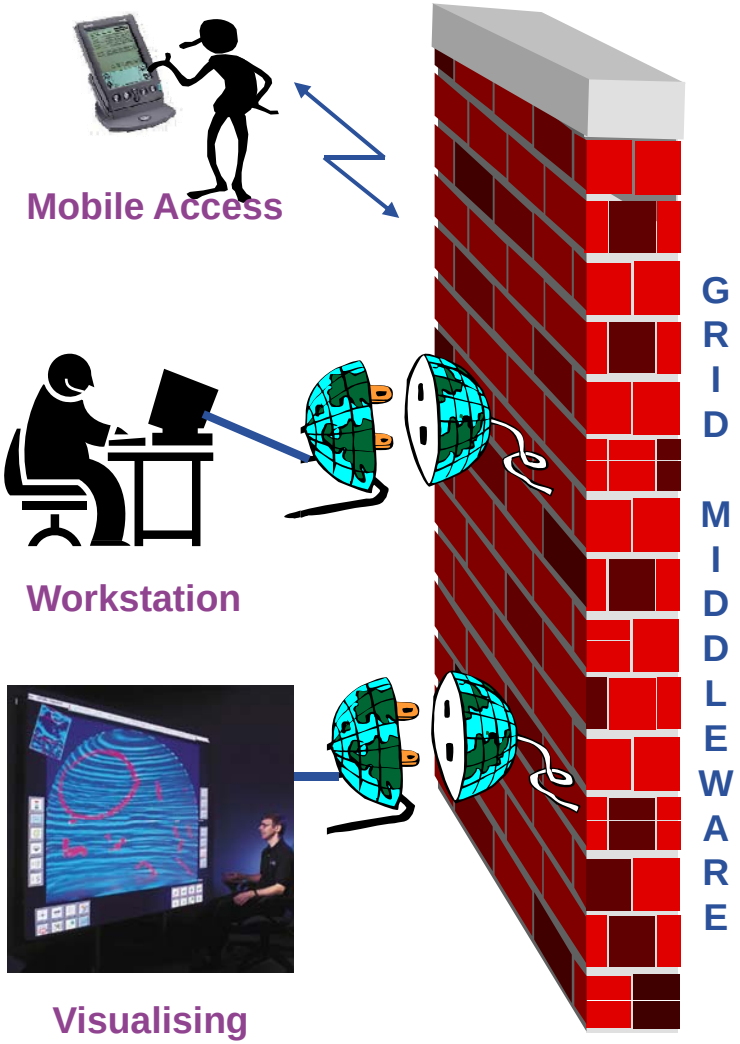
Overview of gLite grid middleware

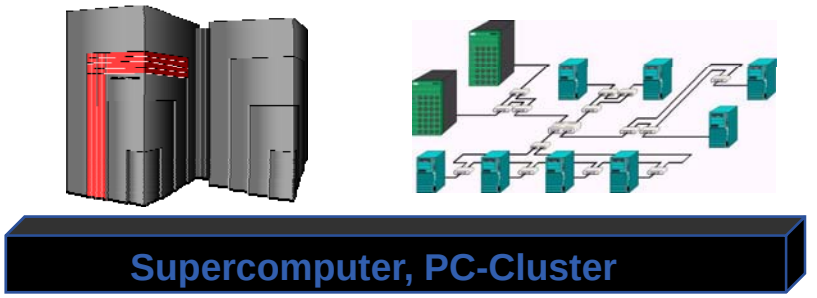
Gergely Sipos
MTA SZTAKI
sipos@sztaki.hu

Outline

- Basic features of gLite middleware
- gLite services
- Security in gLite
- External services

Grid vision


Mobile Access
Workstation
Visualising
GRID MIDDLEWARE


Supercomputer, PC-Cluster


Data-storage, Sensors, Experiments


Internet, networks

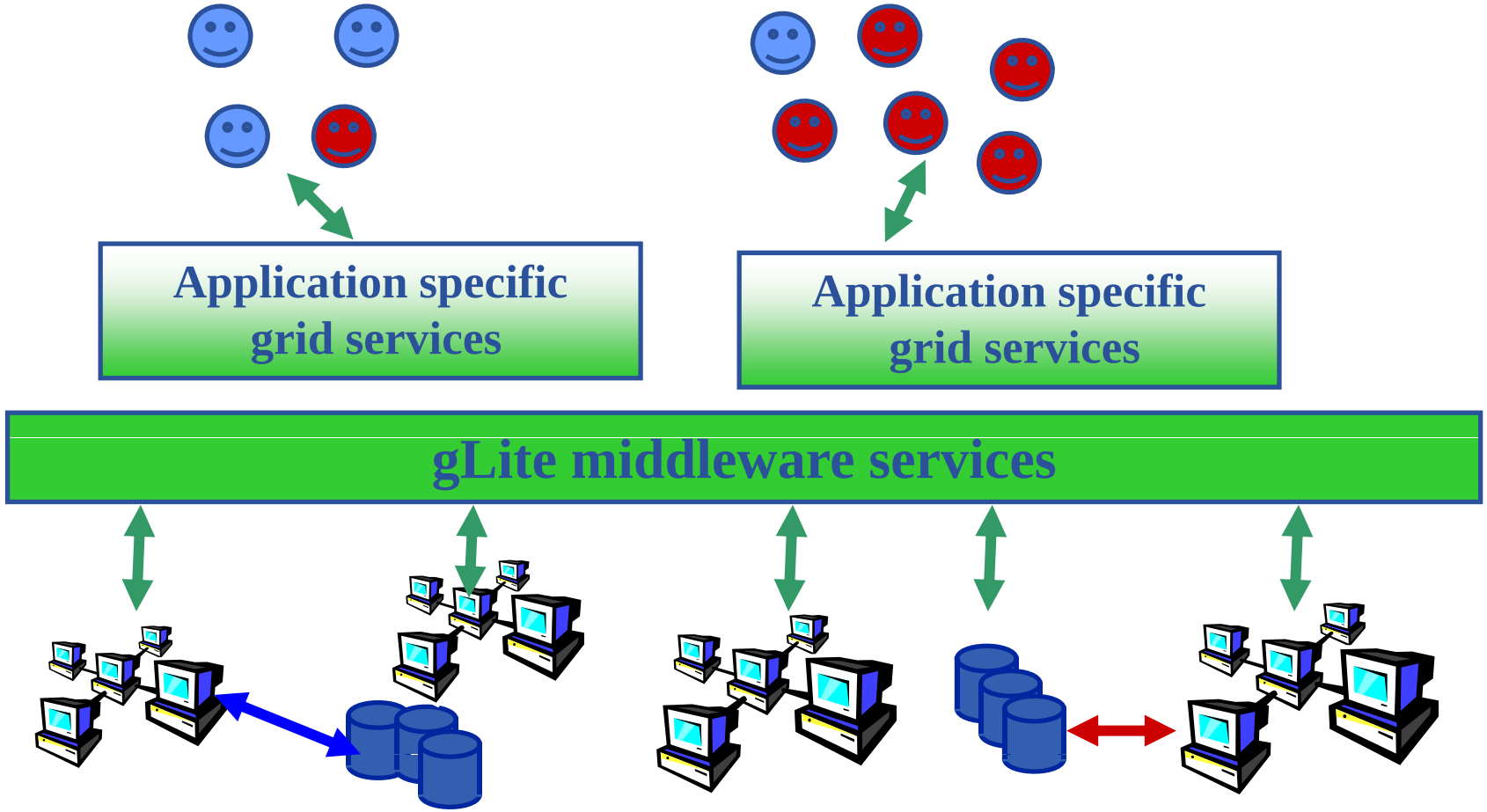
Problems to solve

- **Standardised access to resources**
  - Computers
  - Storages
  - Special equipments
  - Software services
- **Access policy**
- **Load balancing**
- **Monitoring resources and services**
- **Monitoring applications**
- **Fault management**
- **Programming contepts, level of abstraction**
- **User interfaces**
- **...**

