

Grid Computing

Back to basics: the concept, the model,
the infrastructure...

EGEE & Int.EU.Grid Tutorial
Porto, 15th May 2008

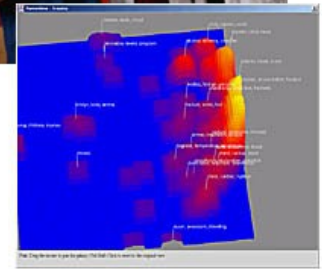
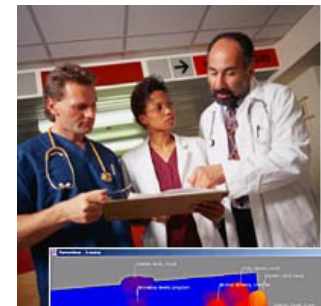
Gonalo Borges, Mrio David, Jorge Gomes
LIP Lisboa

1. Concepts and Definitions

- **GRID** computing is a recent concept which takes distributing computing a step forward
- The name **GRID** is chosen by analogy with the **electric power grid**:
 - Transparent: plug-in to obtain computing power without worrying where it comes from
 - Permanent and available everywhere
- The **World Wide Web** provides seamless access to **information** that is stored in many millions of different geographical locations
- In contrast, the **GRID** is a new computing infrastructure which provides seamless access to **computing power** and **data storage** distributed all over the globe



- Single institutions are no longer able to support the computing power and storage capacity needed for modern scientific research
- **Compute intensive sciences** which are presently driving the GRID development:
 - **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



■ Distributed infrastructures already exist, but...

- they normally tend to be **local & specialized systems**:
 - Intended for a single purpose or user group
 - Restricted to a limit number of users
 - Do not allow coherent interactions with resources from other institutions

■ The GRID goes further and takes into account:

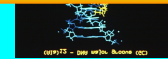
- **Different kinds of resources**:
 - Not always the same hardware, data, applications and admin. policies
- **Different kinds of interactions**:
 - User groups or applications want to interact with Grids in different ways
- **Access computing power / storage capacity across different administrative domains by an *unlimited* set of non-local users**
- **Dynamic nature**:
 - Resources added/removed/changed frequently
- **World wide dimension**

Researchers want to perform their activities regardless their geographical location...

For that, they need to interact with colleagues and share/access data and information distributed all over the globe



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



The GRID: networked data processing centres and "middleware" software as the "glue" of resources

Scientific instruments and experiments provide huge amount of data stored in different geographical places and which can only be processed by large computer resources facilities

Network vs Computer Performance

- Computer speed doubles every 18 months
- Network speed doubles every 9 months

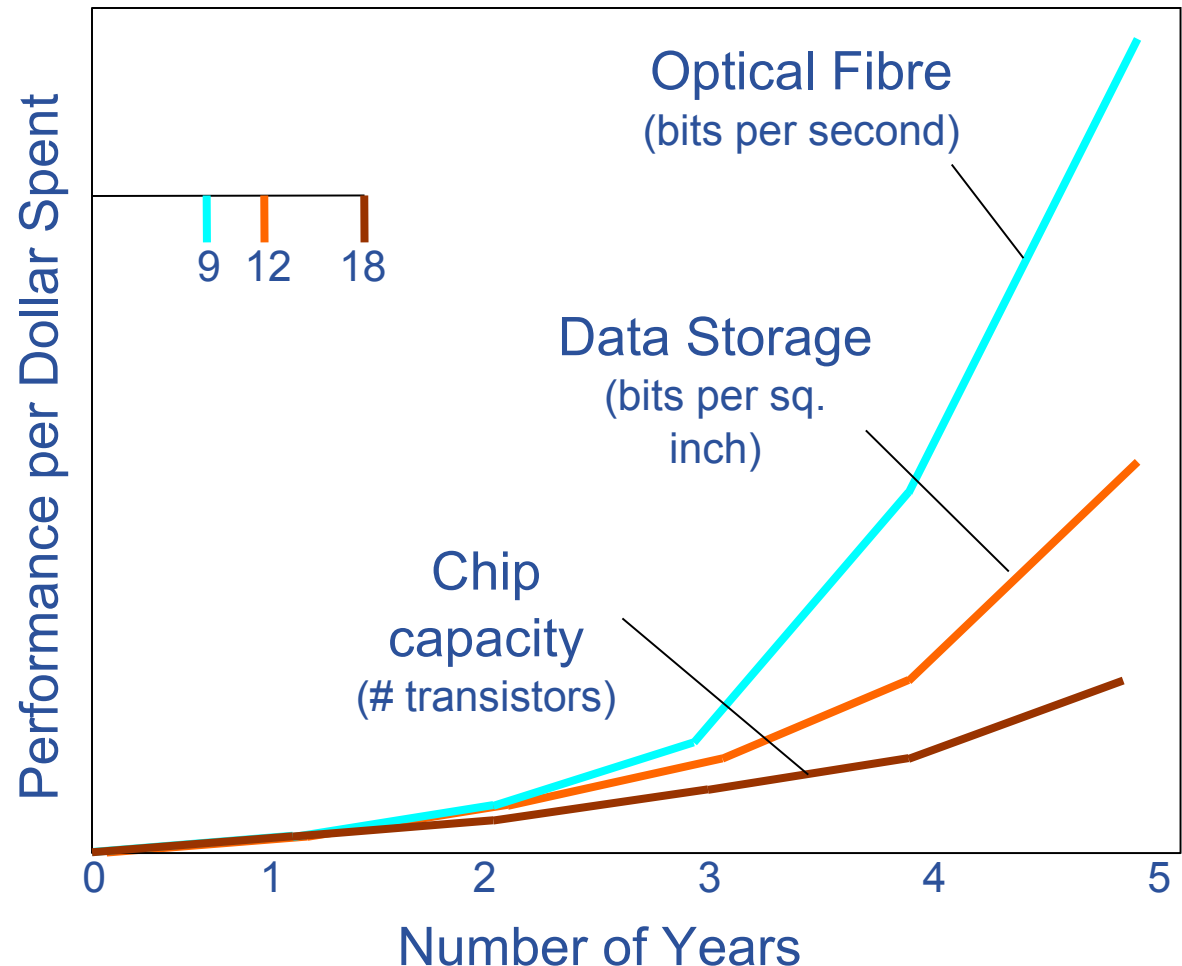
1986 to 2000

- Computers: x 500
- Networks: x 340,000

2001 to 2010

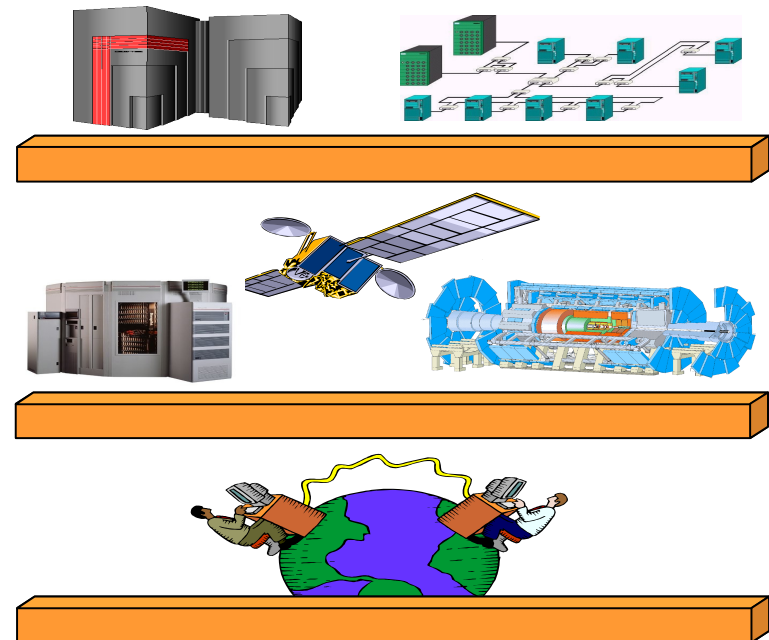
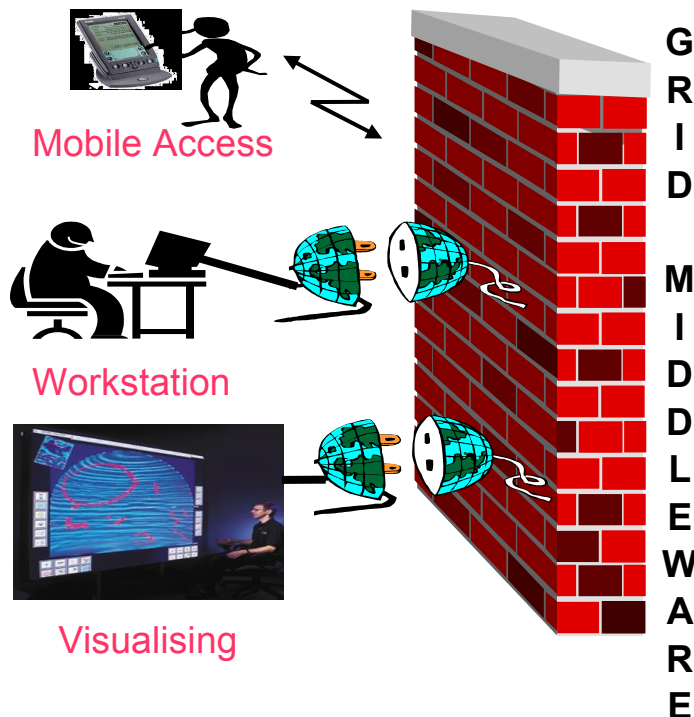
- Computers: x 60
- Networks: x 4000

Triumph of Light – *Scientific American*. George Stix, January 2001



The **transparent interaction between heterogeneous resources** (owned by geographically spread organizations), applications and users is only possible through...

