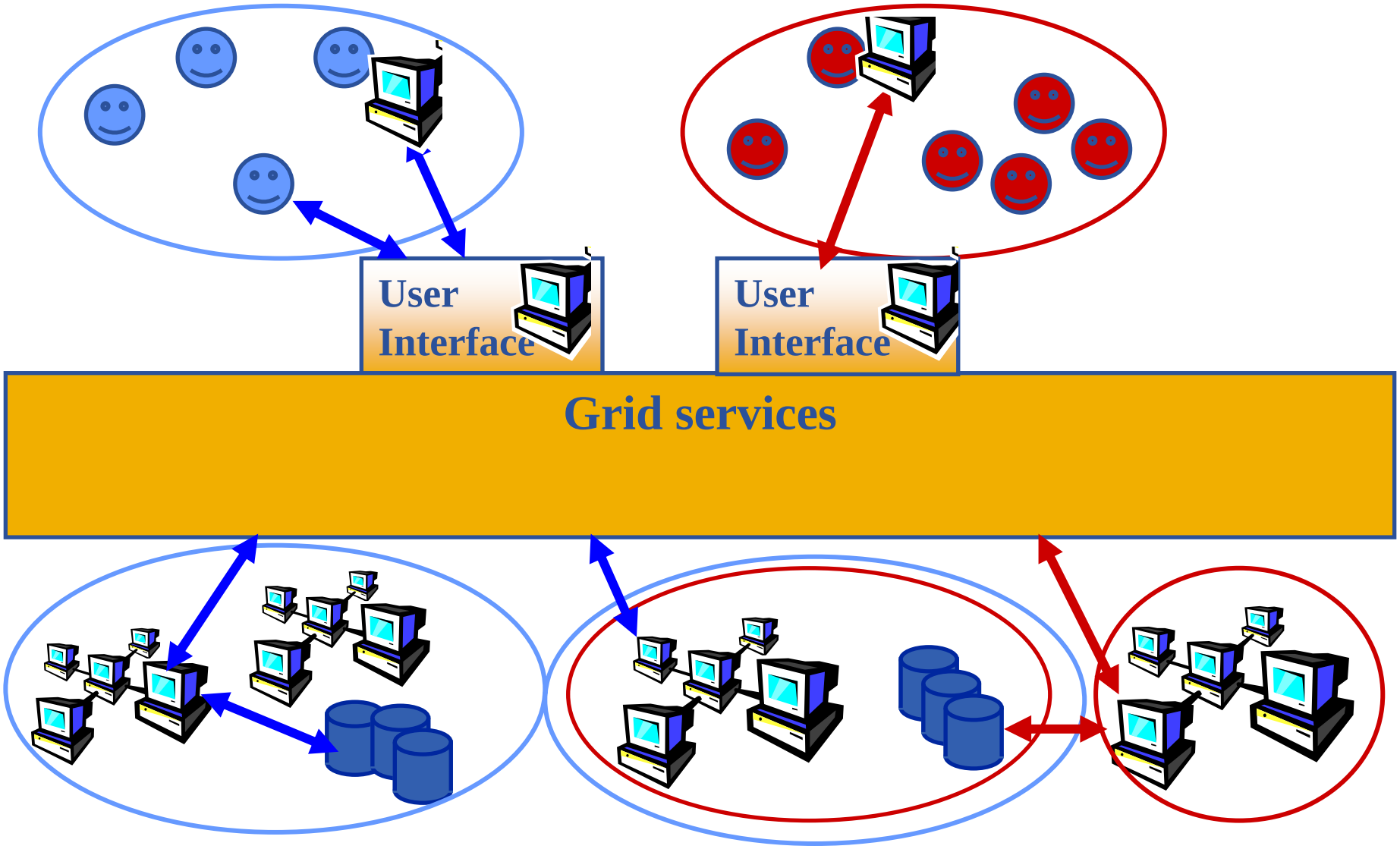
Security

(keep the Grid a safe place)

Virtual Organisation

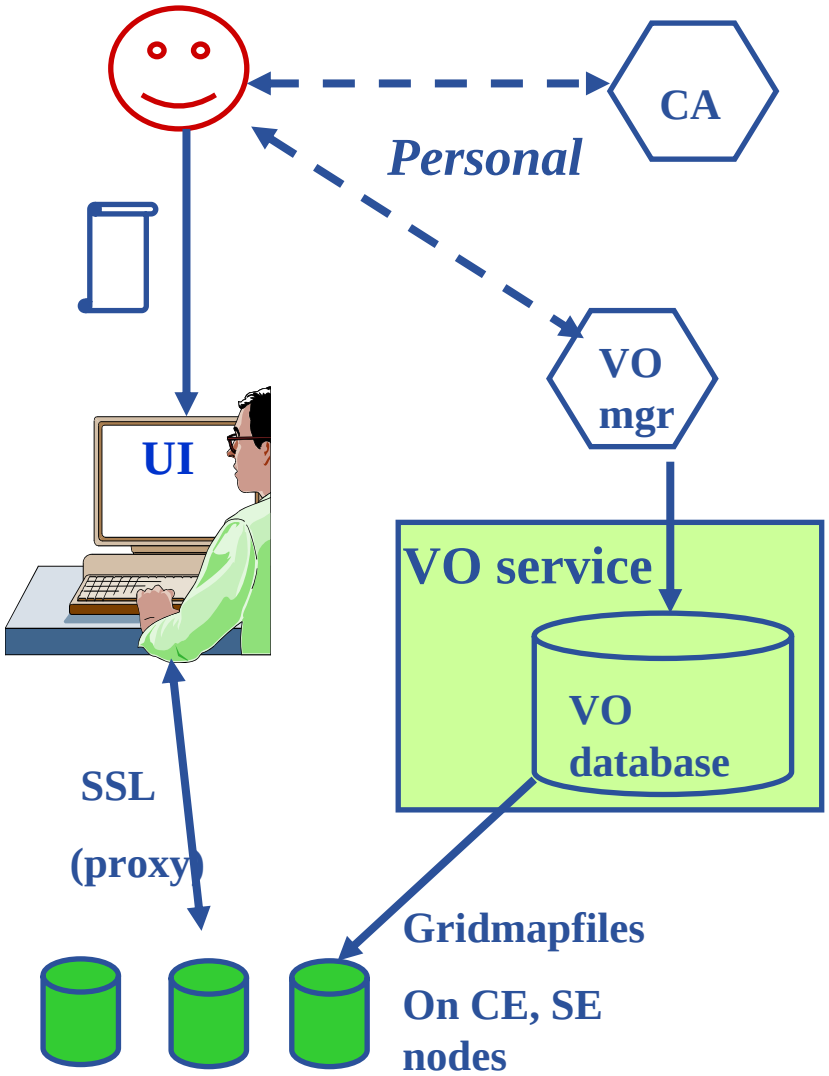
- **A VO can be defined as a group which shares Authentication and Authorization procedures in order to allow the sharing of resources**
- **Other examples**
  - Similar to ‘group’ in a file system
  - In the EU, the *Schengen* countries are a bit like a VO