Mostly application specific

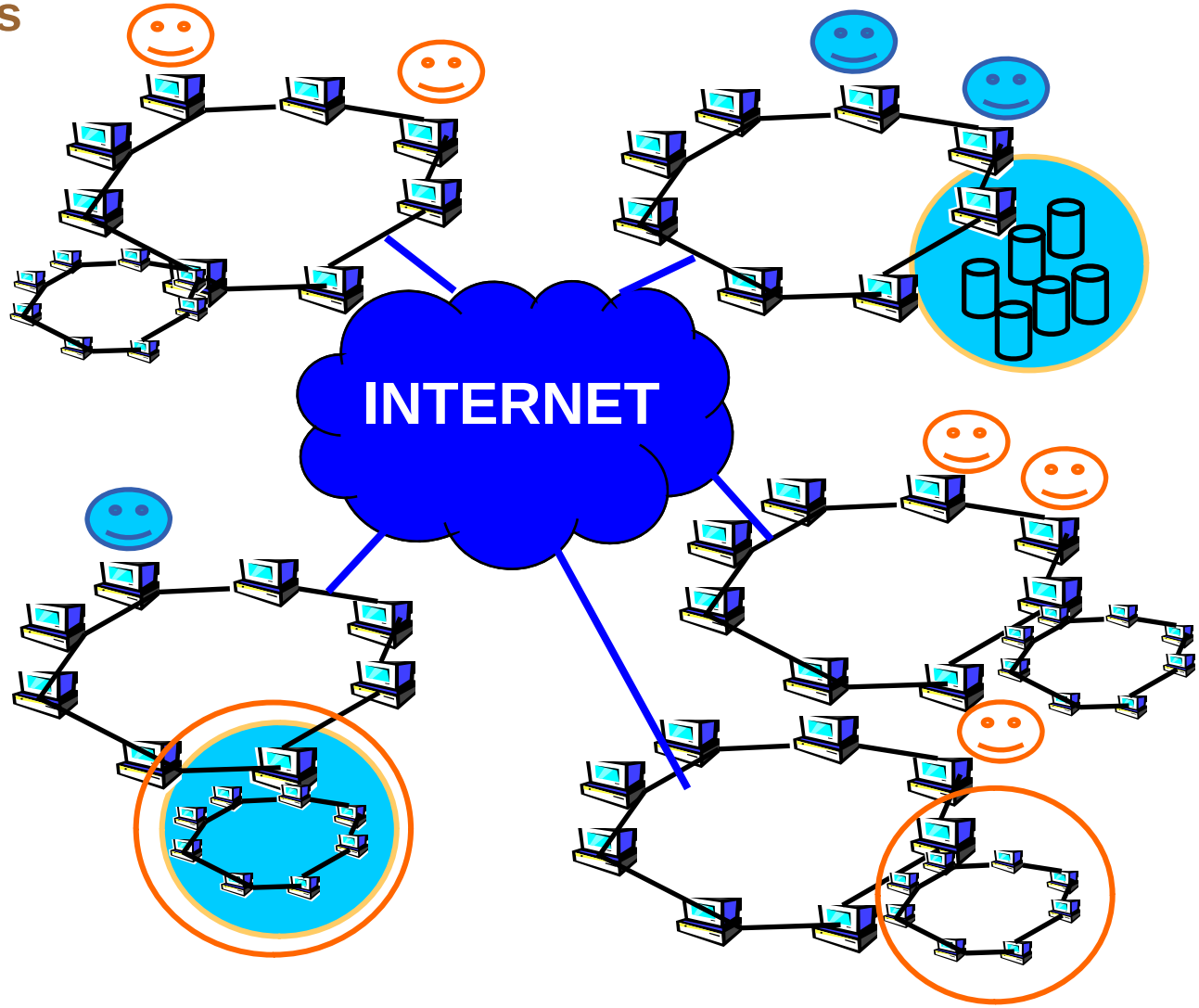
EGEE grid, gLite middleware

- **EGEE is establishing a production grid infrastructure with gLite middleware**



VO concept

- **gLite middleware runs on each shared resource to provide**
  - Data services
  - Computation services
  - Security service
- **Resources and users form Virtual organisations: basis for collaboration**
- **Distributed services (both people and middleware) enable the grid**



Grid Middleware

- **When using a PC or workstation you**
  - Login with a username and password (“Authentication”)
  - Use rights given to you (“Authorisation”)
  - Run jobs
  - Manage files: create them, read/write, list directories
- **Components are linked by a bus**
- **Operating system**
- **One admin. domain**

- **When using a Grid you**
  - Login with digital credentials – single sign-on (“Authentication”)
  - Use rights given you (“Authorisation”)
  - Run jobs
  - Manage files: create them, read/write, list directories
- **Services are linked by the Internet**
- **Middleware**
- **Many admin. domains**

