



Enabling Grids for E-scienceE

Introduction to Grids, EGEE and gLite (Middleware for Grid computing)

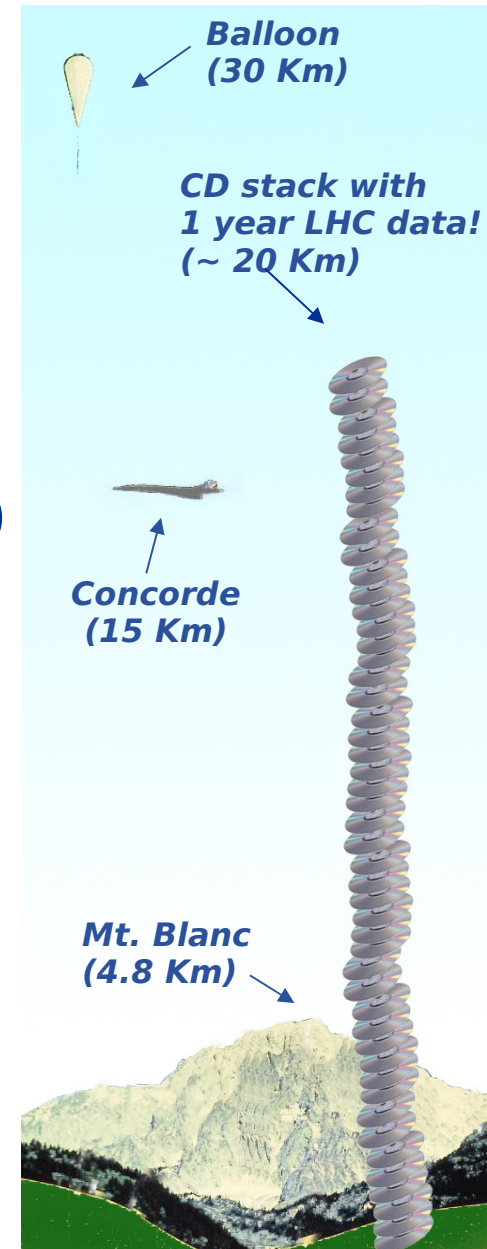
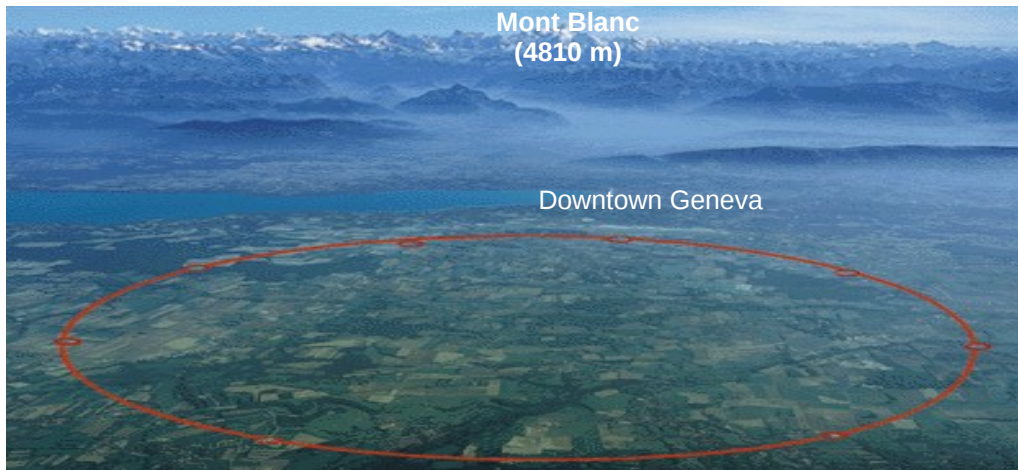
www.eu-egee.org