- the use of a specialized layer of software called **middleware**



The **middleware** hides the infrastructure technical details and allows a **secure integration/sharing of resources**.

- Internet protocols do not provide security mechanisms for resource sharing.

■ **“GRID computing”** is an appealing term. Therefore, it's true meaning is frequently misused to describe:

- **Big computing clusters (in industry)**
- **Harvesting of unused compute cycles**
 - SETI@home (peer-to-peer computing)
 - Climateprediction.net



2. How to access to the GRID?

■ Users need

- single sign-on: the ability to logon to a machine and have the user's identity passed to other resources as required
- to trust owners of the resources they are using

■ Providers of resources (computers, databases,..) need to

- trust users they do not know
- minimise impact on security
- have the ability to trace who did what

■ The solution comes from

- Digital Certificates
- Virtual Organizations

- **Resource providers** are “opening themselves up” to itinerant users:
 - **A Secure Access** to resources is provided through the **X.509 Public Key Infrastructure**
 - Digital certificates identify uniquely its user/service identity
- **Users/Services identity** have to be certified by (mutually recognized) **national Certification Authorities (CAs)**
 - One CA per country
 - CAs are coordinated by global bodies to enable the creation of a world-wide trust zone (**EUGridPMA**)
- **Digital certificates** allow a temporary delegation from users to processes executed “in user’s name” (proxy certificates and myproxy certificates repositories)

■ Virtual Organizations (VO):

- People from different organizations but with common goals get together to solve their problems in a cooperative way.
 - **Virtualized shared computing resources:** VO members have access to computing resources outside their home institutions.
 - **Virtualized shared data resources:** VOs members can store and access data outside their home institutions.
 - **Other resources may be shared and virtualized as well:** Instruments, sensors, software and even people...

■ VO practical role:

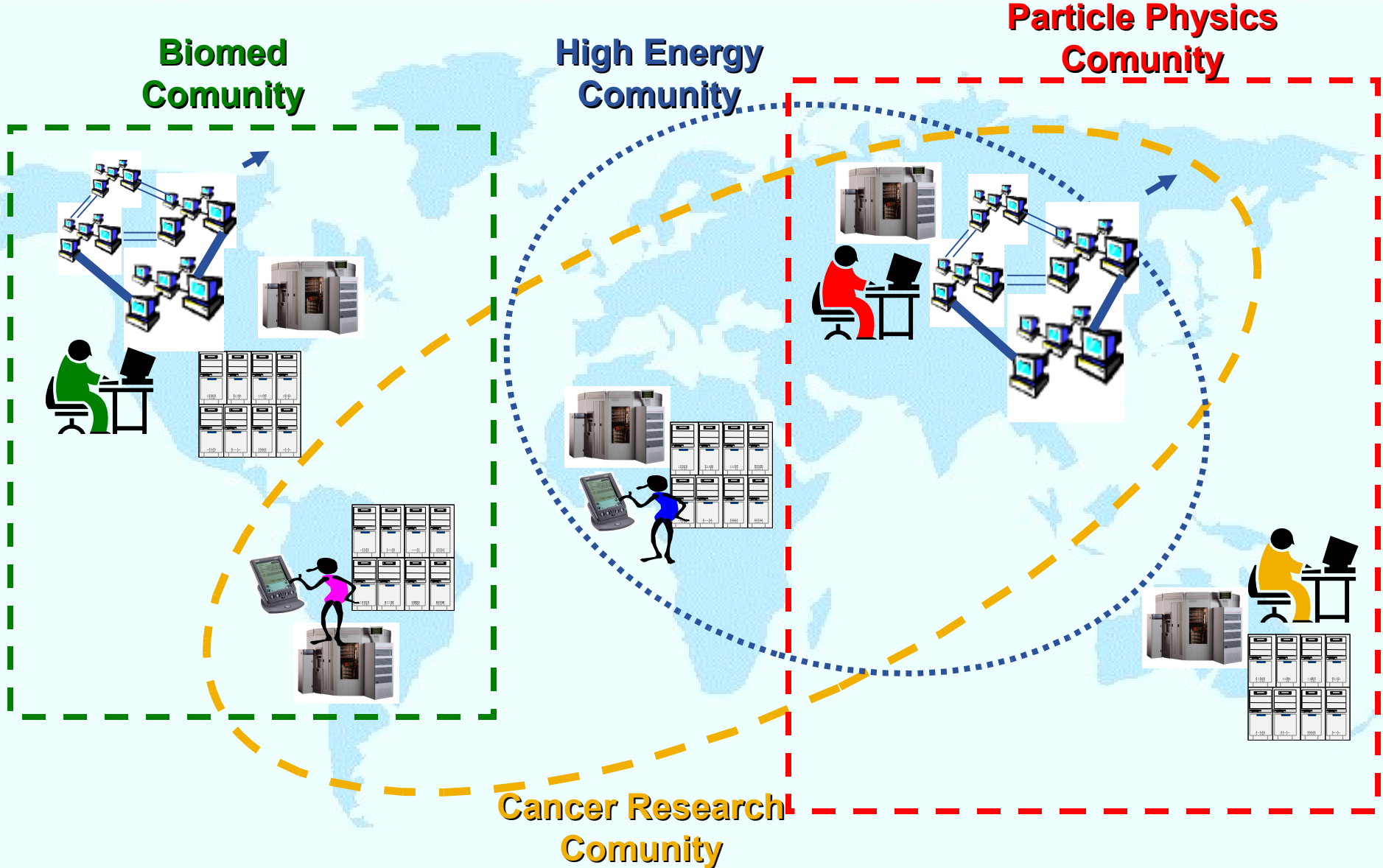
- **Set Common Agreed Policies for accessing resources**
- **Administrates the VO membership list**
 - Before joining a VO, a person must already have a valid certificate
 - VO administrators can reject persons which do not fulfill the VO requirements

Biomed Community

High Energy Community

Particle Physics Community

Cancer Research Community



3. LCG/EGEE projects

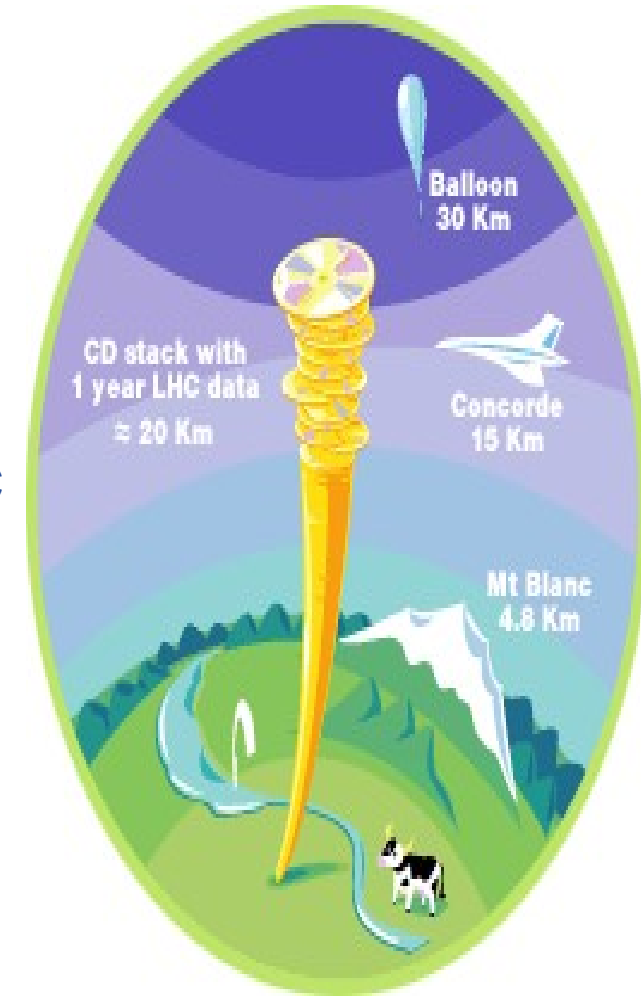
- Among the several projects in place, LIP is involved in:
 - **LHC Computing GRID (LCG)**
 - The biggest worldwide GRID infrastructure
 - Will be used in the data analysis produced by the LHC accelerator built at CERN, the european organization for nuclear research
 - **Enabling GRIDS for E-Science in Europe (EGEE)**
 - An European GRID project
 - The biggest worlwide GRID built for multi-disciplinary sciences



EGEE
Enabling Grids for
E-science in Europe

LCG aims to build/maintain a data storage and analysis infrastructure for the large LHC physics community.