EGEE Middleware: gLite

- **gLite 3.0, gLite 3.1**

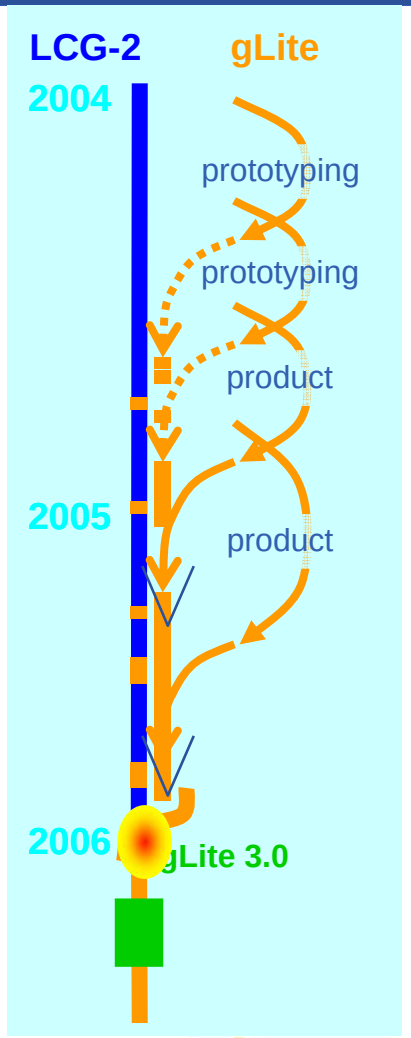
⇨ Merger of LCG 2.7 and gLite 1.5

  - **Scientific Linux v3 and v4**
  - **Ongoing efforts to port to other OS**

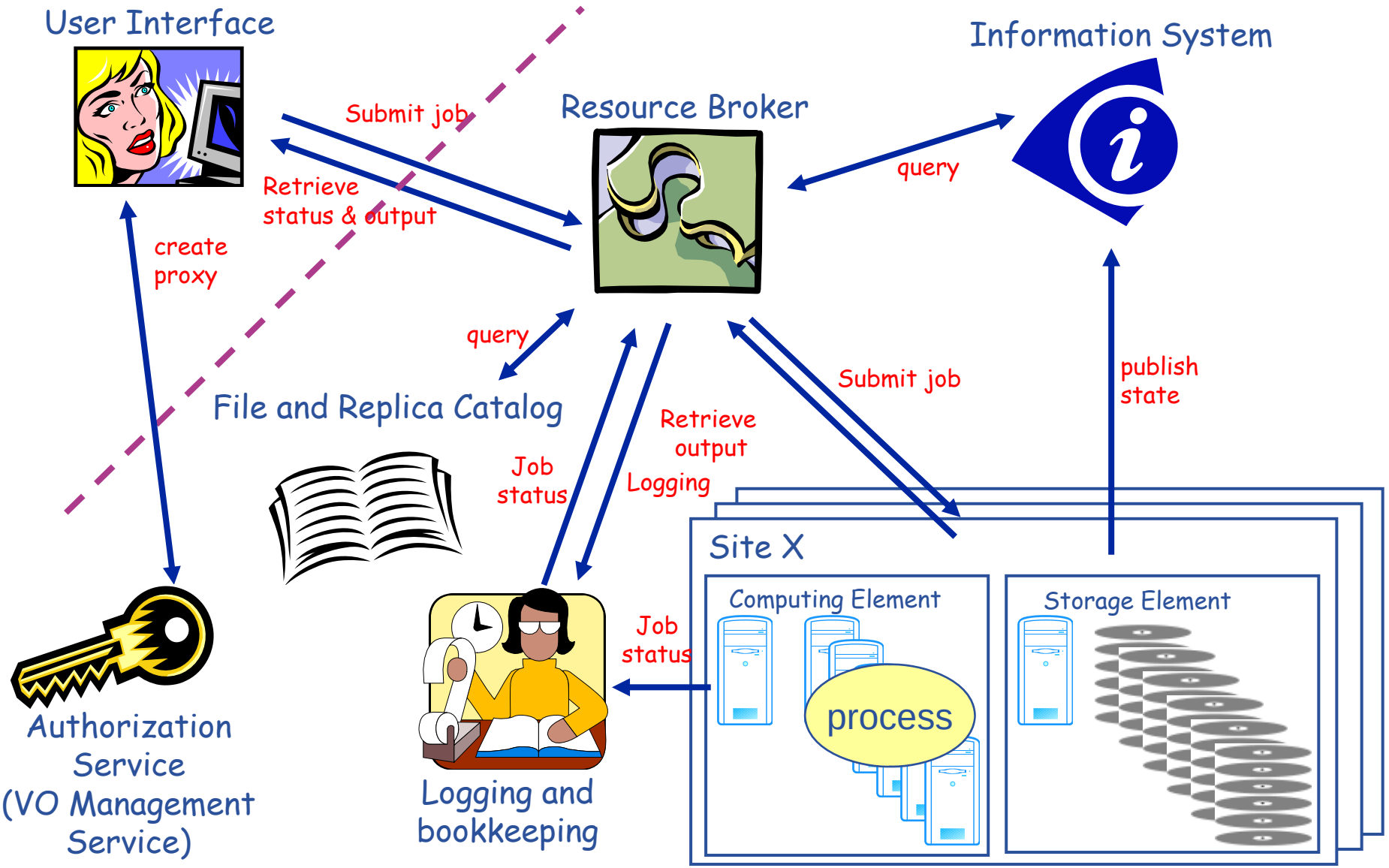
– **Exploit experience and existing components from VDT (Condor, Globus), EDG/LCG, and others**

– Develop a lightweight stack of generic middleware useful to EGEE applications (HEP and Biomedics are pilot applications).

– Focus is on providing a stable and basic infrastructure



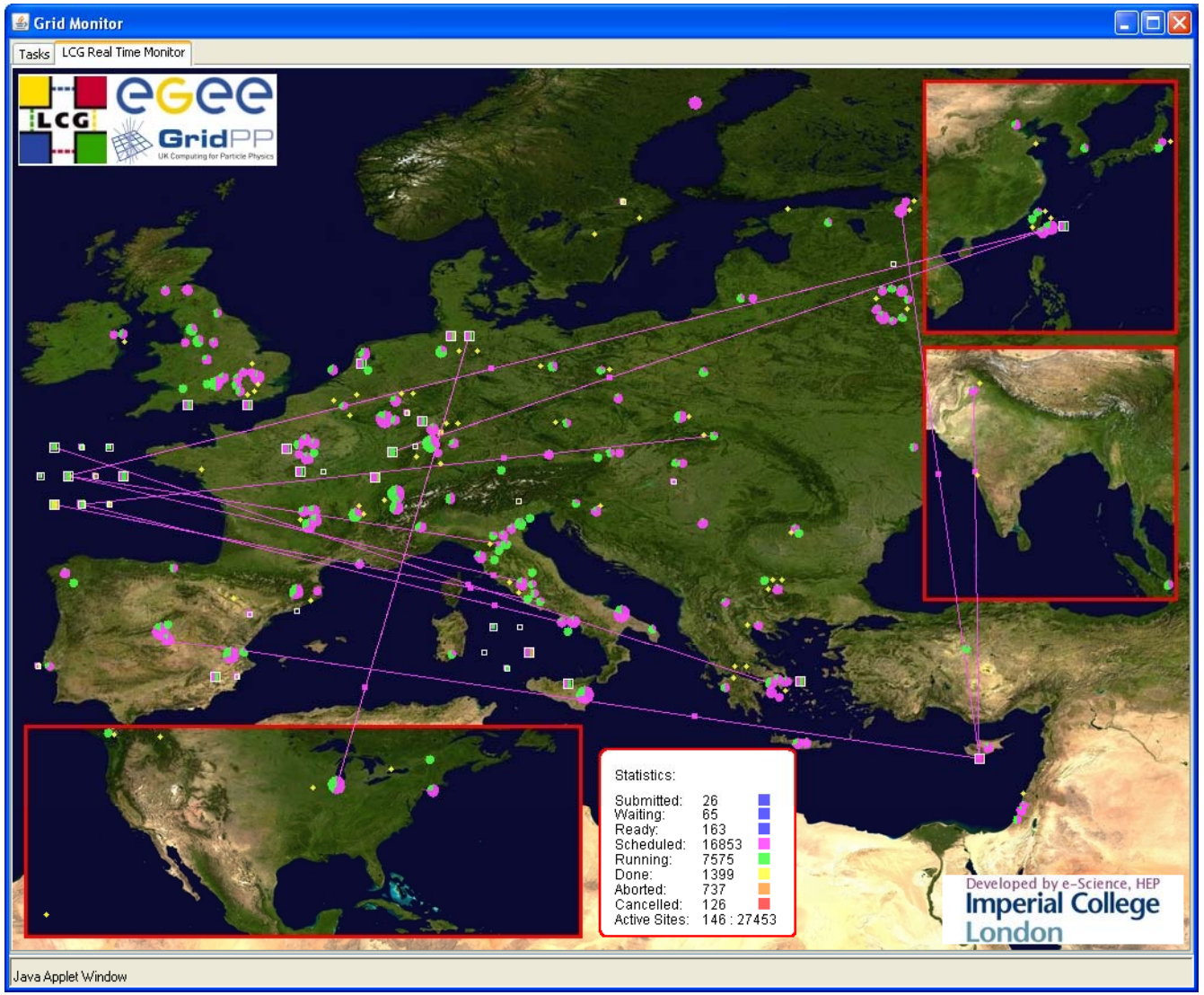
Basic services of gLite



What is happening now?