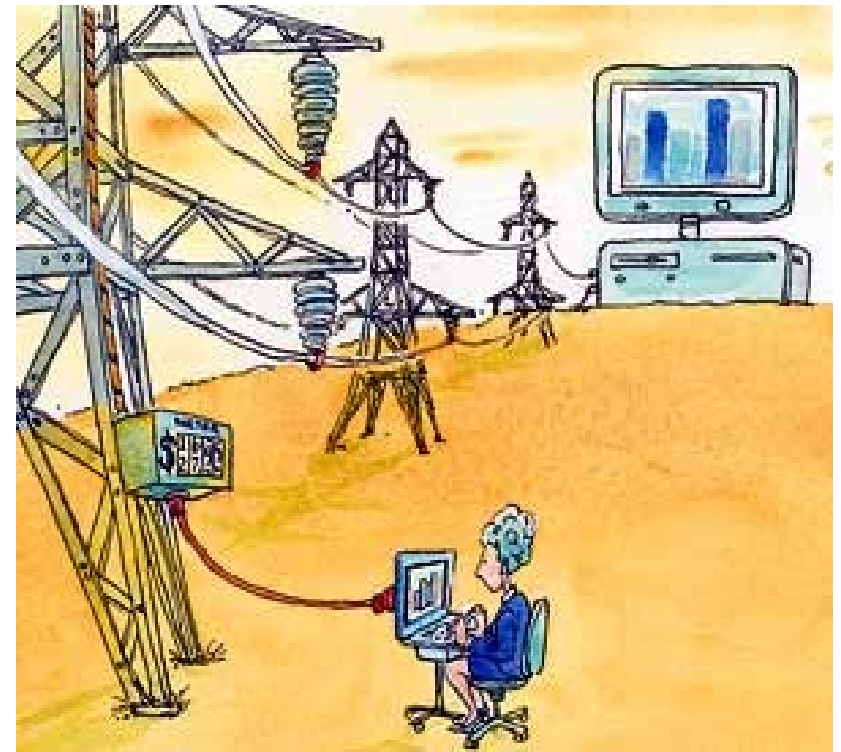
- **Introduction to the Grid**
- **EGEE project**
- **gLite middleware**
 - Overview and architecture

Why the Grid?

- Science is becoming increasingly digital and needs to deal with increasing amounts of data
 - ▽ Large amount of data produced
 - ▽ Large worldwide organized collaborations
- e.g. Large Hadron Collider (LHC) at CERN (Geneva)
 - ~10 petabytes/year (~10 Million GBytes)



... **securely share distributed resources (computation, storage, etc)** so that users can **collaborate within Virtual Organisations (VO)**



Various scopes: production grids/research grids More or less homogeneous grids:

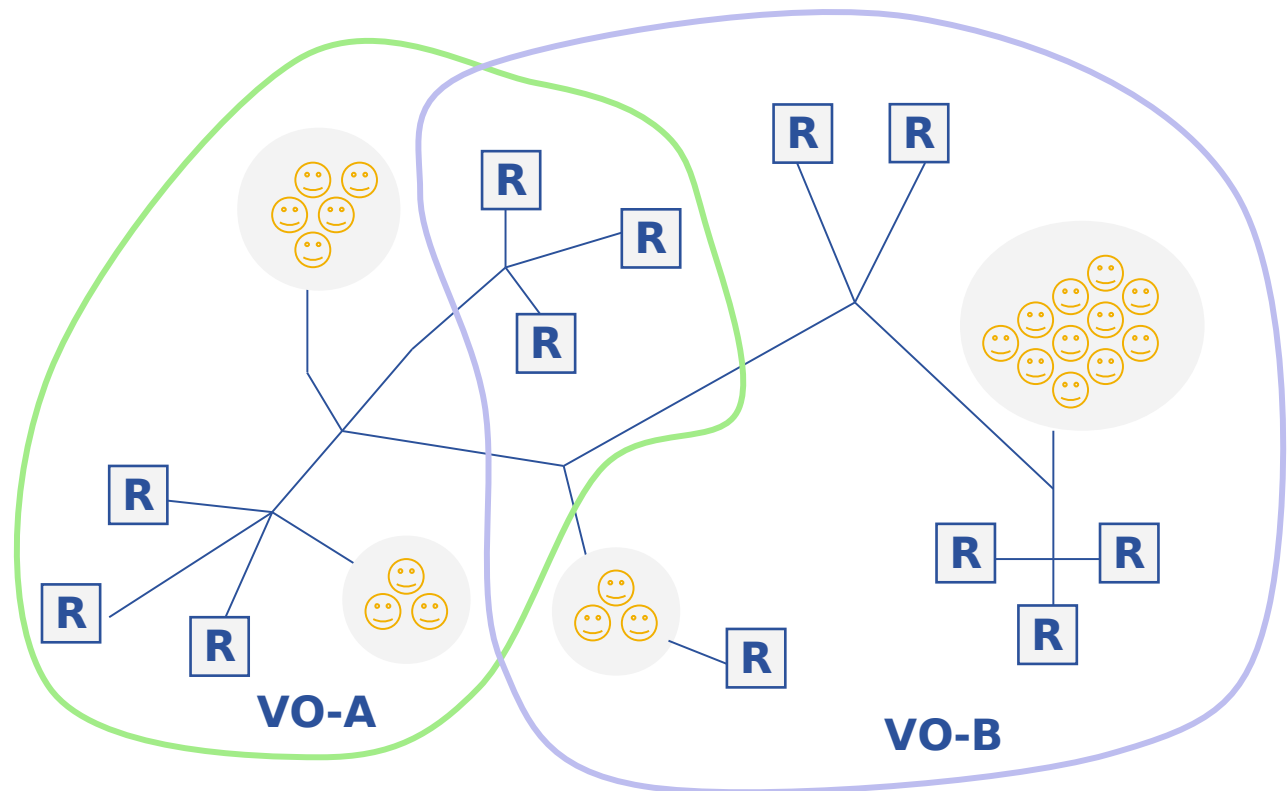
- desktop grids
 - BOINC / @home
 - XtremOS
 - service grids
 - EGEE / OSG / NAREGI ...
 - decryphon
 - HPC grids
 - DEISA, TeraGrid, ...
- eventually interoperable.**