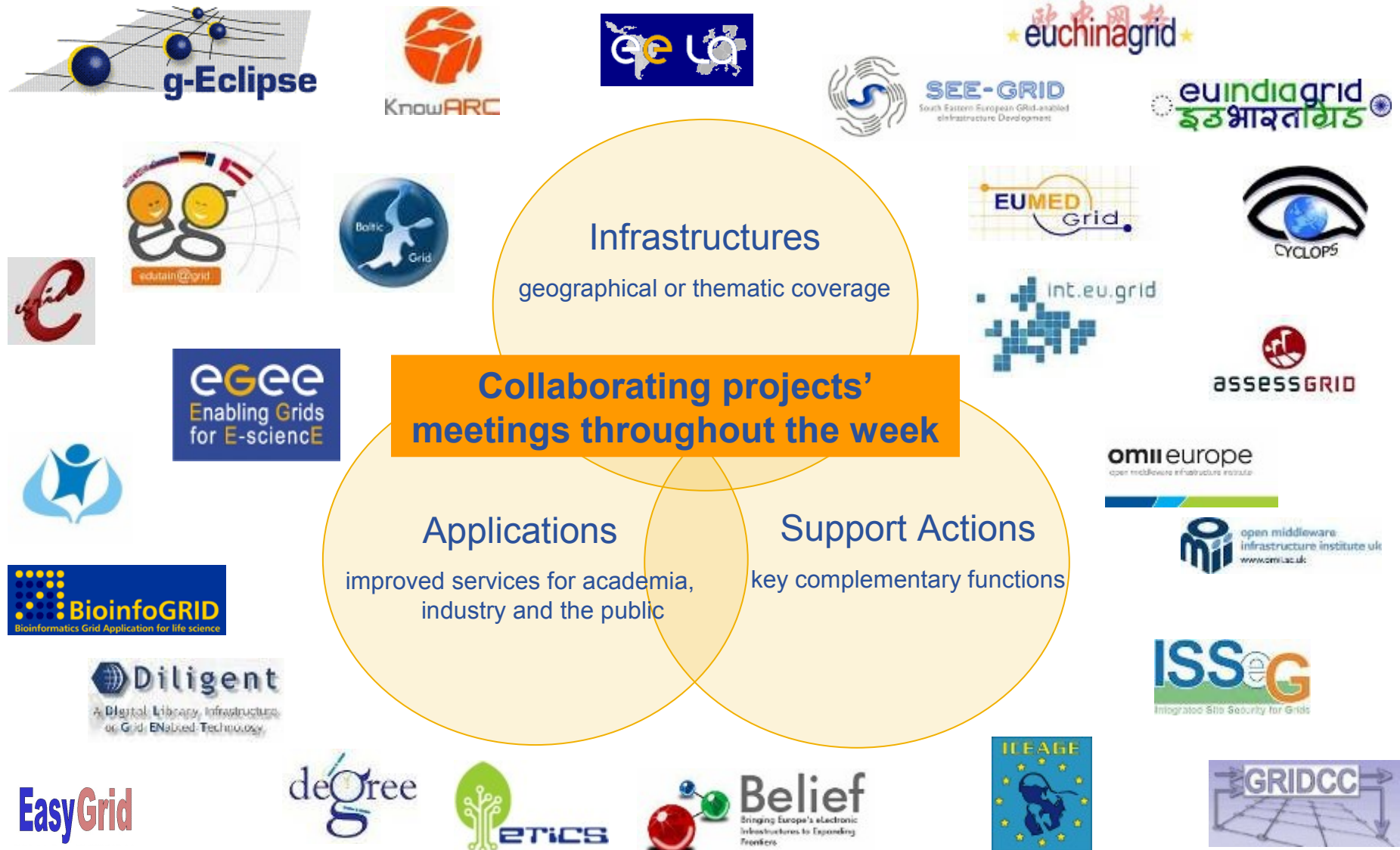
- **LHC experiments are expected to produce 10 Petabytes of experimental data annually**
 - 10^9 collisions/second (1 GHz) – 100Hz (after filtering)
 - 1 collision = 1MB of data
 - $100 \text{ MB/s} \Rightarrow 10 \text{ PB year} = 20 \text{ Km CD stack}$
- **Computing power ~100 000 “today's fastest” PC for analysis and simulation**
- **Needed to be available during the 15 years life time of the LHC machine**
- **Fully accessible to about 5000 scientists from more than 500 institutes around the world.**



LCG depends on two major science grid infrastructures



25 projects have registered as of Sept 2007: [web page](#)



240 sites
45 countries
41,000 CPUs
5 PetaBytes
>10,000 users
>150 VOs
>100,000 jobs/day

Archeology
Astronomy
Astrophysics
Civil Protection
Comp. Chemistry
Earth Sciences
Finance
Fusion
Geophysics
High Energy Physics
Life Sciences
Multimedia
Material Sciences

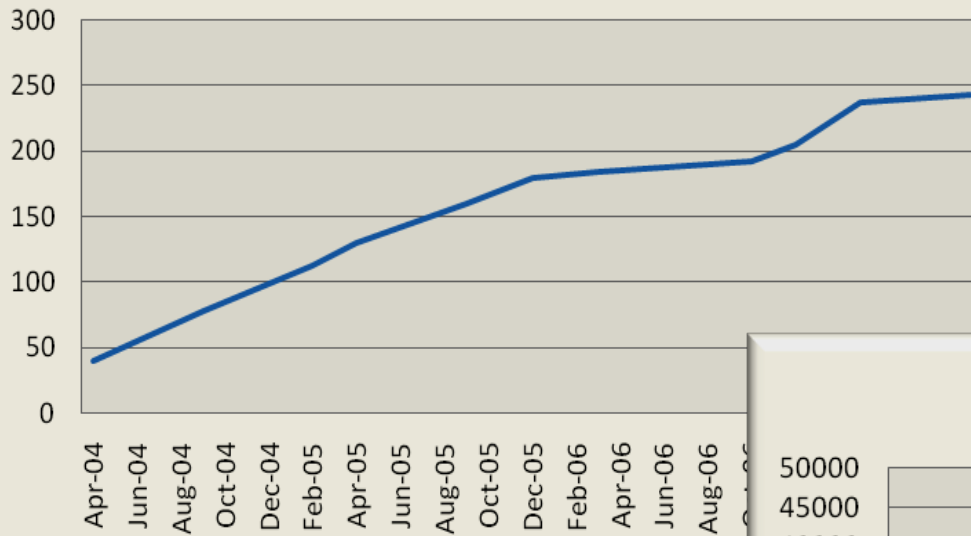
...

91 partners in 32 countries
25 collaborating projects



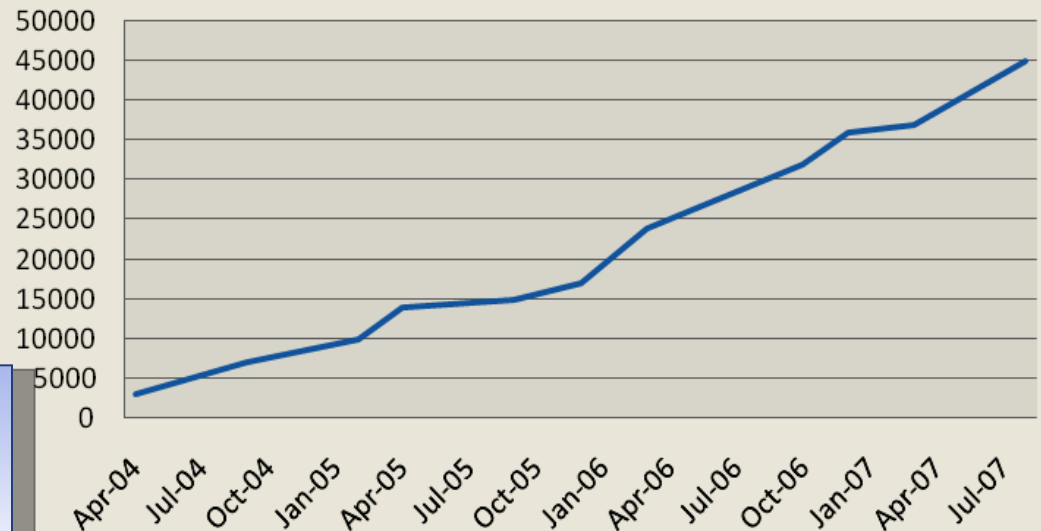
GridPP
UK Computing for Particle Physics