A Multi-VO Grid



Security Framework: Authentication & Authorisation

- **Authentication**
  - User obtains certificate from CA
  - Connects to UI via ssh
  - Downloads certificate
  - Invokes Proxy server
  - Single sign-on – to UI - then Secure Socket Layer with proxy identifies user to other nodes
- **Authorisation - currently**
  - User joins Virtual Organisation
  - VO negotiates access to Grid nodes and resources (CE, SE)
  - Authorisation tested by CE, SE:
    - gridmapfile maps user to local account



Security Framework: Authentication & Authorisation (2)

- **Globus Toolkit™ proposed and implements the Grid Security Infrastructure (GSI)**
  - Protocols and APIs to address Grid security needs
- **GSI protocols extend standard public key protocols**
  - Standards: X.509 & SSL/TLS
  - Extensions: X.509 Proxy Certificates (single sign-on) & Delegation
- **GSI extends standard GSS-API (Generic Security Service)**
  - The GSS-API is the IETF standard for adding authentication, delegation, message integrity, and message confidentiality to applications.
- **Proxy Certificate:**
  - Short term, restricted certificate that is derived form a long-term X.509 certificate
  - Signed by the normal end entity cert, or by another proxy
  - Allows a process to act on behalf of a user
  - Not encrypted and thus needs to be securely managed by file system
- **OpenSSL now includes GSI features in standard distribution**

EGEE: adding a VO

- **Security is not only a software matter, but also human and procedures!!**

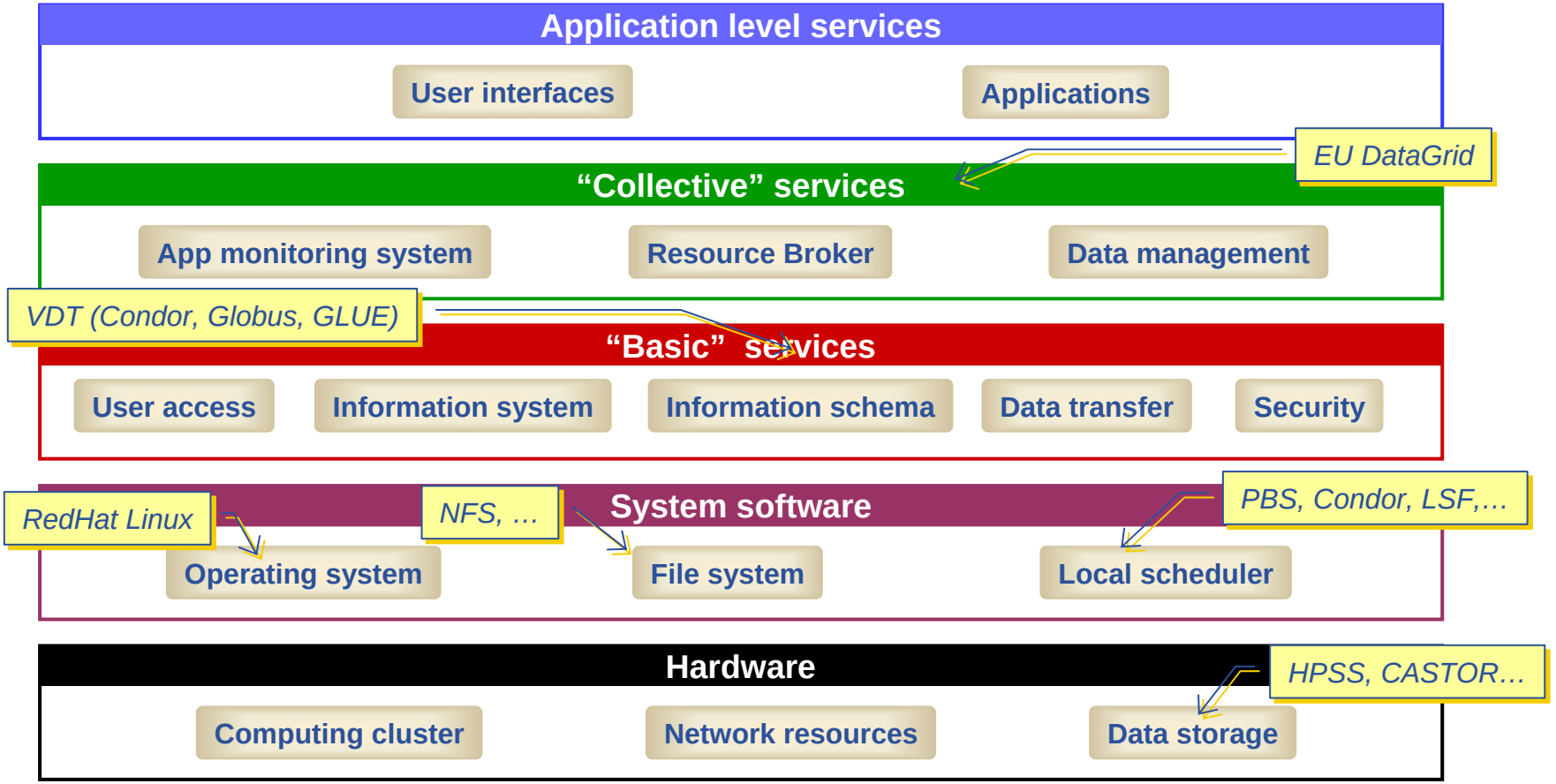
- **EGEE has a formal procedure for adding selected new user communities (Virtual Organisations)**
  - Negotiation with one of the Regional Operations Centres
  - Seek balance between the resources contributed by a VO and those that they consume
  - Resource allocation will be made at the VO level
  - Many resources need to be available to multiple VOs : shared use of resources is fundamental to a Grid