Real Time Monitor

- Java tool
- Displays jobs running (submitted through RBs)
- Shows jobs moving around world map in real time, along with changes in status



http://gridportal.hep.ph.ic.ac.uk/rtm/

(snapshot 16 January 2007)

Main components

User Interface (UI): The place where users logon to the Grid

Resource Broker (RB) (Workload Management System (WMS):
Matches the user requirements with the available resources on the Grid

Information System: Characteristics and status of CE and SE

File and replica catalog: Location of grid files and grid file replicas

Logging and Bookkeeping (LB): Log information of jobs

Computing Element (CE): A batch queue on a site's computers where the user's job is executed

Storage Element (SE): provides (large-scale) storage for files

Main components

User Interface (UI): The place where users logon to the Grid

Resource Broker (RB) (Workload Management System (WMS): Matches the user requirements with the available resources on the Grid

Inf SE

Fil eplicas

Lo

All built upon authorisation, authentication, security

Computing Element (CE): A batch queue on ere the user's job is executed

Storage Element (SE): provides (large-scale) storage for files

gLite User Interface (UI)

- **Server where the user has a login**
- **gLite client programs are installed**
  - Command line clients
  - Programming APIs
- **Typical UI scenario**
  (UI is central for the VO)
  - Upload program to UI with SCP
  - Login to UI with SSH
  - Compile code
  - Write job description (JDL file)
  - Create proxy certificate
  - Submit job
  - Check job status, download result from grid to UI when DONE
  - Download result from UI with SCP
- **Typically high level environments hide gLite UI.**
  - P-GRADE Portal
  - GANGA
  - GridWay
  - …

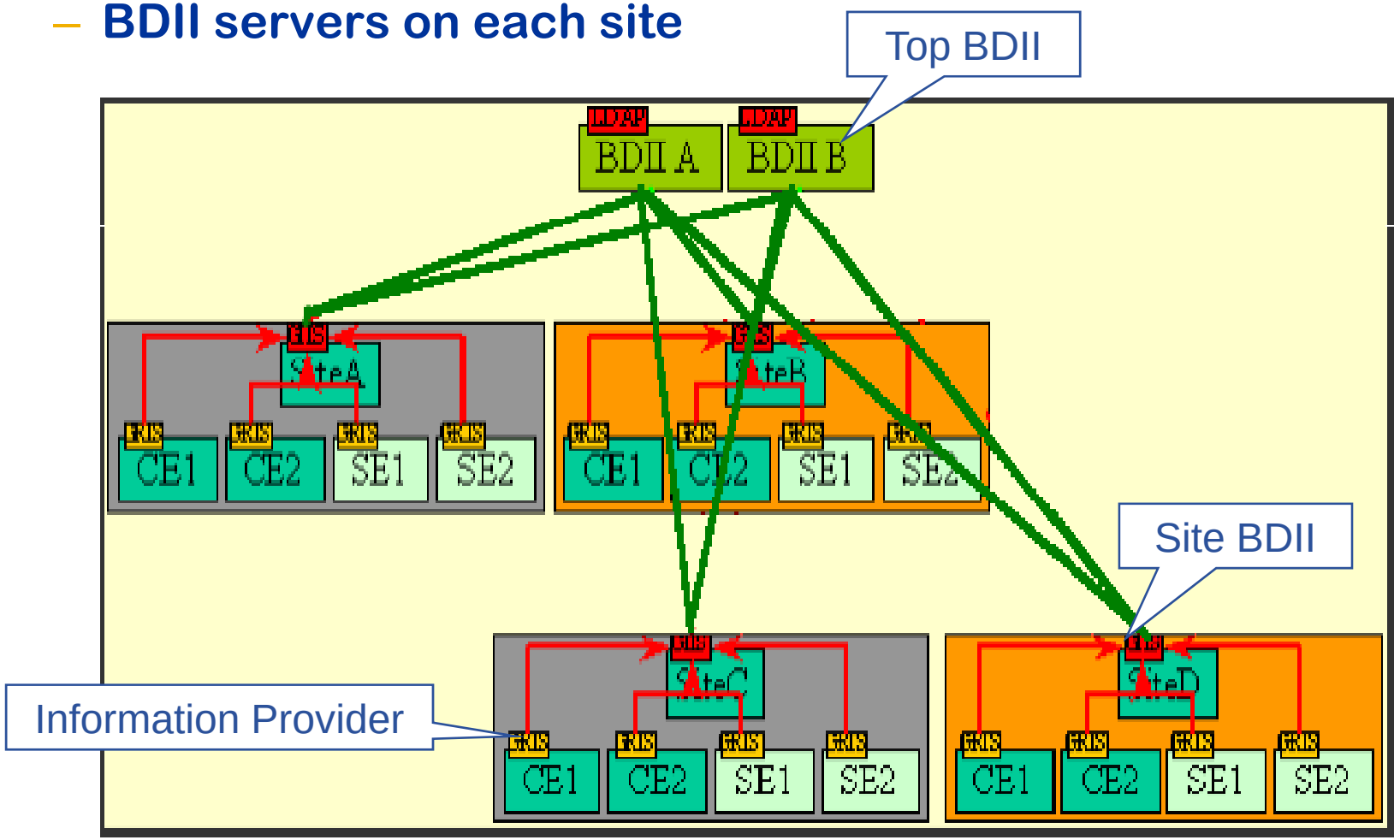
Resource Broker

- Official name: Workload Management System
- Key service in gLite
- Accepts Job Description Language files from User Interface, execute jobs on Computing Resources

- Detailed lecture later…

Information system (IS)

- **the user or a service (e.g. broker) can query**
  - **the top level BDII (usual mode)**
  - **BDII servers on each site**



Information system (IS)

- BDII server
  - LDAP server
  - Structures data as a tree
  - Tree model is defined by GLUE schema
  - Optimized for frequent queries
- Interacting with information system
  - Programming API
  - Command line tools (on UI)
    - lcg-infosites: simple, meets most needs
    - lcg-info: supports more complex queries
  - Portals (e.g. P-GRADE Portal)

Example – CE query

$ lcg-infosites --vo alice ce

#CPU	Free	Total Jobs	Running	Waiting	ComputingElement
14	0	0	0	0	ce002.ipp.acad.bg:2119/jobmanager-lcgpbs-alice
15	4	0	0	0	ce001.ipp.acad.bg:2119/blah-pbs-alice
80	8	0	0	0	ce02.grid.acad.bg:2119/jobmanager-pbs-alice
10	10	0	0	0	ce.hpc.iit.bme.hu:2119/blah-pbs-alice
96	94	0	0	0	grid109.kfki.hu:2119/jobmanager-lcgpbs-alice
3409	6	493	493	0	ce101.cern.ch:2119/jobmanager-lcglsf-grid_alice
3409	6	493	493	0	ce102.cern.ch:2119/jobmanager-lcglsf-grid_alice
3409	6	493	493	0	ce105.cern.ch:2119/jobmanager-lcglsf-grid_alice

[. . .]

Example – SE query

$ lcg-infosites --vo atlas se

Avail	Space(Kb)	Used Space(Kb)	Type SEs
39657488	106362948	n.a	se.phy.bg.ac.yu
31400000	18580000	n.a	se1.egee.man.poznan.pl
569586792	47148288	n.a	clrauvergridse01.in2p3.fr
1200000000	410000000	n.a	koala.unimelb.edu.au
22903032	42994124	n.a	se-lcg.sdg.ac.cn
457865076	663121389	n.a	atlasse01.ihep.ac.cn
29593756	80561288	n.a	se001.grid.bas.bg
931135488	41943040	n.a	se001.ipp.acad.bg

[. . .]