No. Sites

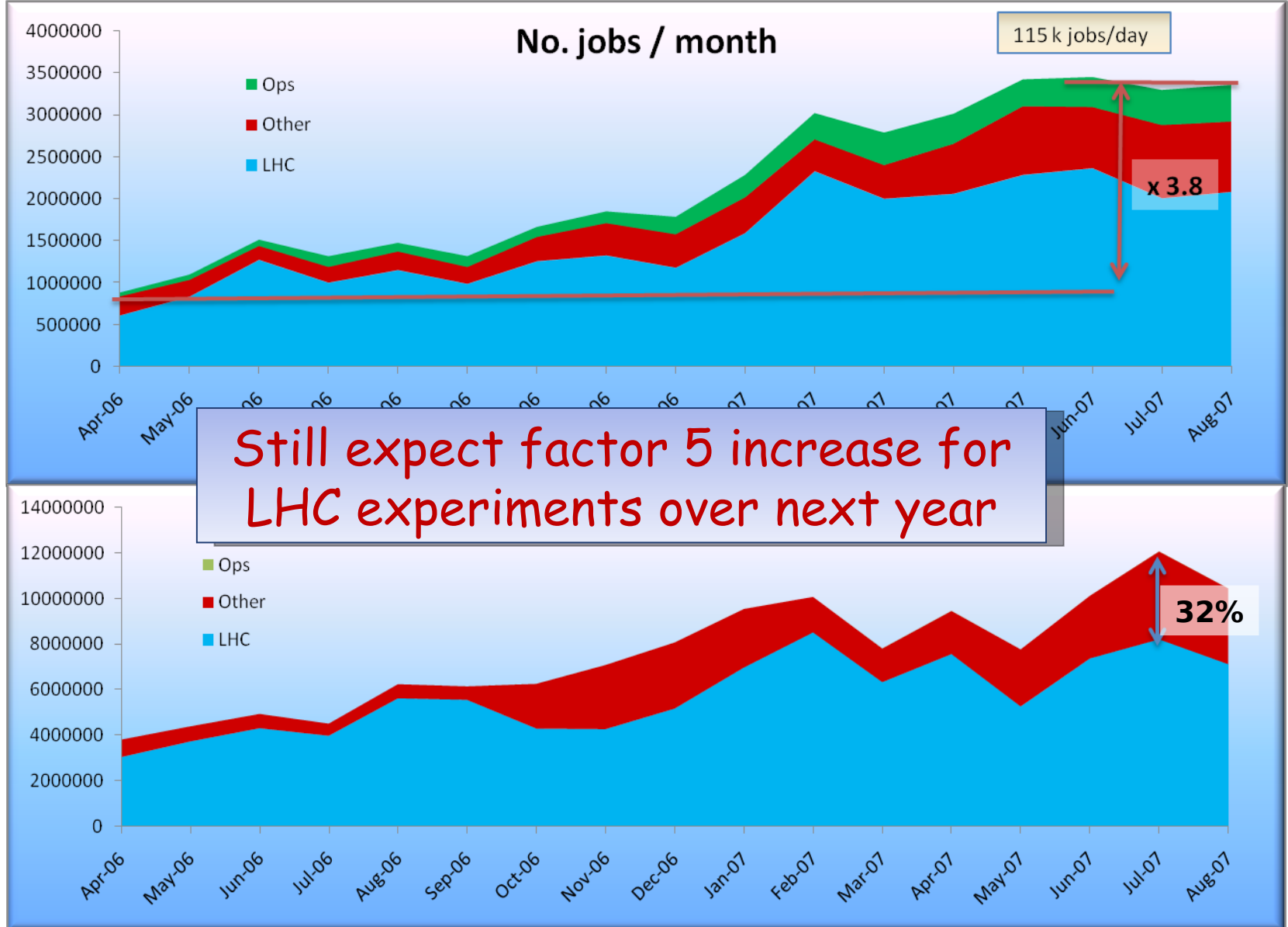


5 times more sites in
3 years.

No. CPU



9 times more CPUs in
3 years.





■ Portugal together with Spain form the EGEE Southwest federation

■ LIP coordinates EGEE in Portugal

■ Portuguese sites

○ LIP

- Lisbon (core services, production and pre-prod)
- Coimbra

○ Univ Lusiada

- Famalicão

○ Univ Porto

- Porto (3 clusters)

○ Univ Minho

- Braga

○ IEETA

- Aveiro (pre-prod)



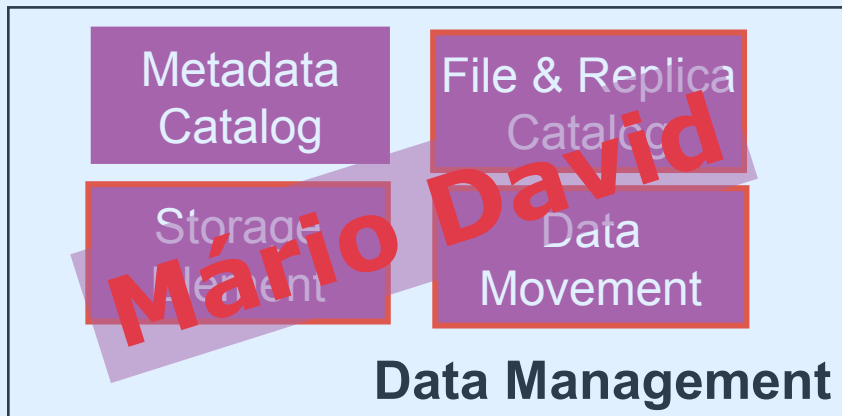
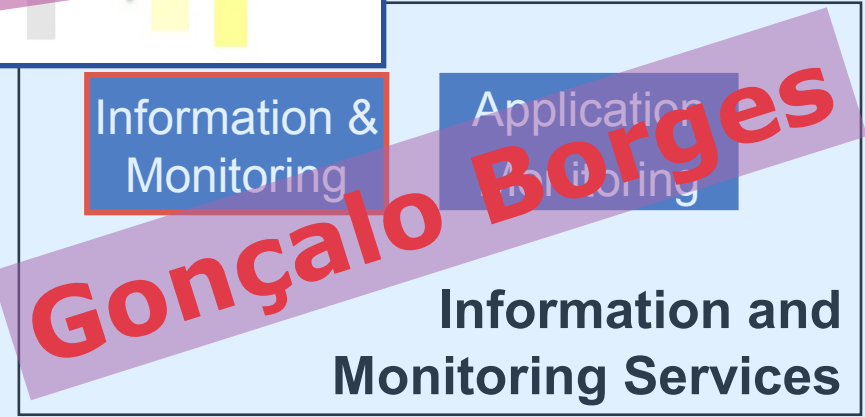
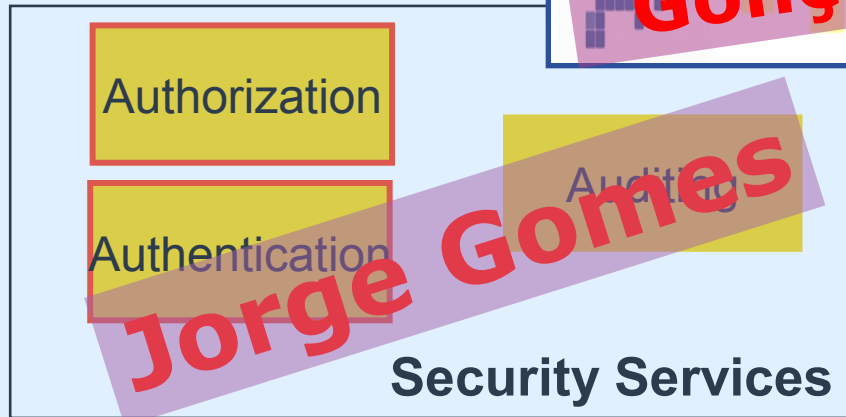


- Total number of CPUs in the federation > 1400
- Total number of CPUs in Portugal ~ 262

4. The Grid Infrastructure



Services used in
this Grid Track



Interactivity and Parallel Processing

■ Grid services are divided in two different sets:

○ **Local services:** deployed and maintained by each participating site

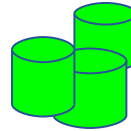
- **Computing Element (CE)**



- **Storage Element (SE)**

- **Monitoring Box (MonBox)**

- **User Interface (UI)**



○ **Core services:** central services installed only in some Resource Centres (RC's) but used by all users to allow interaction with the global infrastructure