Overview

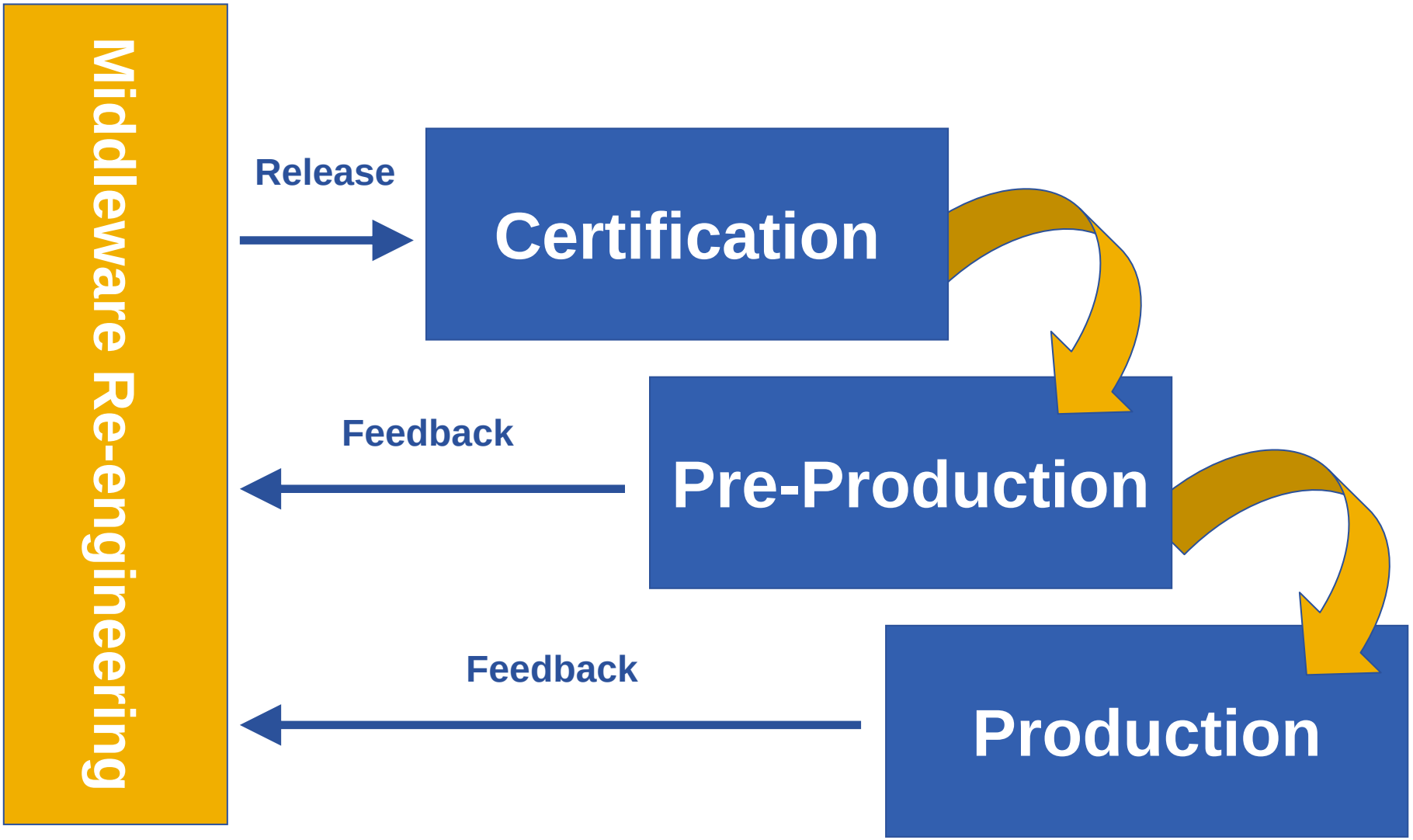
- EGEE - what is it and why is it needed?
- Middleware – current and future
- Operations
- Applications running in Production
- Summary

Current production middleware: LCG-2

- **Regular updates (latest is LCG-2.x)**
  - short term developments driven by operational priorities