File and replica catalog (LFC)

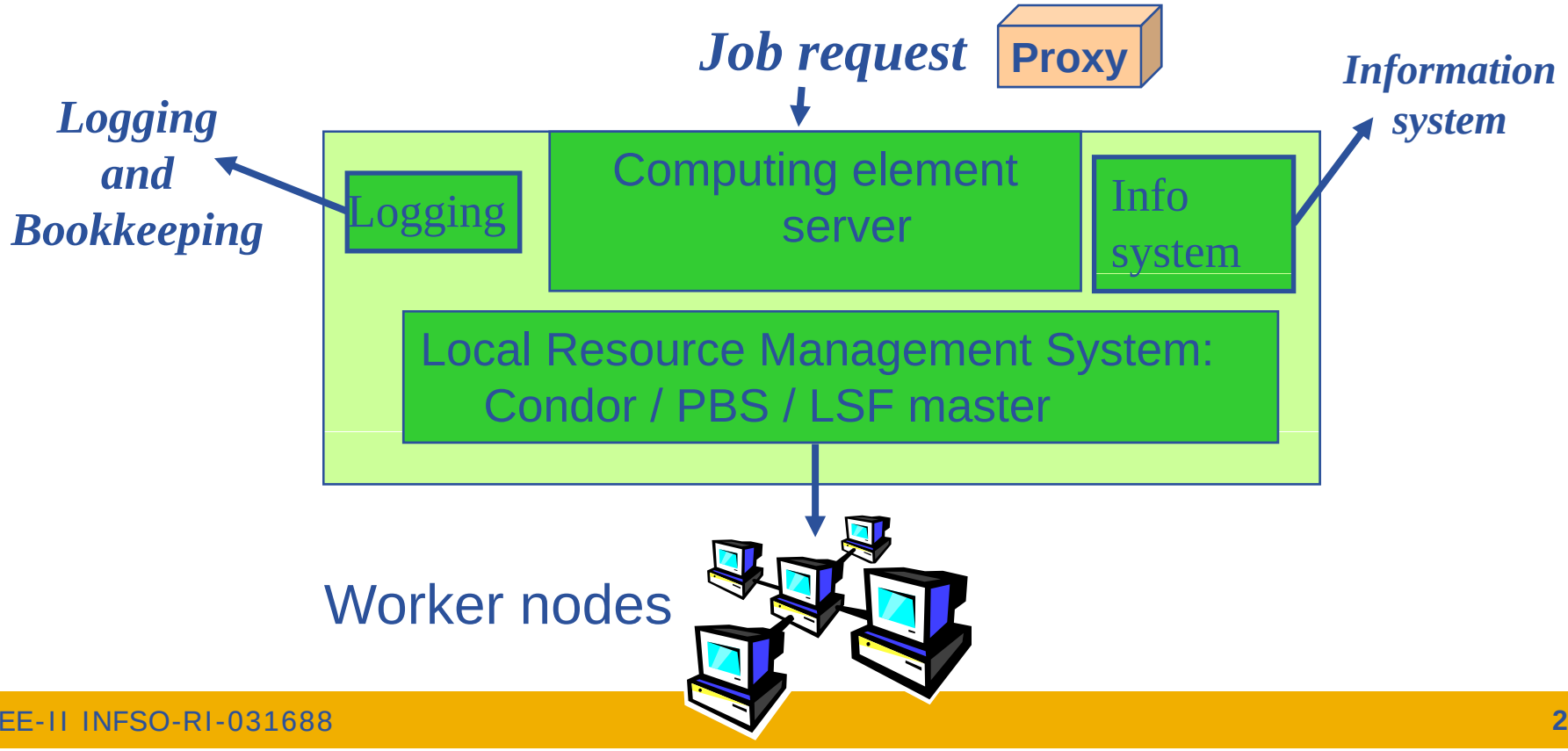
- **Users and their jobs refer to files by logical file name**
  - lfn:/grid/gilda/sipos/matrix_computation/input1
- **File catalog is used to map logical file name to physical file name(s)**
  - sfn://grid005.iucc.ac.il/storage/gilda/generated/2007-06-23/fileb233d43f-5bc6-4ede-a5fe-611d48be2ba5
- **LFC is a central database in the VO**

Logging and bookkeeping (LB)

- **Job history stored here**
  - When, what, where
- **Detailed information, not only a job status value**
- **LB is a central database in the VO**
- **Will not be used during the course**
  - Command line tools to query LB

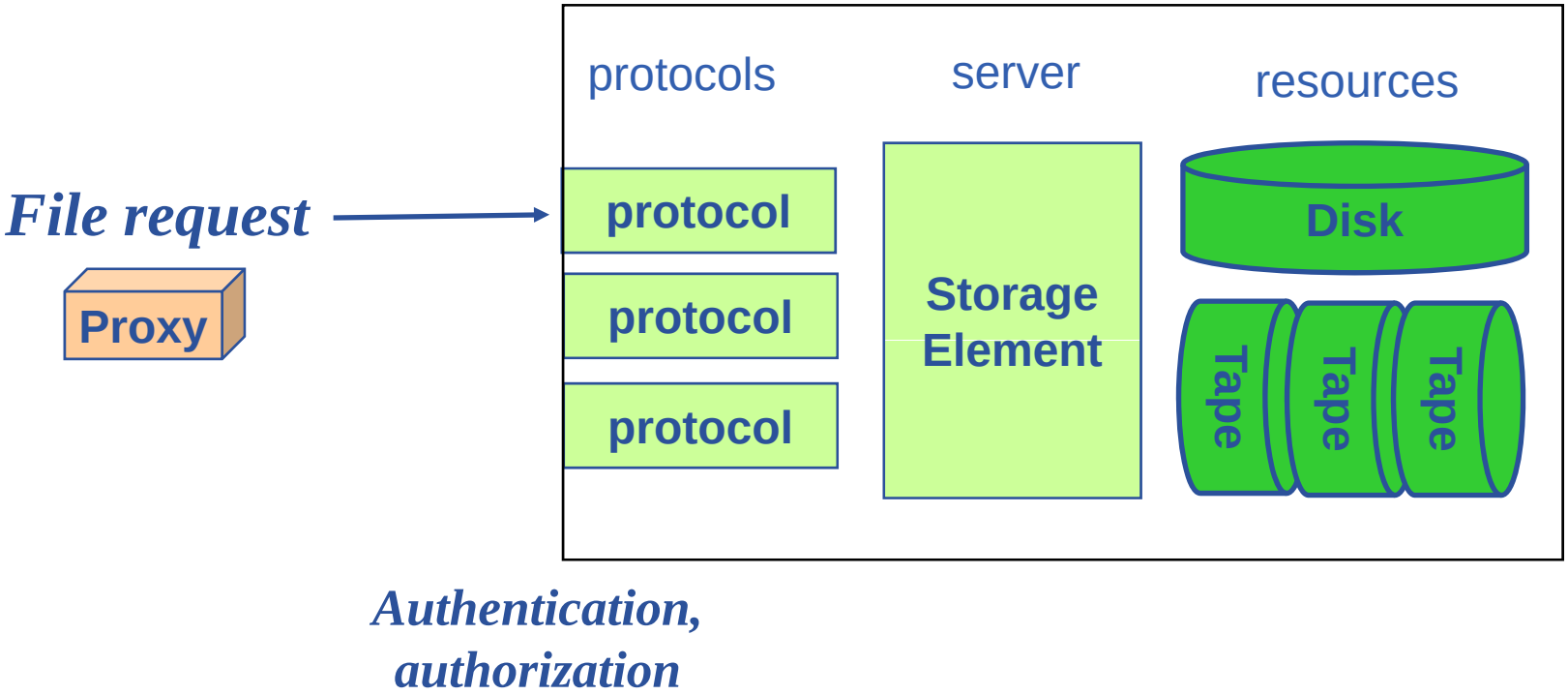
Computing Element (CE)

- **A “core” grid services**
  - One installation at every grid site
- **Expose computational facility - CPU**
- **Typically installed on a cluster**



Storage Element (SE)

- **A “core” grid services**
  - One installation at every grid site
- **Expose storage facility – Hard disk / tape**



Who provides the resources?!

Service	Provider	Note
User interface	User / institute / VO	Computer with client SW
Resource Broker (WMS)	VOs - EGEE does not fund RBs	
Information System	Grid operations - EGEE funded effort	
File and replica catalog	VOs - EGEE does not fund catalogs	
Logging and Bookkeeping	VOs - EGEE does not fund LB servers	
Computing Element (CE)	VOs - EGEE does not fund CEs	VOs provide resources to match average need
Storage Element (SE)	VOs - EGEE does not fund SEs	VOs provide resources to match average need
External services	User / institute / VO	To extend the capabilities of the core infrastructure

Security

- **Grid Security Infrastructure (GSI) enables VOs**
- **gLite security extended GSI**
- **Two levels of grid security problems**
  - Network level:
    - Mutual authentication of endpoints
    - Encrypted messages
    - Non repudiation
    - Integrity (protection against 3rd party changes)
  - VO level:
    - Who can be the member of a VO, who cannot?
    - What a VO member is allowed to do?
    - How can a SW (e.g broker) act on your behalf?