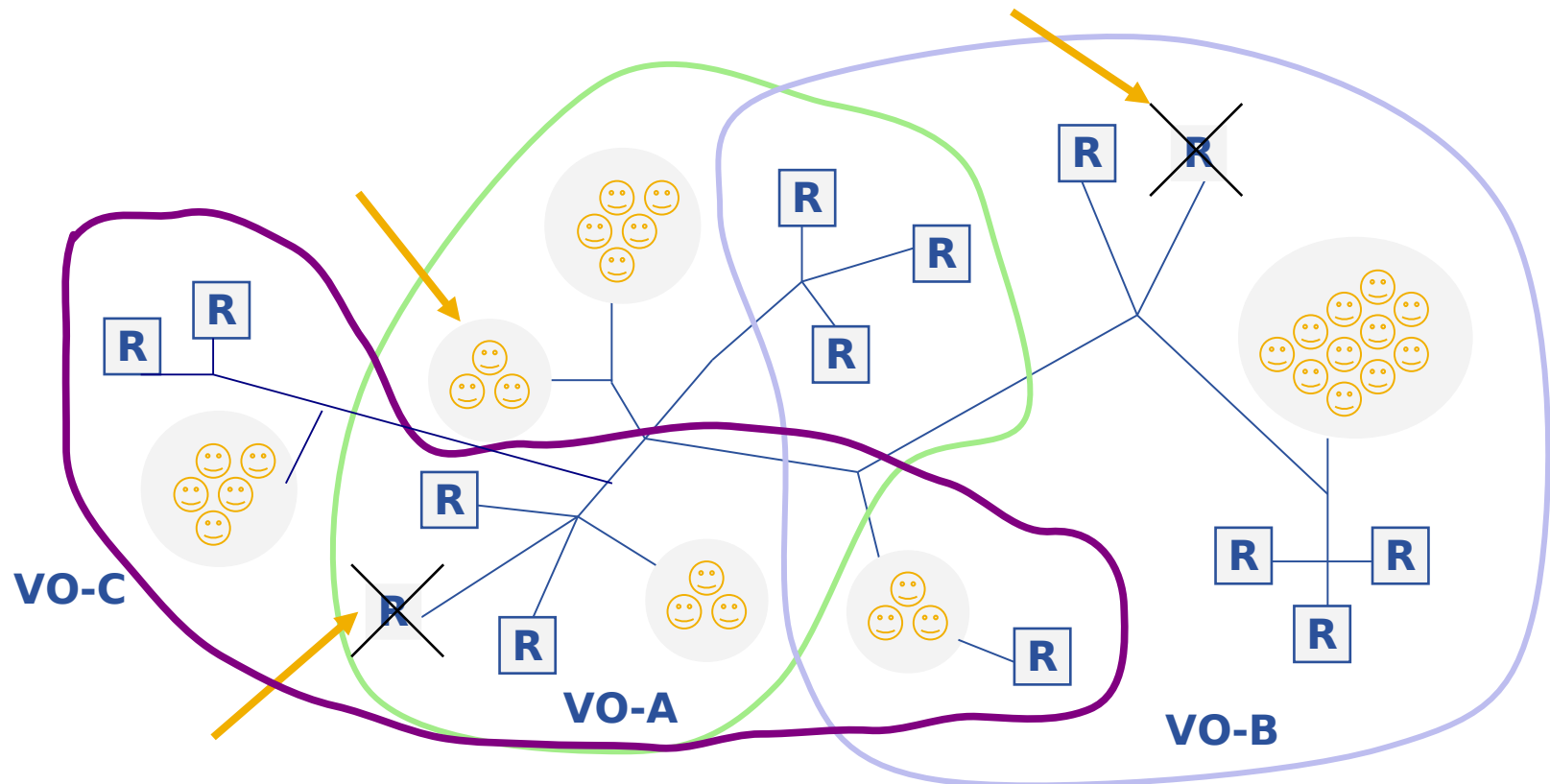
On many Grids, users are grouped in Organisations: Virtual

- **A group of people from different institutions working on a common goal**
- **Sharing distributed processing and data resources**
 - Computers
 - Data files
 - Scientific instruments
 - Codes
 - ...