Release process



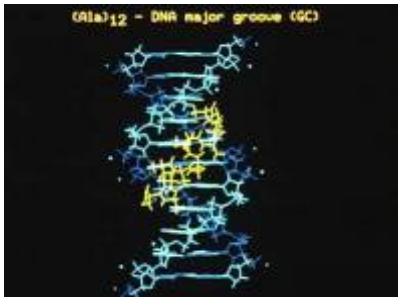
Overview

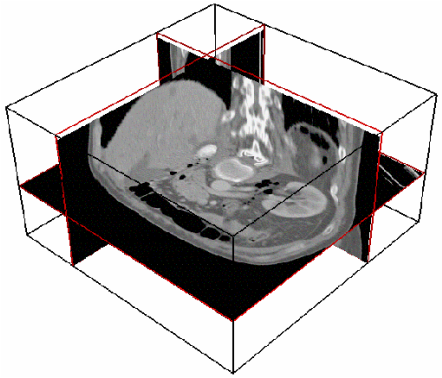
- EGEE - what is it and why is it needed?
- Middleware – current and future
- Operations
- Applications running in Production
- Summary

Pilot applications

- **High-Energy Physics (HEP)**
  - Provides computing infrastructure (LCG)
  - Challenging:
    - thousands of processors world-wide
    - generating petabytes of data
    - ‘chaotic’ use of grid with individual user analysis (thousands of users interactively operating within experiment VOs)

- **Biomedical Applications**
  - Similar computing and data storage requirements
  - Major challenge: security



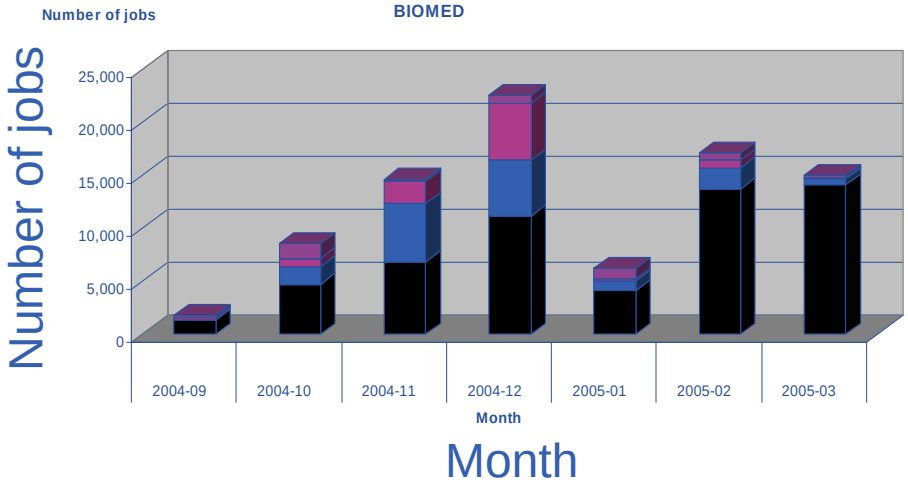


BioMed Overview

- **Infrastructure**
  - ~2.000 CPUs
  - ~21 TB disks
  - in 12 countries

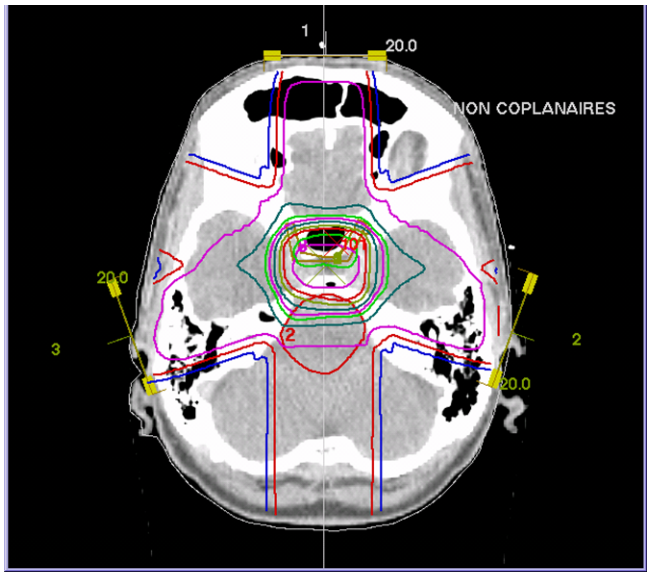
- **>50 users in 7 countries working with 12 applications**
- **18 research labs**

- **~80.000 jobs launched since 04/2004**
- **~10 CPU years**



GATE

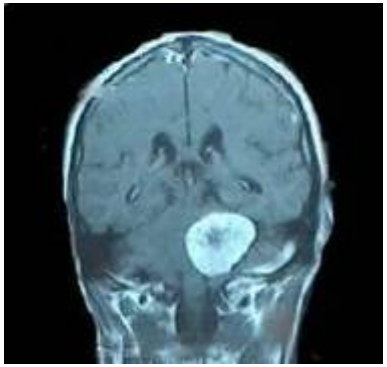
- **GEANT4 Application to Tomography Emission**
  - Scientific objectives
    - Radiotherapy planning to improve treatment of tumors computed from pre-treatment MR scans
  - Method
    - GEANT4-based software to model physics of nuclear medicine
    - Monte Carlo simulation to improve accuracy of computations



  - Grid added value
    - Splitting the random number sequences needed for Monte Carlo simulations enables independent computations
    - Parallelization reduces the total computation time
  - Results and perspectives
    - computation time reduced BUT not sufficiently for clinical practice
      → further optimizations are on-going
    - large community of users is interested in GATE