- **Resource Broker (RB)**



- **Top-Berkeley-Database Information Index (BDII)**



- **File Catalogues (FC)**



- **Virtual Organization Membership Service (VOMS)**

- **MyProxy server (PX server).**

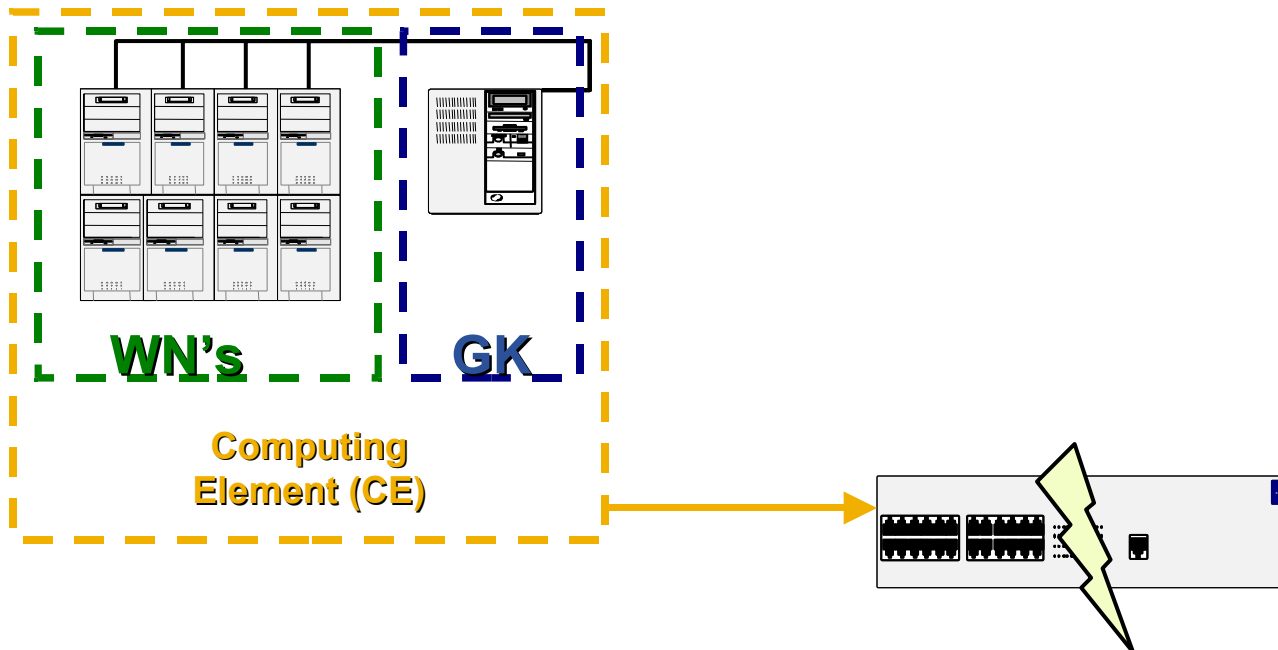
■ **Computing Element (CE)** is one of the key elements of the site.

○ **Gatekeeper service (GK)**

- Authentication and authorization;
- Interacts with the local batch system (PBS, LSF, Condor, SGE);
- Runs a local information system (GRIS) publishing information regarding local resources.

○ **Worker Nodes (WN)**

- Where jobs are really executed.



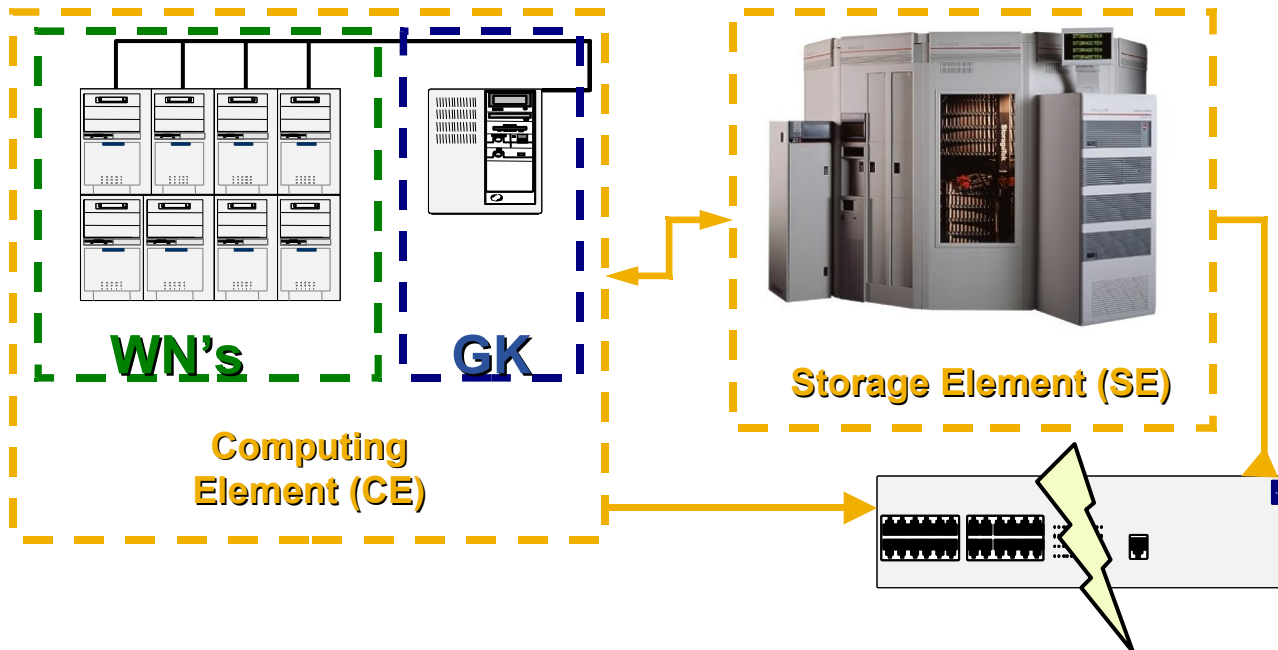
■ **The Storage Element (SE)** is the other key service.

○ **SRM - Storage Resource Management (DPM, dCache and CASTOR)**

- Provides an interface for the grid user to access the local storage system (disk pools, HSM with tape backend, etc...);
- Implements **gridFTP**, an extension of the ftp protocol adding GSI security.

○ **GFAL API protocol**

- Used on top of the SRM to provide POSIX-like access, enabling open/read/write operations in files currently stored in a site SE.

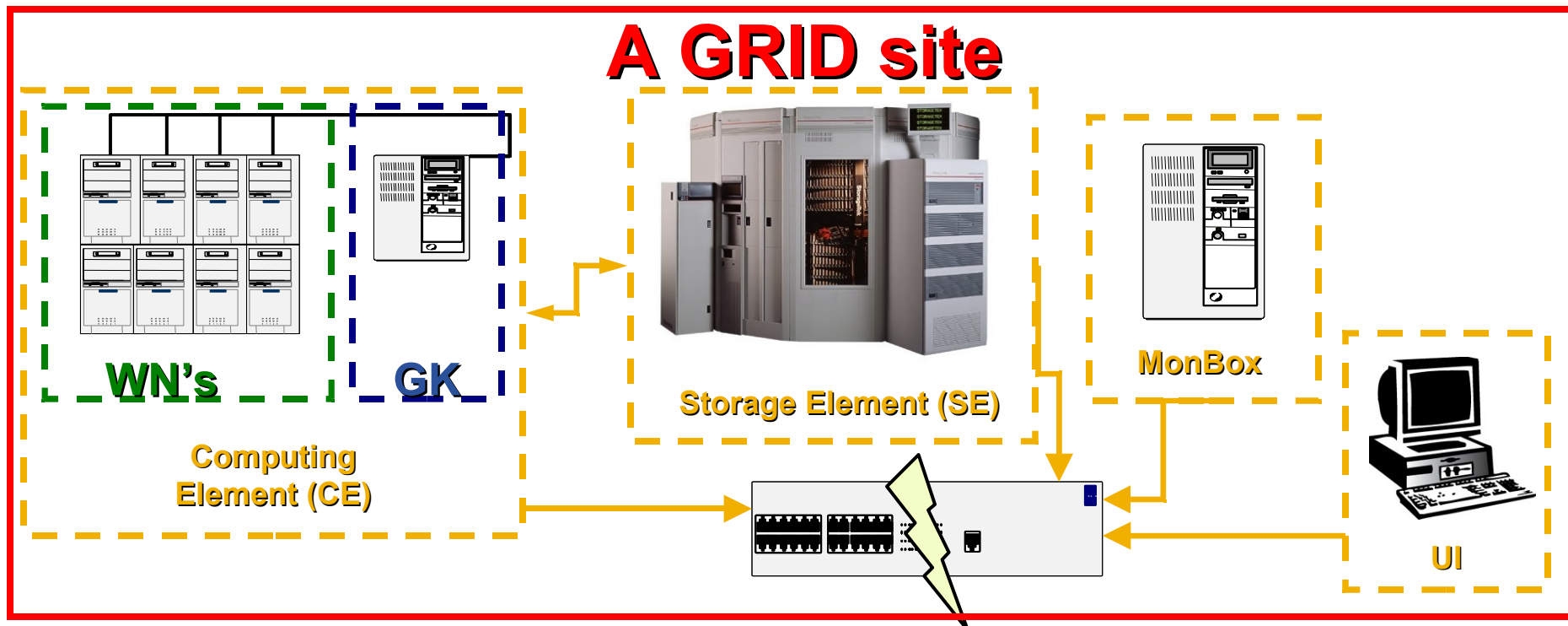


The Monitoring Box (MonBox) service

- Collects information given by sensors installed in the site machines.