- Distributed resources and people
- Linked by networks, cross-administrative domains
- Sharing resources, common goals



- Distributed resources and people
- Linked by networks, cross-administrative domains
- Sharing resources, common goals
- Dynamic, fault tolerant



- **Virtual Organization Membership Service (VOMS)**
 - EGEE/gLite enhancement for VO management
 - Provides information on user's relationship with Virtual Organization (VO)
 - Membership
 - Group membership
 - Roles of user
 - Has an Account Database
 - Comparable to a Kerberos server
 - Serving information in a special format (VOMS credentials)
 - Administration via command line & web interface

- **Single login creating a proxy at the beginning of a session**
 - Attaches VOMS certificate to user proxy certificate
- **Expiration time**
 - Authorization valid for a limited time (may differ from proxy duration)
- **Multiple VO**
 - User can register to multiple VOs
- **Backward compatibility**
 - Extra VO related information in users proxy certificate
 - Users proxy can still be used with non VOMS-aware services
- **Security**
 - Client/Server communication are secured and authenticated

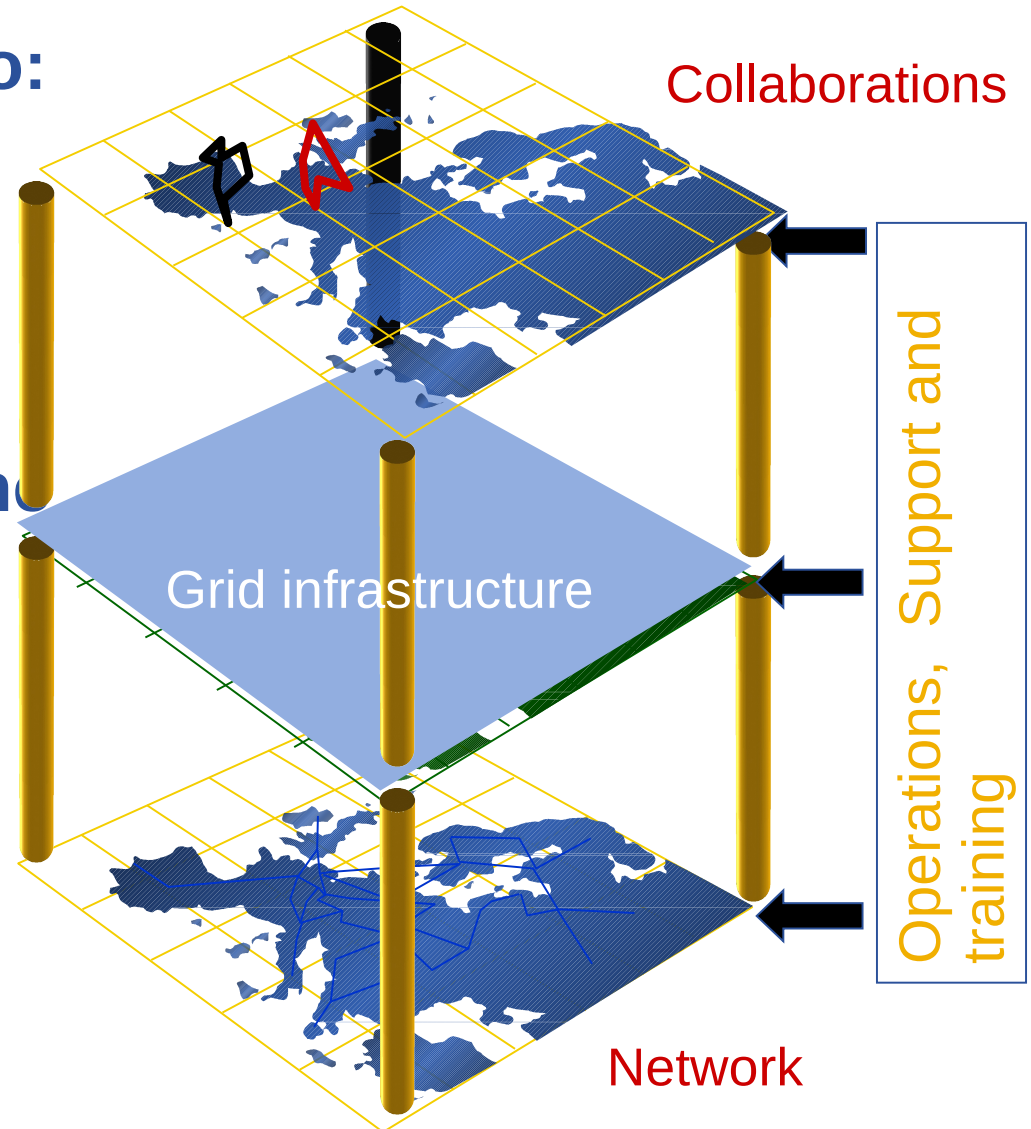
EGEE : Enabling Grids for E-science

- The flagship Grid Infrastructure project of the EU
- Funded by the European Commission
- Primary Objectives
 - consistent, robust and secure service grid infrastructure
 - improving and maintaining the middleware
 - attracting new users from industry and science
 - *ensure they receive high standard of training and support*
- Structure
 - over 250 computing centers in more than
 - 48 countries, federating in regional Grids
 - 60.000 CPUs, > 5 Petabytes storage

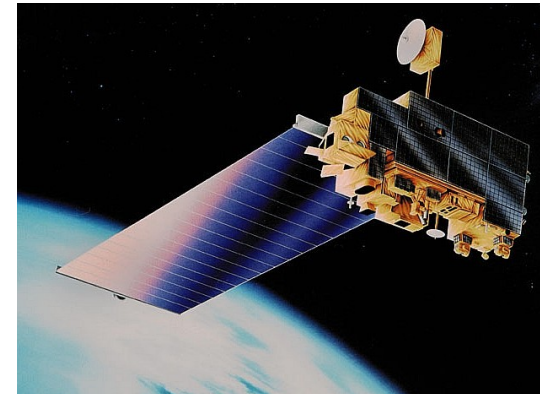
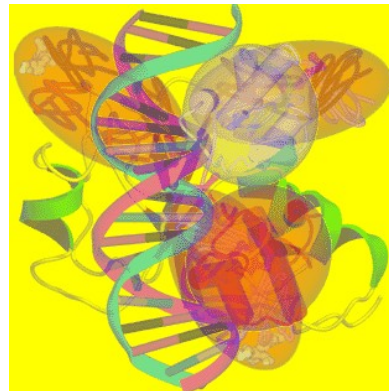


Build a large-scale production grid service to:

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



- Particle Physics
- Bioinformatics
- Industry
- Astronomy
- Chemistry
- Earth Observation
- Geophysics
- Biodiversity
- Nanotechnology
- Climate Modeling



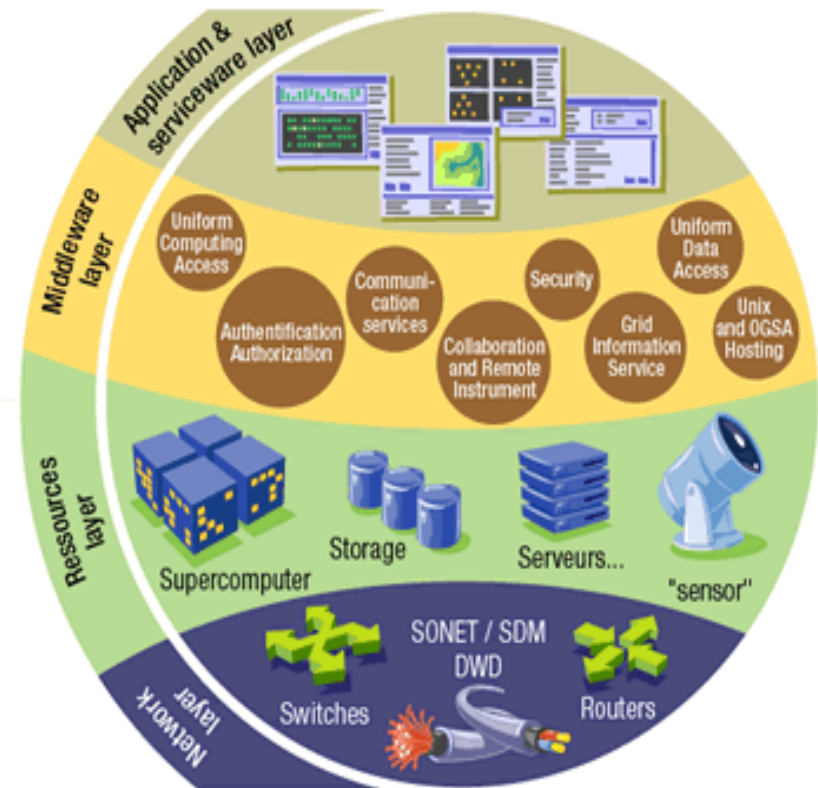
- **Exploring big parameter space**
- **Deterministic / probabilistic**
- **“Gridifying “ legacy code / designed for Grids**

but also:

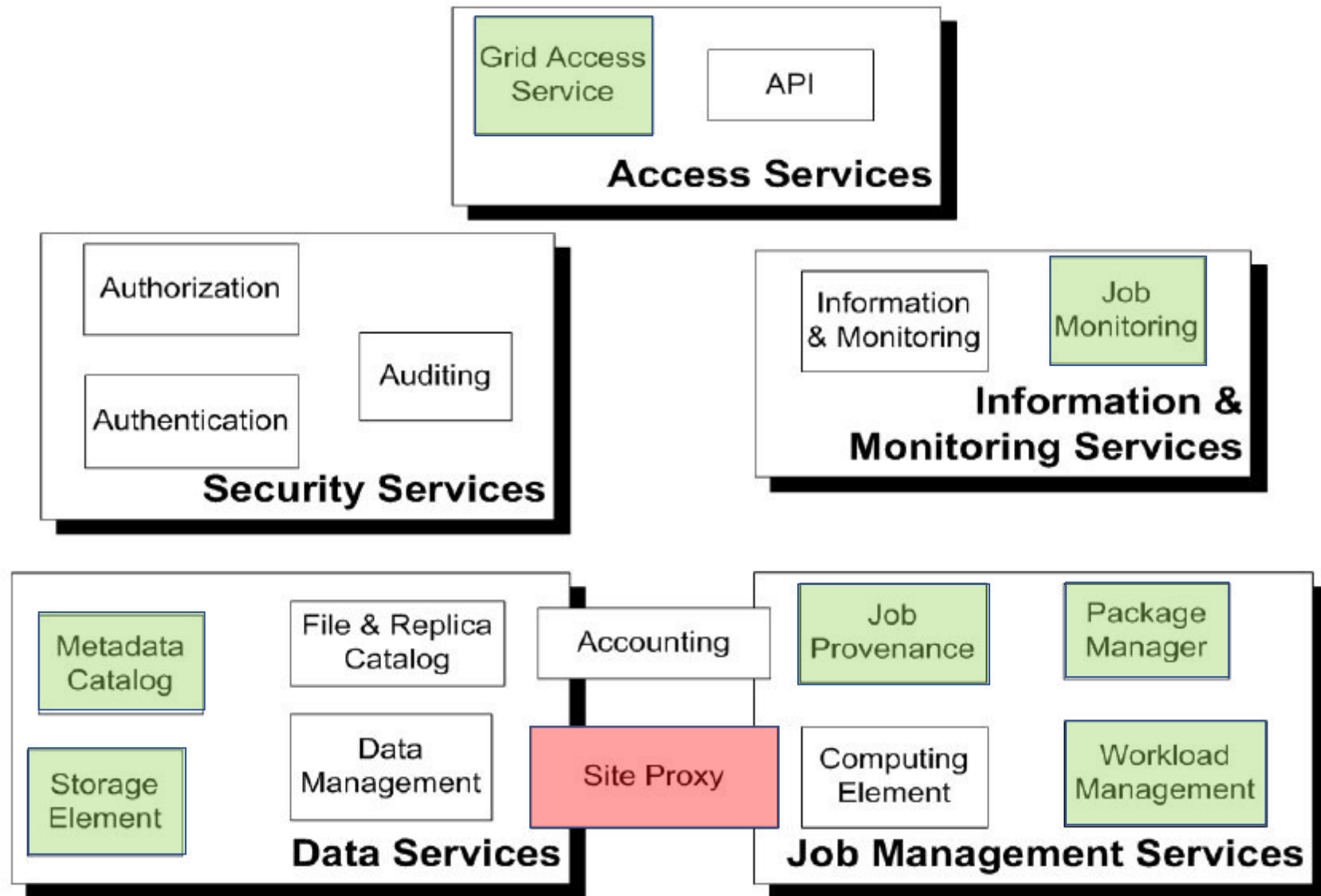
- **standalone / parallel (MPI)**
- **short runs / long runs (checkpoints)**