CDSS

- **Clinical Decision Support System**
  - Scientific objectives
    - Extract clinically relevant knowledge to guide practitioners in their clinical practice
  - Method
    - Starting from trained databases
    - Use classifier engines
    - Compare to annotated databases to classify data

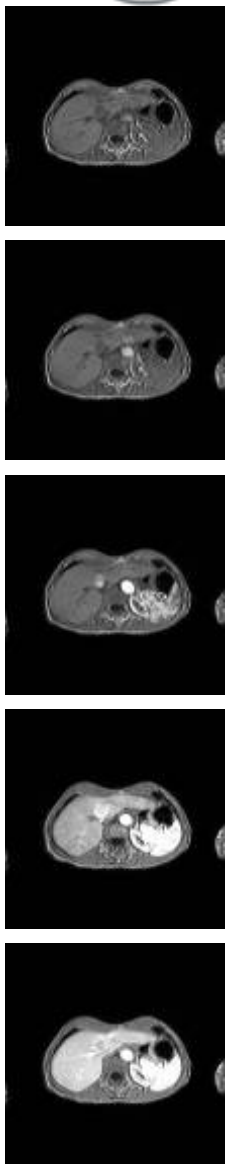


Classification of tumours in soft tissues

  - Grid added value
    - Ubiquitous access to distributed databases and classifier engines
    - Grid information system to publish and discover data sources and engines
    - Automatic management of login and security
  - Results and perspectives
    - 12 classification engines available
    - 1000 medical cases registered
    - Dynamic discovery of all engines can be implemented on top of the grid information system
    - Accounting will be provided by the grid

Pharmacokinetics

- **Co-registration of Medical Images**
  - Scientific objectives
    - Contrast Agent Diffusion to characterize tumour tissues without biopsy
  - Method
    - Co-registration requires deformable registration methods
      → compute intensive
  - Grid added value
    - Processing of compute intensive co-registration and generation of diffusion maps for the 3D MRI Studies.
    - Parallel & independent computations on different input data sets
  - Results and perspectives
    - Last clinical test:
      12 patients with 13 MRI studies each
      each study comprises 24 512x512 12-bit slices
    - Processing of the registration algorithm takes around 12 hours per study
    - Registration parameters tuned with four possible combinations
    - Each combination of parameter took 2 hours
      → 72 times faster than with a single computer

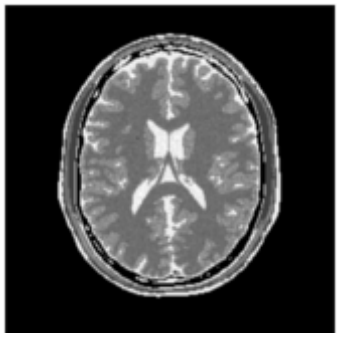


GPS@

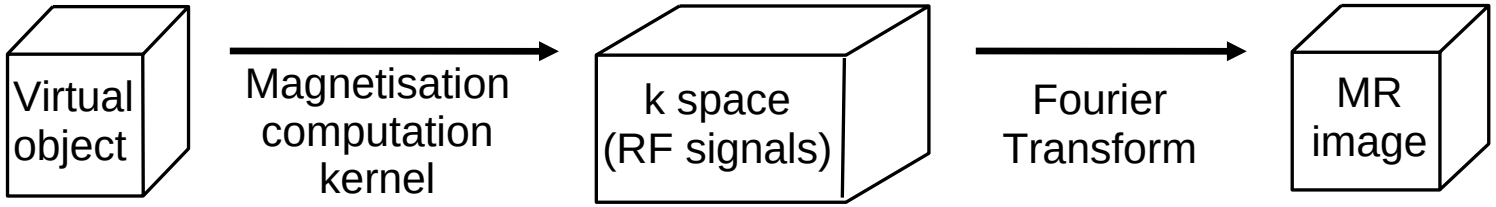
- **Grid Protein Structure Analysis**
  - Scientific objectives
    - Integrating up-to-date databases and relevant algorithms for bio-informatic analysis of data from genome sequencing projects
  - Method
    - Protein databases are stored on the grid as flat files
    - Protein sequence analysis tools run unchanged on grid resources
    - Output is analysed and displayed in graphic format through the web interface
  - Grid added value
    - Convenient way to distribute and access international databanks, and to store more and larger databases
    - Compute larger datasets with available algorithms
    - Open to a wider user community
  - Results and perspectives
    - 9 bioinformatic softwares gridified so far
    - large number of rather short jobs (few minutes each)
    - Optimizations on-going to
      - *speed up access to databases*
      - *lower short jobs latencies*

SiMRI3D

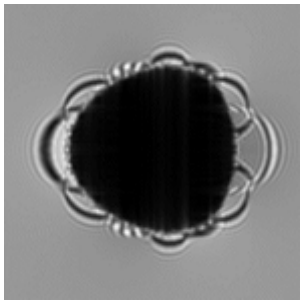
- **3D Magnetic Resonance Image Simulator**
  - Scientific objectives
    - Better understand MR physics by studying MR sequences *in silico* and MR artefacts
    - Validate MR Image processing algorithms on synthetic but realistic images