User Interfaces (UI)

- Contain client middleware tools allowing the user to perform a large set of operations with grid resources (submission of jobs and storage and management of files).



■ The Resource Broker/WMS (RB/WMS)

○ Implements the **matchmaking** process

- MatchMake the user requirements to the available resources. The status of site resources is obtained querying the top-BDII
- Communicates with File Catalogues to determine in which sites files can be read or stored (if jobs require file manipulation)

○ Runs the **Logging and Bookkeeping service (LB)**

- Stores the status of all the jobs submitted to the RB in a database which can be queried by the user through UI client command tools.

■ The Resource Broker/WMS (RB/WMS)

○ Implements the matchmaking process

- MatchMake the user requirements to the available resources. The status of site resources is obtained querying the top-BDII
- Communicates with File Catalogues to determine in which sites files can be read or stored (if jobs require file manipulation)

○ Runs the Logging and Bookkeeping service (LB)

- Stores the status of all the jobs submitted to the RB in a database which can be queried by the user through UI client command tools.

■ The top-BDII

- Collects the information provided by the sites information systems (GIS) providing “on-line” information to other grid services.

■ The Resource Broker/WMS (RB/WMS)

○ Implements the matchmaking process

- MatchMake the user requirements to the available resources. The status of site resources is obtained querying the top-BDII
- Communicates with File Catalogues to determine in which sites files can be read or stored (if jobs require file manipulation)

○ Runs the Logging and Bookkeeping service (LB)

- Stores the status of all the jobs submitted to the RB in a database which can be queried by the user through UI client command tools.

■ The top-BDII

- Collects the information provided by the sites information systems (GIS) providing “on-line” information to other grid services.

■ Grid File Catalogues (FC's)

- Database services which map the Logical File Names (LFN's) attributed to files, their Global Unique Identifiers (GUID's) and their storage locations (SURL's)

Who provides the resources?!

<u>Service</u>	<u>Provider</u>	<u>Note</u>
<u><i>User interface</i></u>	Users / institutes / VOs	Computer with client software
<u><i>Resource Broker/WMS</i></u>	VOs / ROCs(?) EGEE does not funds RBs	
<u><i>Information System</i></u>	Grid operations: EGEE funded effort	
<u><i>File and replica catalog</i></u>	VOs / ROCs (?)	
<u><i>Computing Element</i></u>	VOs / Institutes	Scalability requires that VOs provide resources to match average need
<u><i>Storage Element</i></u>	VOs / Institutes	
<u><i>External services</i></u>	User / institute / VO	To extend the capabilities of the core infrastructure

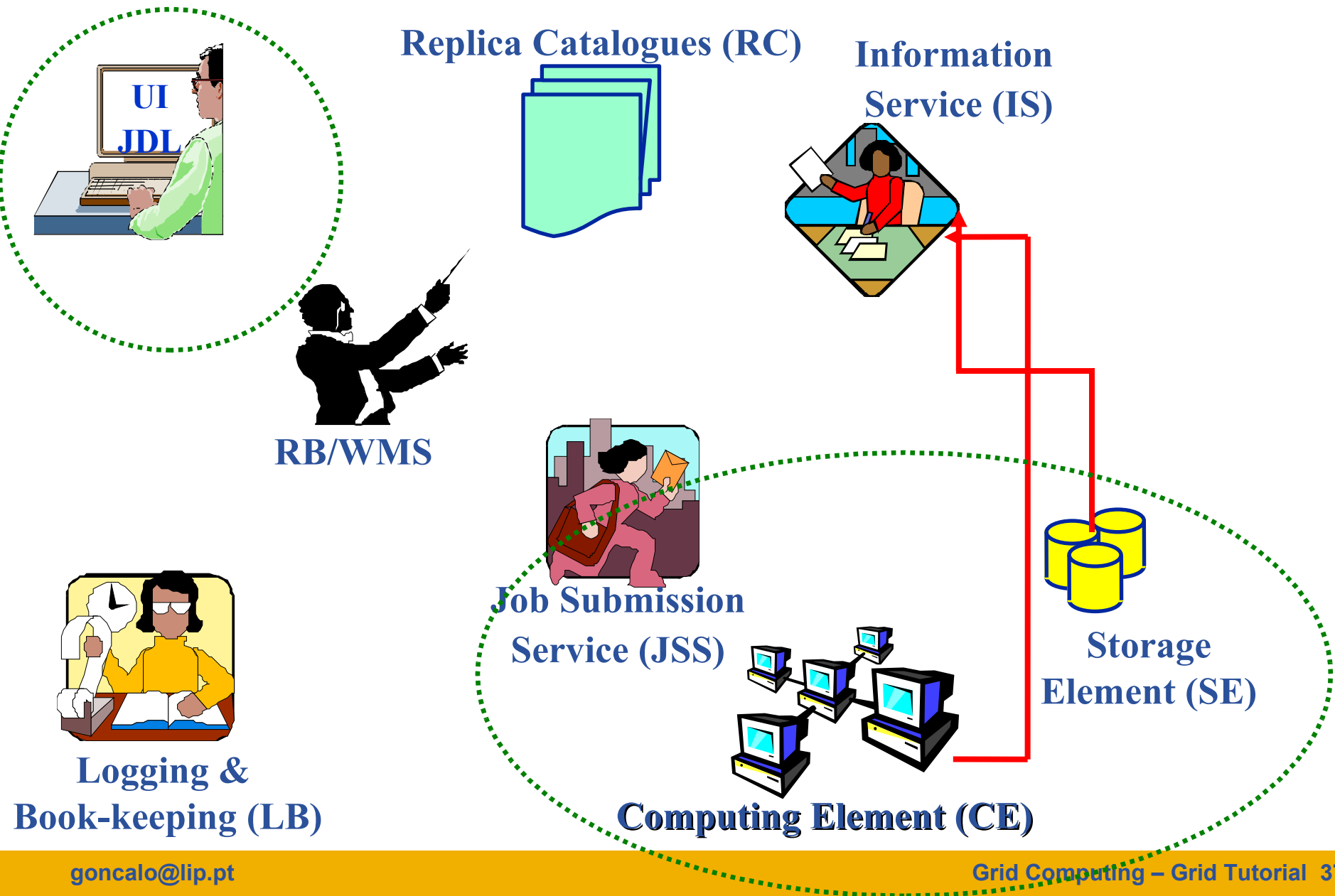
5. Job Submission

■ Jobs are sent to the RB/WMS by the user and Job specifications are encoded in the “**Job Description Language**”