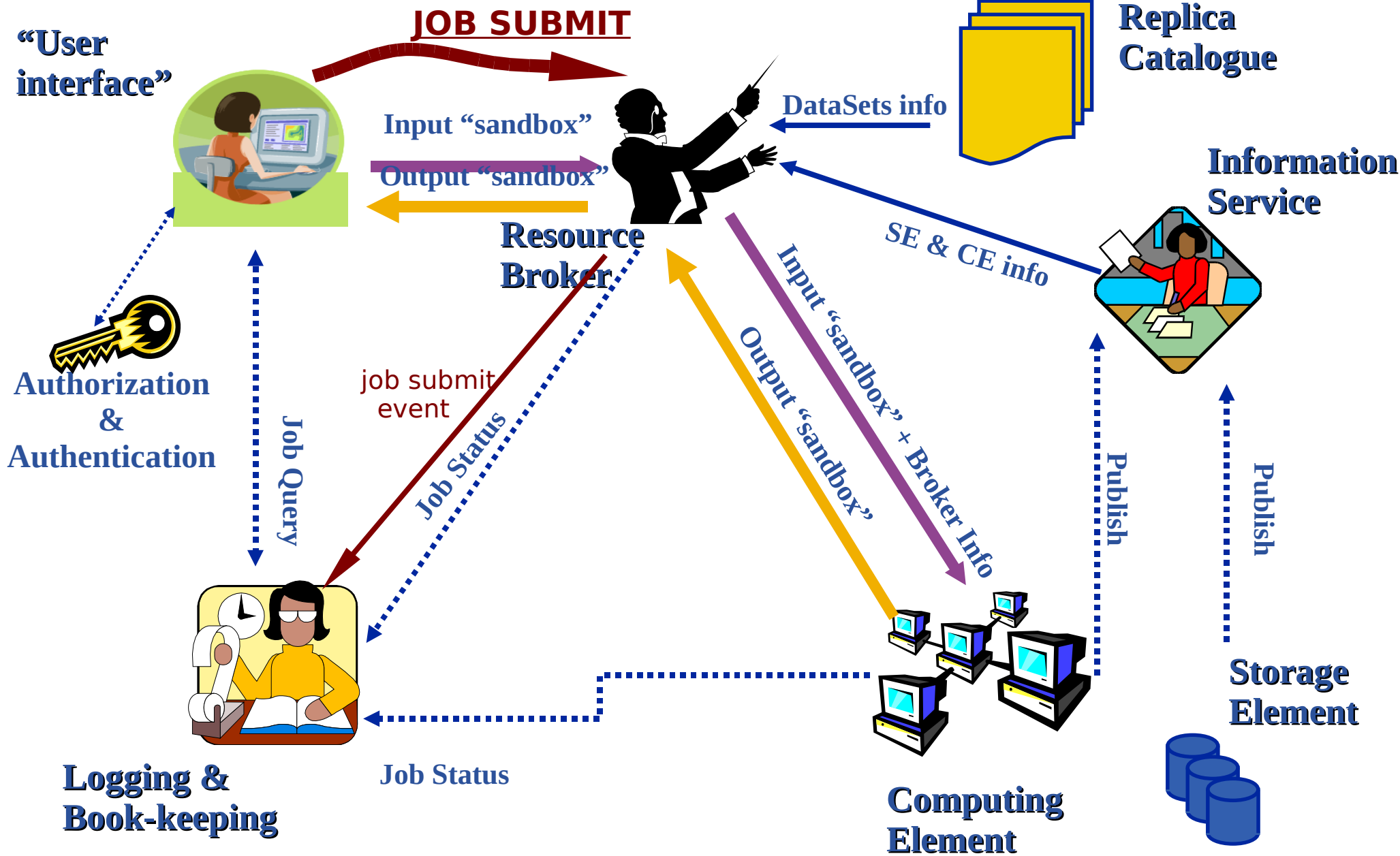
**Often complex data-flow and control-flow
(Grid workflow applications)**