  - Method

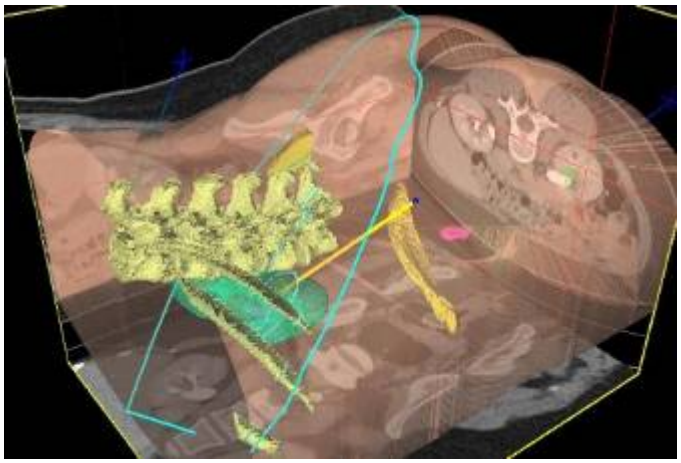


  - Grid added value
    - Speeds up the simulation time
    - Enables simulation of high resolution images
    - Offers an access to MPI-enabled clusters
  - Results and perspectives
    - Manageable computation time for medium size images
    - Development of a portal to ease access to the application
    - Implementation of new artifacts



gPTM3D

- **3D Medical Image Analysis Software**
  - Scientific objectives
    - Interactive volume reconstruction on large radiological data
  - Method
    - Starting from hand-made initialization
    - Algorithm segments each slice of a medical volume
    - 3D reconstruction by triangulating contours from consecutive slices
  - Grid added value
    - Interactive reconstruction time: less than 2mins and scalable
    - Permanent availability of resources for fast reconstruction
    - Access to users at non grid-enabled sites (e.g. hospital)
    - Unmodified medically optimized interface
  - Results and perspectives
    - Successfully ported and demonstrated at first EGEE review
    - Streams to/from non EGEE-enabled sites specific protocol, CrossGrid glogin will be considered
    - Resource access QoS: ongoing work



xmipp_ML_refine

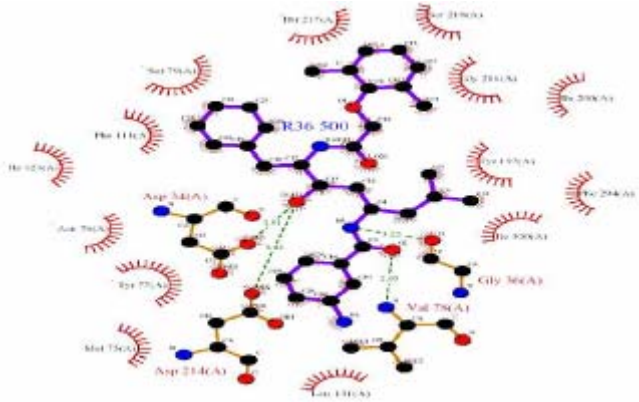
- **Macromolecules structure analysis from electron microscopy**
  - Scientific objectives
    - 3D reconstruction of molecular structural information from cryo-electron microscopy
  - Method
    - Multi-reference refinement of electron microscopy structures through a maximum likelihood statistical approach
  - Grid added value
    - Very compute intensive analysis of multiple structures
      - *2D: one to several weeks on a single CPU*
      - *3D: even more costly*
    - Computation can be split in independent jobs that are executed in parallel
  - Results and perspectives
    - First results on 2D analysis show significant time gain:
      two months on a local cluster (20 CPUs) versus one month on the grid
    - algorithm still being optimized and ported to 3D case
    - MPI implementation is currently being developed that should significantly improve the computation time

xmipp_ML_CTFs

- **Electron microscope images correction**
  - Scientific objectives
    - Electron microscopy images impaired by electron sources and defocus of magnetic lenses used in experimental practice
    - Image aberrations are described by a Contrast Transfer Function (CTF) that need to be estimated to fix images
    - CTF estimation lead to drastic image enhancement
  - Method
    - Auto regressive modelling is used to estimate parameters of the CTF and produce more reliable results than classical Fourier transform-based approaches
  - Grid added value
    - Very compute intensive: complex functional, slow optimisation process
    - Parallelisation on different grid resources
  - Results and perspectives
    - → 2 months on a single CPU
    - → 2 days on a local 20-CPUs cluster
    - → 14 hours on the grid

In silico Drug Discovery

- Scientific objectives
  - Provide docking information to help in the search for new drugs
  - Propose new inhibitors (drug candidates) addressed to neglected diseases
  - *In silico* virtual screening of drug candidate databases
- Method
  - Large scale molecular docking on malaria to compute millions of potential drugs with different software and parameters settings
- Grid added value
  - Drug discovery usually takes up to 12 years to complete
  - Docking much faster, but large databases lead to heavy computations
    → split candidate drug input on different grid resources
- Results and perspectives
  - Limited size computation (105 candidate drugs tested for 1 protein target) achievable in 2 days using the Grid compared to 6 months of CPU time
  - Full data challenge planed
    - *3x106 candidate drugs against 5 protein targets*
    - *Total computing time will reach 80 years of CPU and 6 TB of storage*