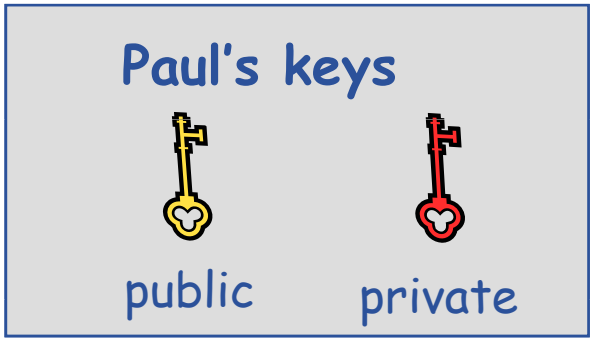
Basic concept of GSI

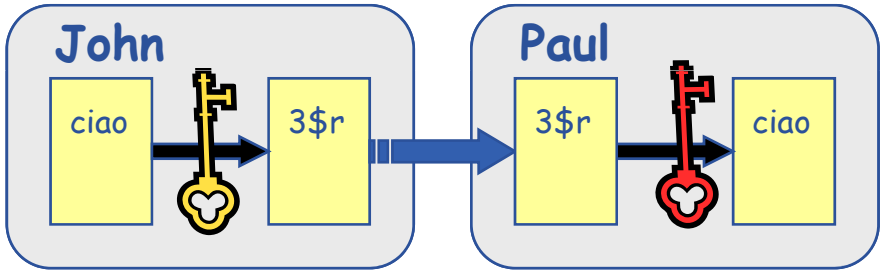
- **Asymmetric encryption…**

- **Every networked entity (user/machine/software) is assigned with two keys: one private key and one public key**
  - Private: only owner knows
  - Public: everybody else knows
- **Communication concept (simplified version):**
  1. Public keys are exchanged
  2. The sender encrypts message using
     - receiver's public key and
     - Sender's own private key
  3. The reciever decrypts using
     - Receiver's own private key
     - Sender's public key

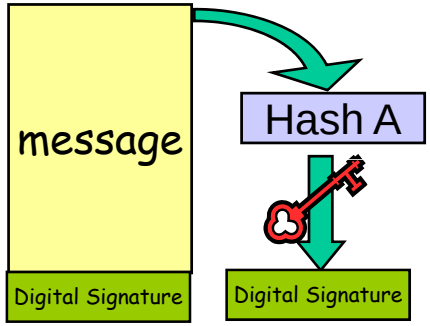
PKI in action

- **Encription**
  - Encription with recipient's public key
  - Only recipient can decript the message

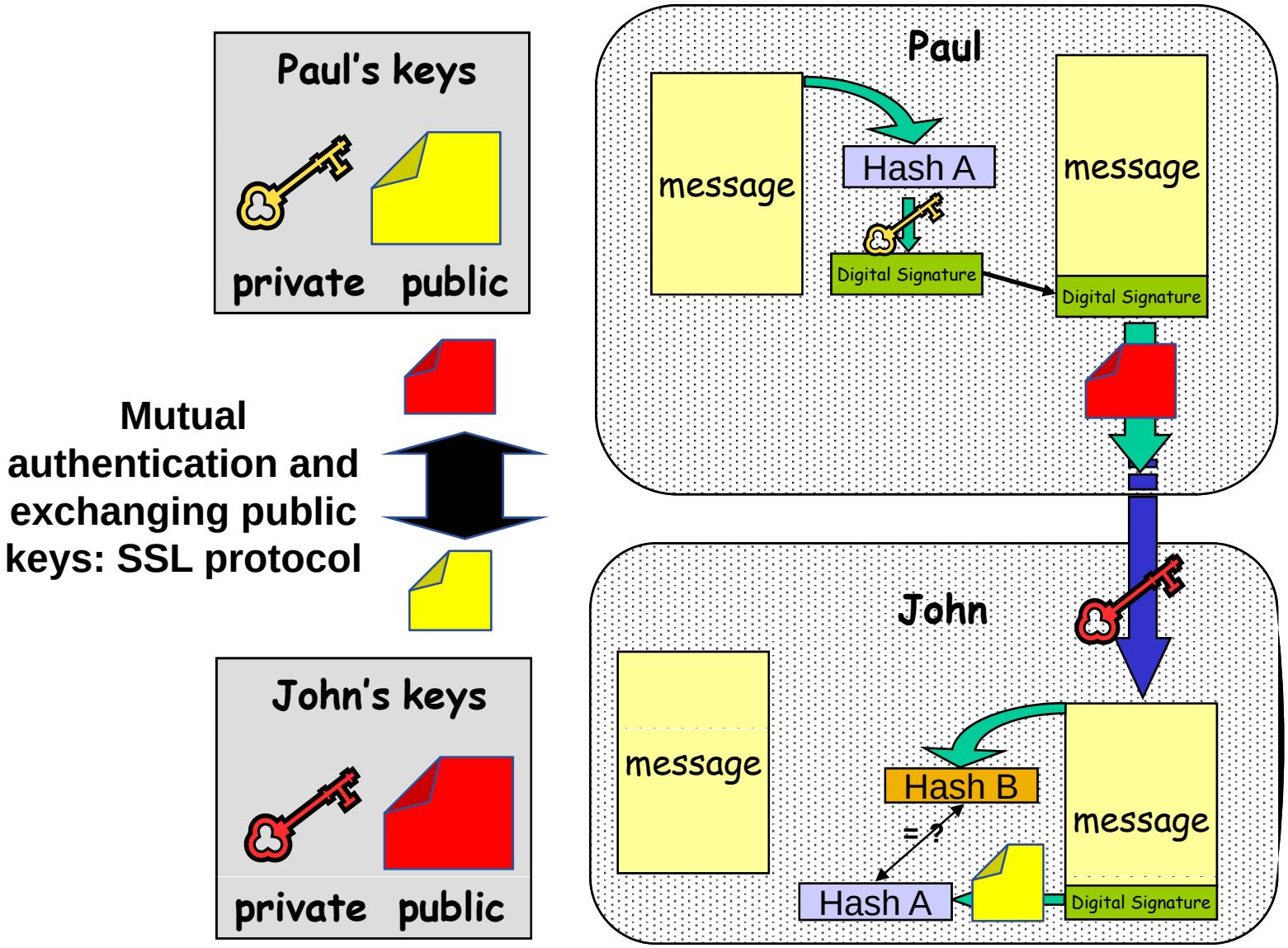




- **Non-repudiation**
  - **Naiv approach:** encript message with sender's private key
    - Too costly for long messages
  - **Solution:**
    - generate hash of the message
    - Encript hash with sender's private key
    - Attach encripted hash to message → **Digital signature**
  - Additional benefit: Integrity (hash is constant)



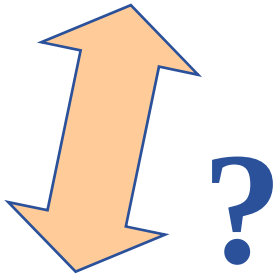
PKI in action – the big picture



Entity identity

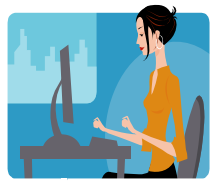
- Since I'm the only one with access to my private key, you know I signed the data associated with it
- But, how do you know that you have my correct public key?