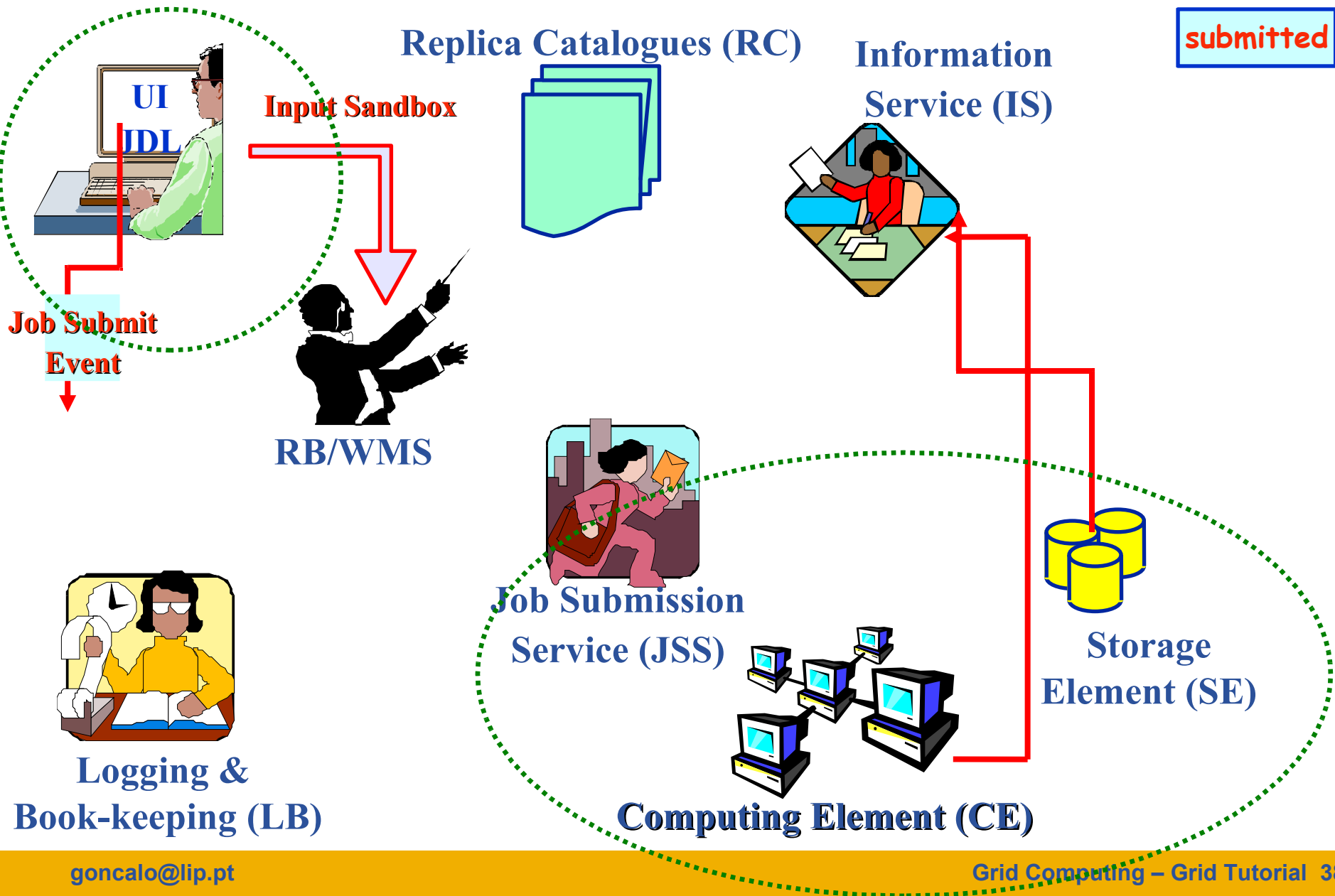
```
[i2g-ui02.lip.pt ~]# cat onsimple.jdl
JobType           = "Normal";
Executable        = "./myexecutable";
Arguments         = "--config small-file.cfg --data BIG-
file.dat";
StdOutput         = "std.out";
StdError          = "std.err";
InputSandbox      = {"myexecutable", "small-file.cfg"};
OutputSandbox     = {"std.out", "std.err", "run.log"};
InputData         = {"lfn:/grid/V0/user/BIG-file.dat"};
DataAccessProtocol = {"gridftp"};
[i2g-ui02.lip.pt ~]# i2g-job-submit onsimple.jdl
```

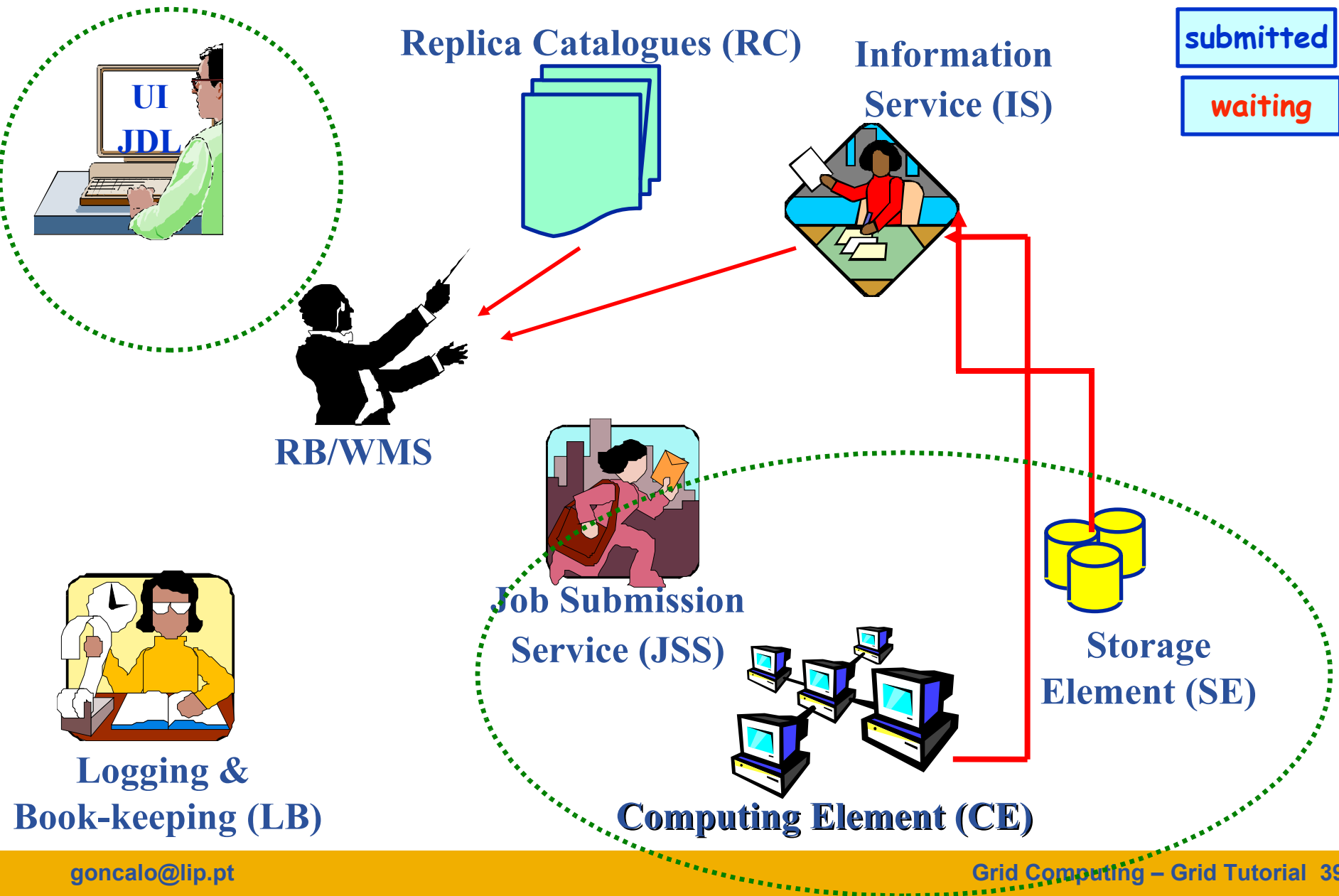
■ According to the meta-scheduler used, the client command tools are also different:

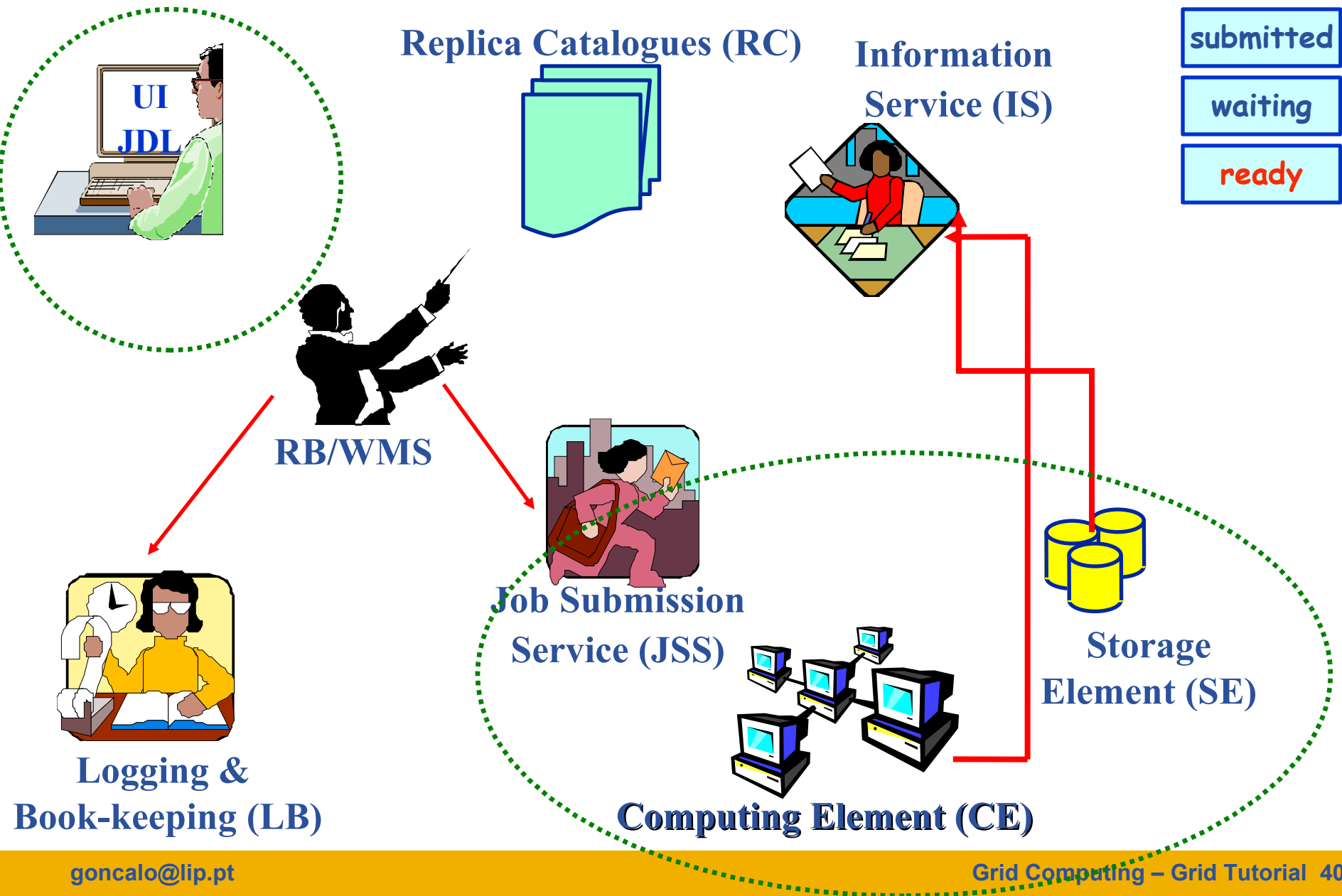
- edg-job-submit (RB), glite-wms-job-submit (WMS)
- i2g-job-submit (CrossBroker)