SPLATCHE

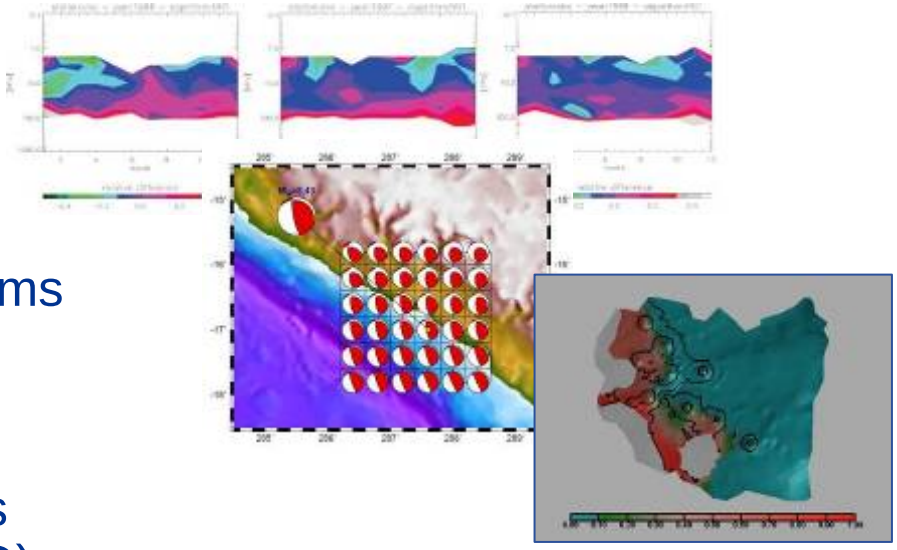
- **Genome evolution modeling**
  - Scientific objectives
    - Study human evolutionary genetics and answer questions such as
      - *geographic origin of modern human populations*
      - *genetic signature of expanding populations*
      - *genetic contacts between modern humans and Neanderthals*
  - Method
    - Simulate past demography of human populations in a geographically realistic landscape
    - Generate molecular diversity of samples of genes drawn from the current human's range, and compare to observed contemporary molecular diversity
  - Grid added value
    - Due to the Bayesian approach used, the SPLATCHE application is very compute intensive
    - Independent simulations can be executed in parallel
  - Results and perspectives
    - Application prototype ported on the EGEE middleware
    - Scale tests on the full grid infrastructure underway

Generic Applications

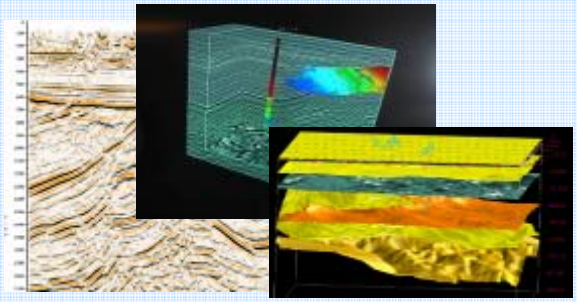
- **6 applications selected so far:**
  - Earth sciences
  - MAGIC
  - Computational Chemistry
  - PLANCK
  - Drug Discovery
  - GRACE (end February 2005)

Earth sciences applications

- **Earth Observations by Satellite**
  - ozone profiles
- **Solid Earth Physics**
  - Fast Determination of mechanisms of important earthquakes
- **Hydrology**
  - Management of water resources in Mediterranean area (SWIMED)



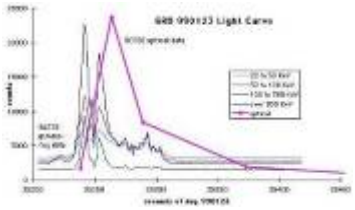
- **Geology**
  - Geocluster: R&D initiative of the Compagnie Générale de Géophysique

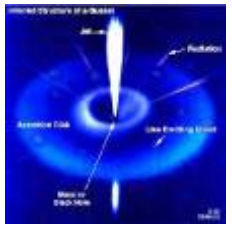


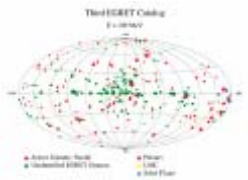
- **A large variety of applications ported on EGEE which incites new users**
- **Interactive Collaboration of the teams around a project**

MAGIC

- **Ground based Air Cerenkov Telescope 17 m diameter**
- **Physics Goals:**
  - Origin of VHE Gamma rays
  - Active Galactic Nuclei
  - Supernova Remnants
  - Unidentified EGRET sources
  - Gamma Ray Burst
- **MAGIC II will come 2007**
- **Grid added value**
  - Enable “(e-)scientific“ collaboration between partners
  - Enable the cooperation between different experiments
  - Enable the participation on Virtual Observatories

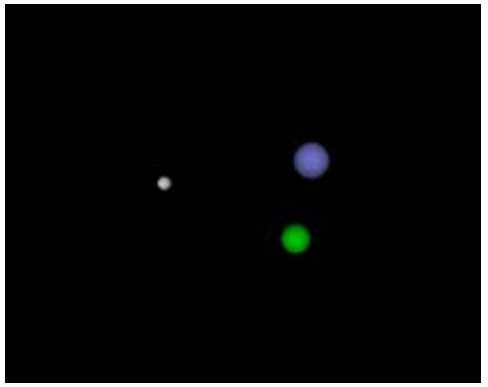


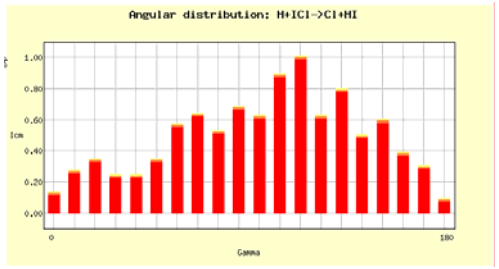


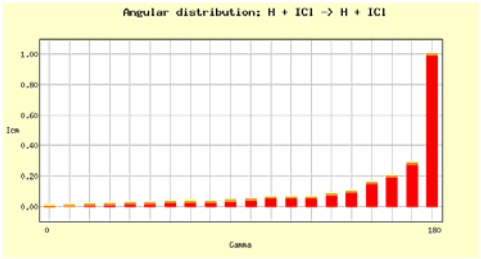


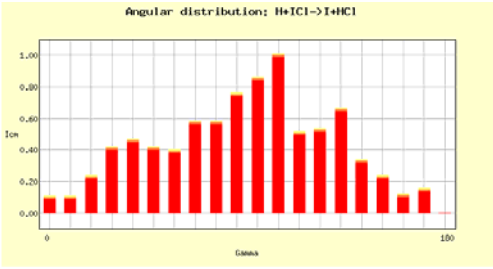
Computational Chemistry

- **The Grid Enabled Molecular Simulator (GEMS)**
  - Motivation:
    - Modern computer simulations of biomolecular systems produce an abundance of data, which could be reused several times by different researchers.
      → data must be catalogued and searchable
  - GEMS database and toolkit:
    - autonomous storage resources
    - metadata specification
    - automatic storage allocation and replication policies
    - interface for distributed computation