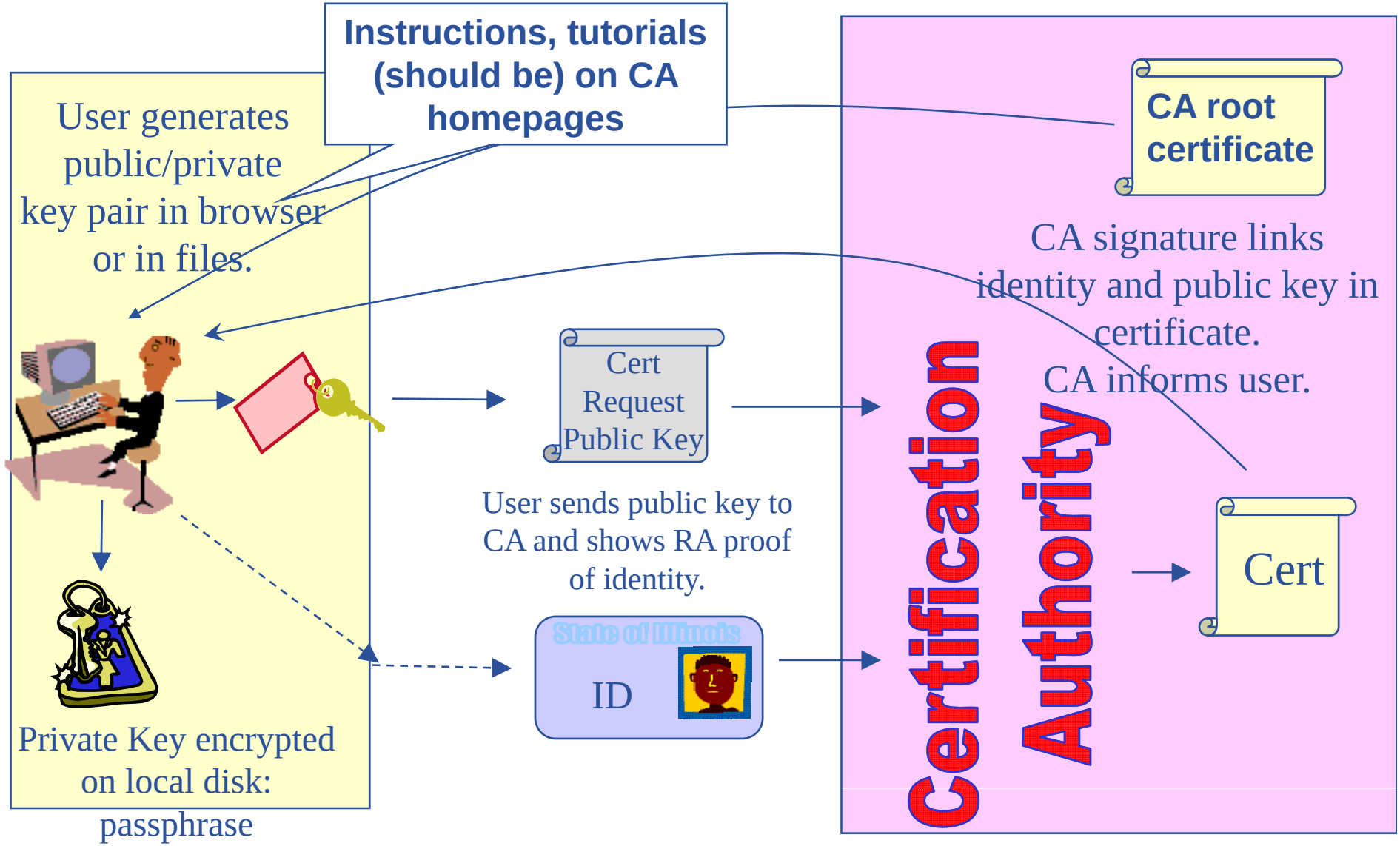








Your public and private keys

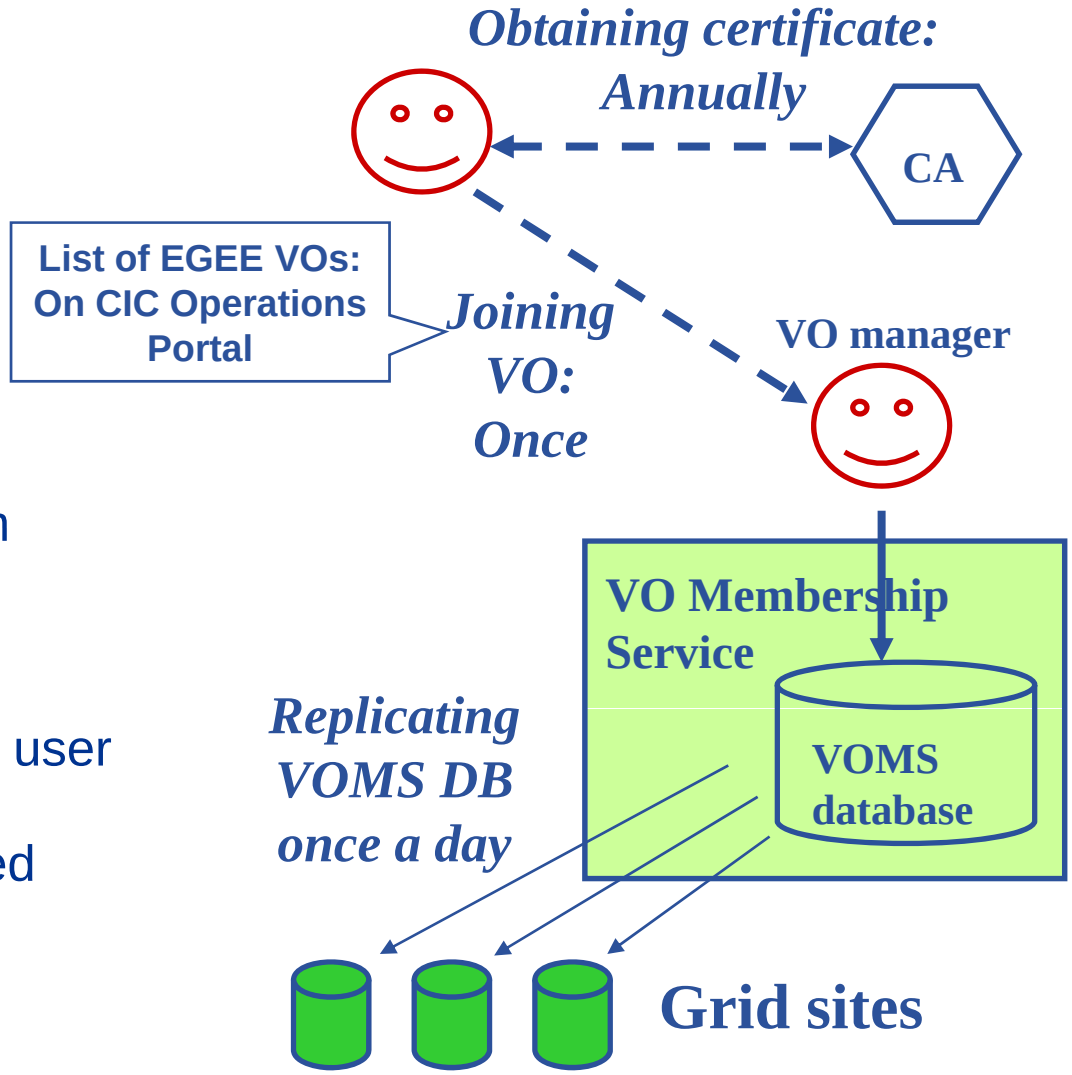


Your certificate and private key

- **Do not loan your certificate and private key to anyone!**
  - **Report to CA if your files were compromised → Certificate revocation list**
- **Where to store them?**
  - Store them in your browser
  - Store them in a file sysytem you trust
    - Different file formats (PEM, P12, …)
  - Store them on a USB key
  - Store them in MyProxy server
    - Obtain short term certificate just before grid interaction
- **Every Grid which recognizes your CA will trust your certificate**
- **CAs recognized by EGEE: http://www.gridpma.org**