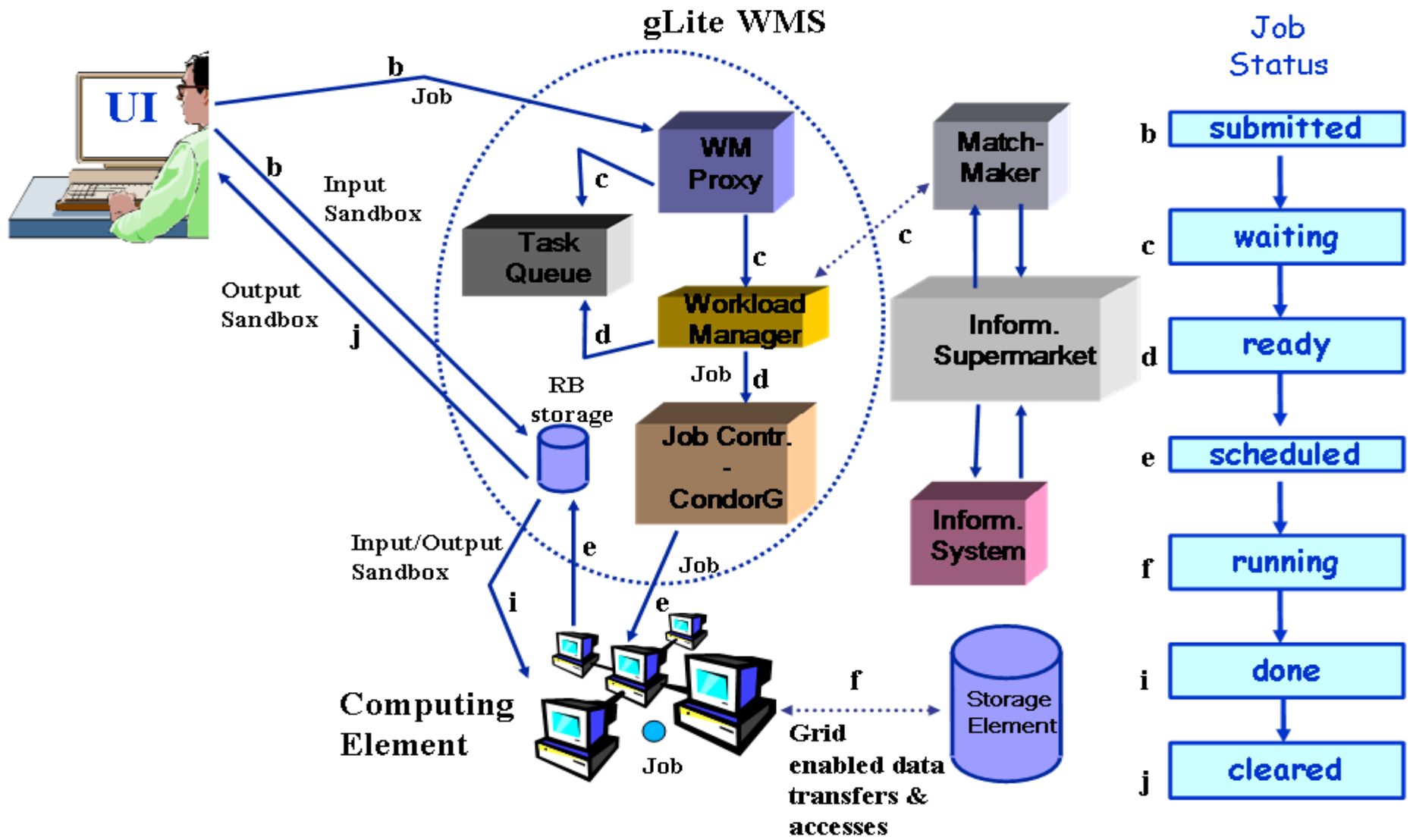
- The Grid relies on advanced software – the middleware - which interfaces between resources and the applications
- The GRID middleware
 - Finds convenient places for the application to be executed
 - Optimises use of resources
 - Organises efficient access to data
 - Deals with authentication to the different sites that are used
 - Run the job & monitors progress
 - Transfers the result back
 - accounts all operations



- **gLite**
 - First release 2005 (currently gLite 3.1)
 - Next generation middleware for grid computing
 - Intended to replace present middleware with production quality services
 - Developed from existing components (globus, condor,..)
 - Interoperability & Co-existence with deployed infrastructure
 - Robust: Performance & Fault tolerance
 - Open Source license
 - Platform: Currently only Scientific Linux supported







File Transfer Service

Metadata management (interface to databases management systems)

- **AMGA**
- **GRELC / GDSE / OGSA-DAI**

Alternate WMS

- **gridway**

gLite user documentation

- <https://edms.cern.ch/document/722398/>

User & Applications portal

- <https://egeena4.lal.in2p3.fr>

Troubleshooting

- <https://gus.fzk.de>

Job monitoring

- <http://gridview.cern.ch>

Grid monitoring

- <http://gridportal.hep.ph.ic.ac.uk/rtm/>
- <http://goc.grid.sinica.edu.tw/gstat>



Questions?