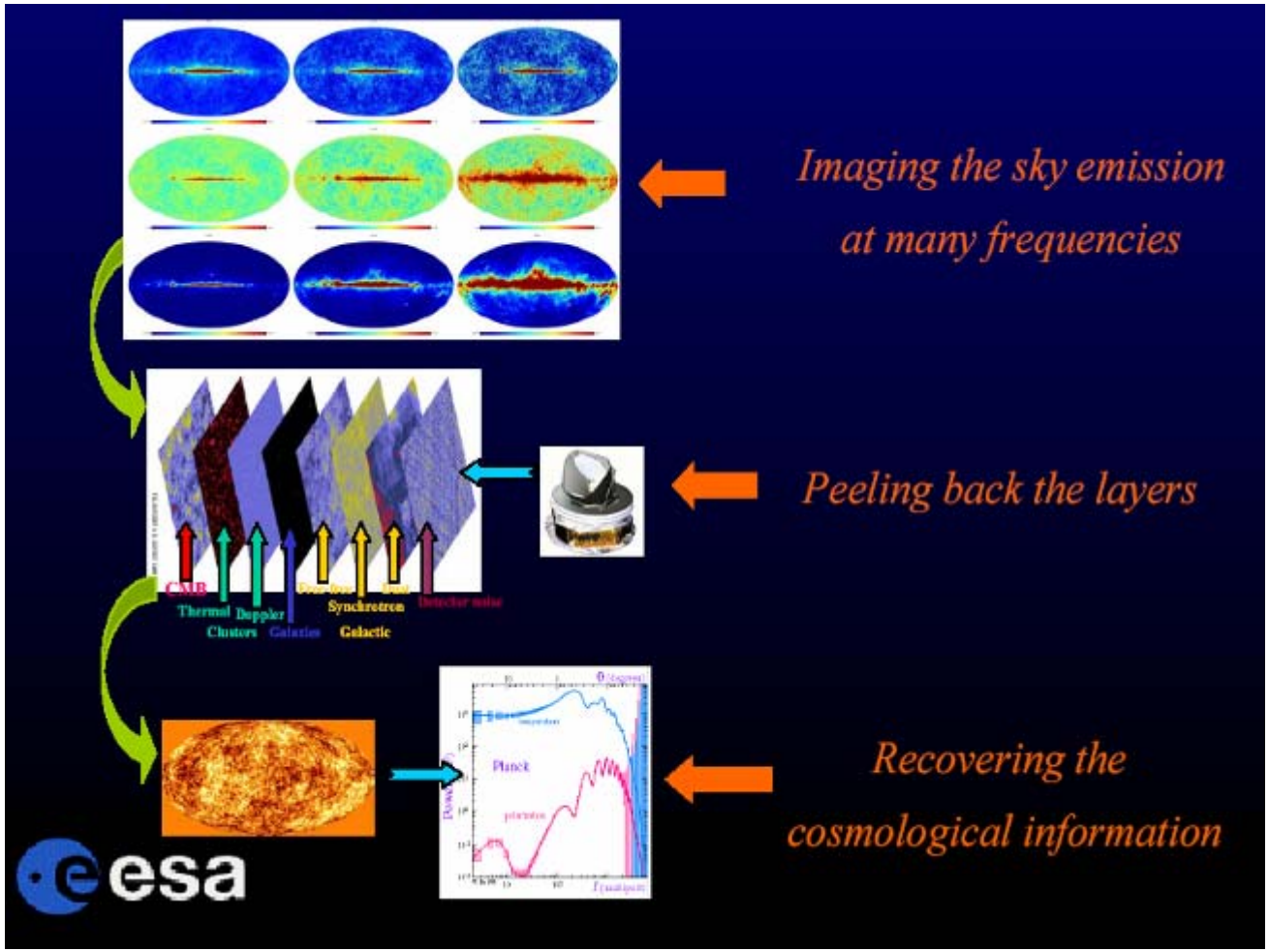






Planck

- **On the Grid:**
  > 12 time faster
  (but ~5% failures)
- **Complex data structure**
  → data handling important
- **The Grid as**
  - collaboration tool
  - common user-interface
  - flexible environment
  - new approach to data and S/W sharing



Almost Over…

- EGEE - what is it and why is it needed?
- Middleware – current and future
- Operations
- Applications running in Production
- Summary - (can we run jobs now?)

Summary

- The Grid attempts to provide scientists with the next generation IT infrastructure
- EGEE is the first attempt to build a worldwide Grid infrastructure for data and compute 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
- EGEE is deploying the next generation middleware: gLite

Things I didn't have time to cover

- **Grid Operation**
- **Support**
  - Application
  - Resource Centers
- **Application induction and incubation process**
- **External relations**
- **EGEE-II and beyond…**

- **Enjoy the tutorial…**

Questionnaire

Your feedback is important…

Please fill-in the questionnaire!!

Contacts

- **EGEE Website**
  http://www.eu-egee.org
- **How to join**
  http://public.eu-egee.org/join/
- **EGEE Project Office**
  project-eu-egee-po@cern.ch
- **How to learn more about Grids**
  http://public.eu-egee.org/test/