Grid security at VO level

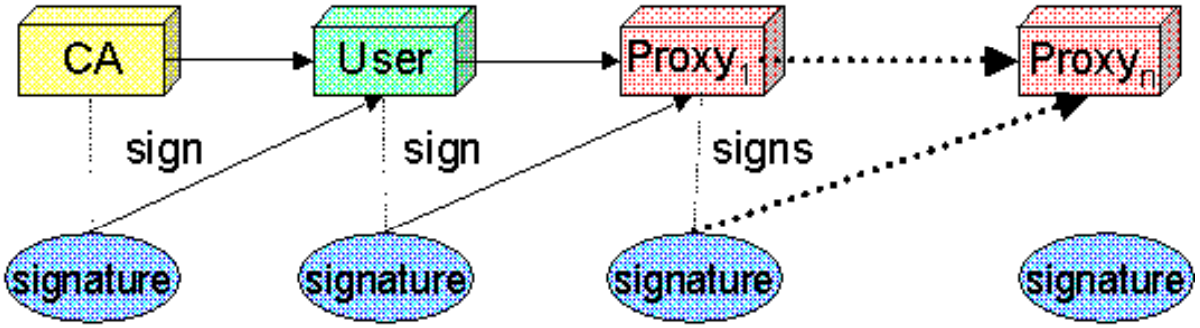
- **Users (and machines) are identified by certificates.**
- **VO Management Service: tool for VO level security**
- **Steps**
  - User obtains certificate from Certification Authority
  - User registers at the VO
    - usually via a web form
  - VO manager authorizes the user
    - VO DB updated
  - User information is replicated onto VO resources typically within 24 hours



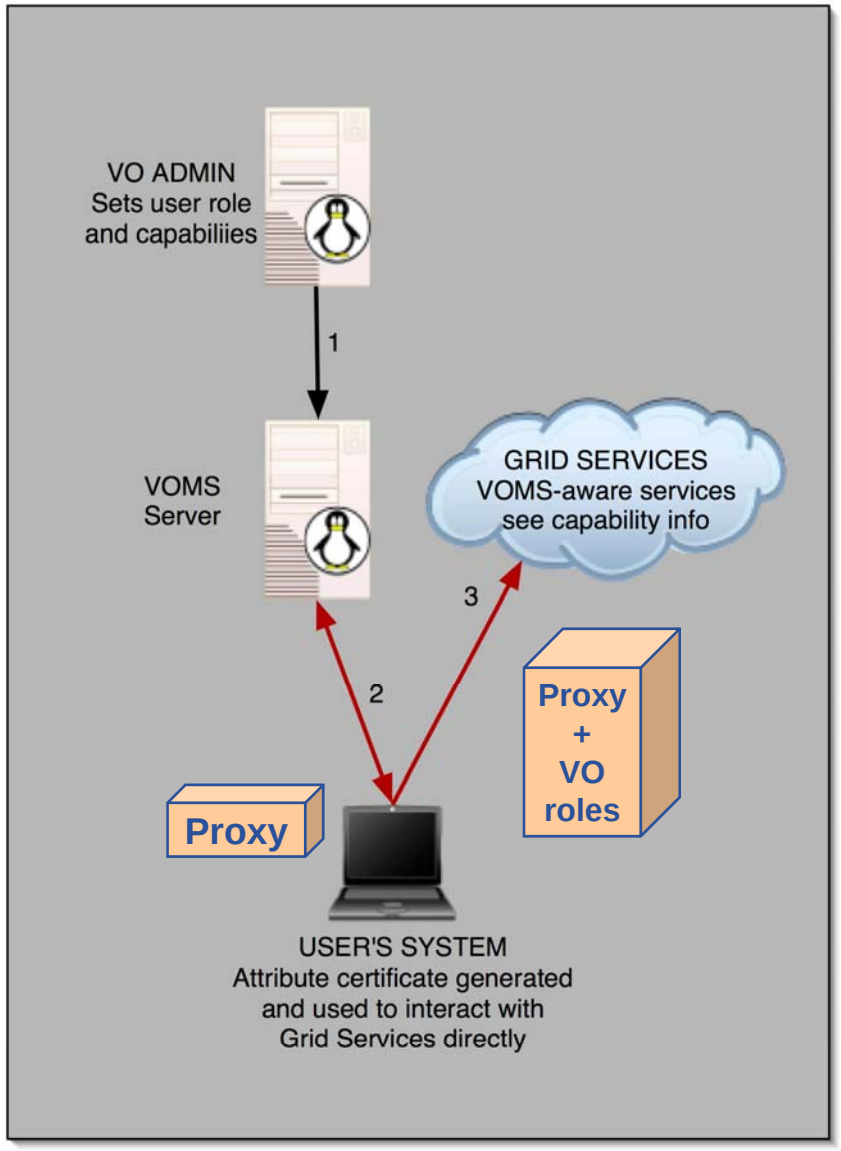
User's identity in the Grid = Subject of grid certificate:
/C=HU/O=NIIF CA/OU=GRID/OU=NIIF/CN=Gergely Sipos/Email=sipos@sztaki.hu

Proxy certificates

- **Delegation - allows remote process and services to authenticate on behalf of the user**
  - Remote process/service **"impersonates"** the user
- **Achieved by creation of next-level private key–certificate pair from the user's private key–certificate.**
  - New key-pair is a single file: Proxy credential
  - Proxy private key is not protected by password
  - Proxy may be valid for limited operations
  - Proxy has limited lifetime
- **The client can delegate proxies to services, processes**
  - Each service decides whether it accepts proxies for authentication



Proxy in action



- **Login to the grid:**
  Generate a proxy certificate

Command line tools on a User Interface:

voms-proxy-init –voms gilda

From a portal:

See P-GRADE Portal tomorrow

External services to gLite

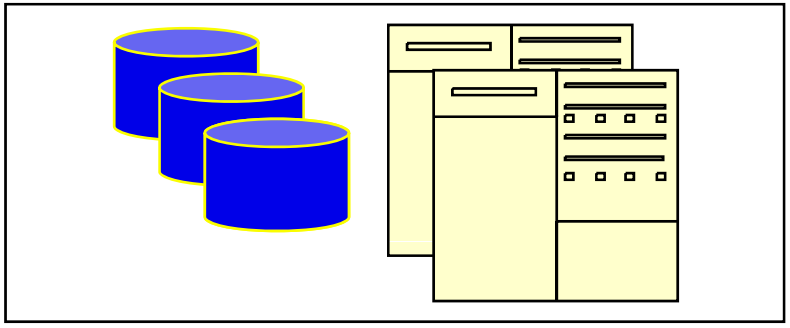
Application

Application toolkits

Command line & APIs

Higher-level gLite services (WMS,…)

Basic gLite services: CE, SE, info, security



- **Where computer science meets the application communities!**
- **The tools, services used by the VO's applications**
- **NA4 Recommended External Software Packages for Egee CommuniTies**
  - Current RESPECT tools:
    - GridWay
    - P-GRADE Portal
  - http://egeena4.lal.in2p3.fr/ → "Grid software" menu

Production infrastructure contains these services

- Basic services: Must be complete and robust; Should not assume the use of Higher-Level Grid Services
- High level services: help the users building their computing infrastructure but should not be mandatory

Further information, references

- EGEE
  - http://www.eu-egee.org/
- gLite middleware
  - http://www.glite.org
- gLite manuals, documentation
  - http://glite.web.cern.ch/glite/documentation/ (gLite user guide)
- Recommended External Software Packages for Egee CommuniTies (RESPECT)
  - http://egeena4.lal.in2p3.fr/