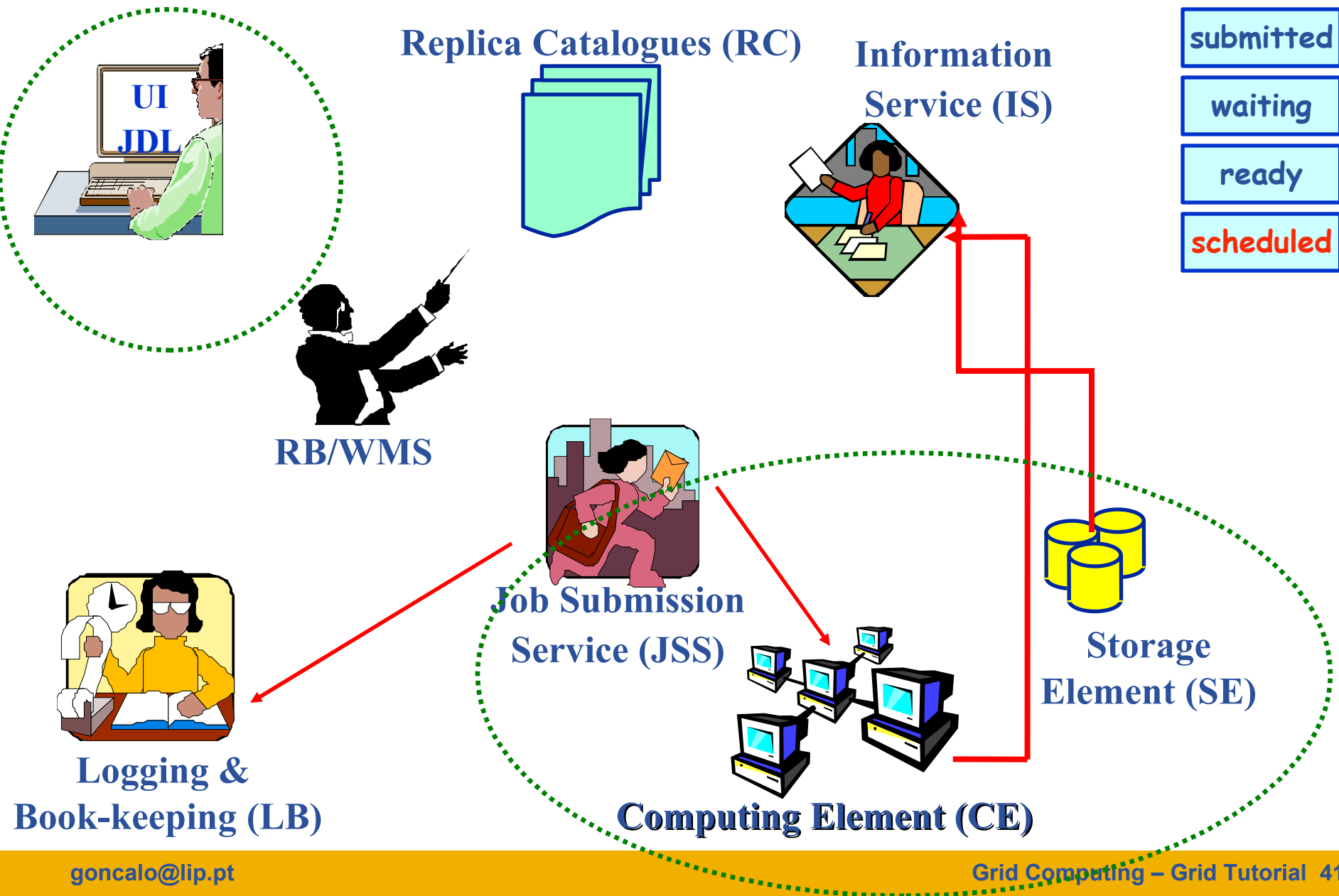


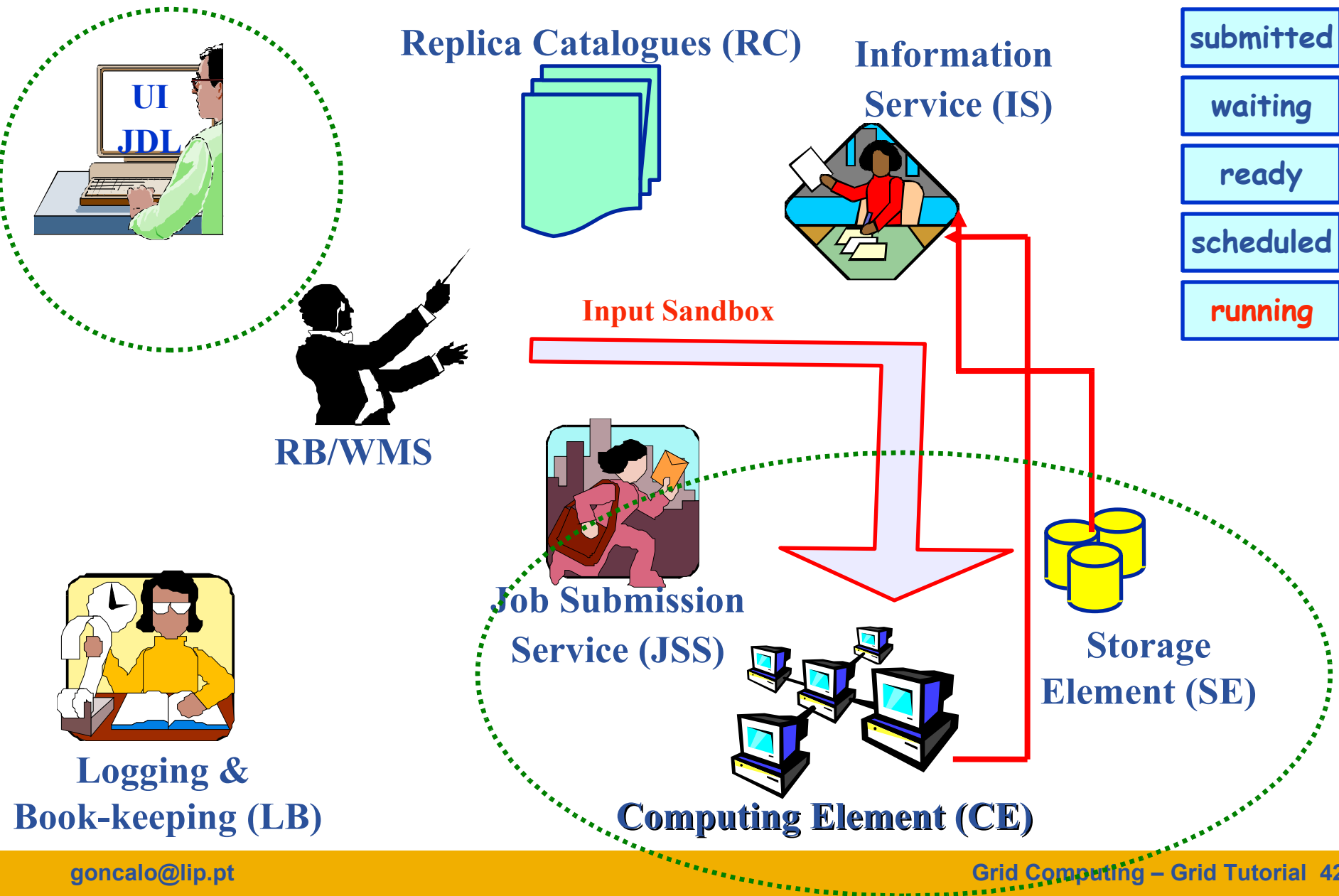
submitted

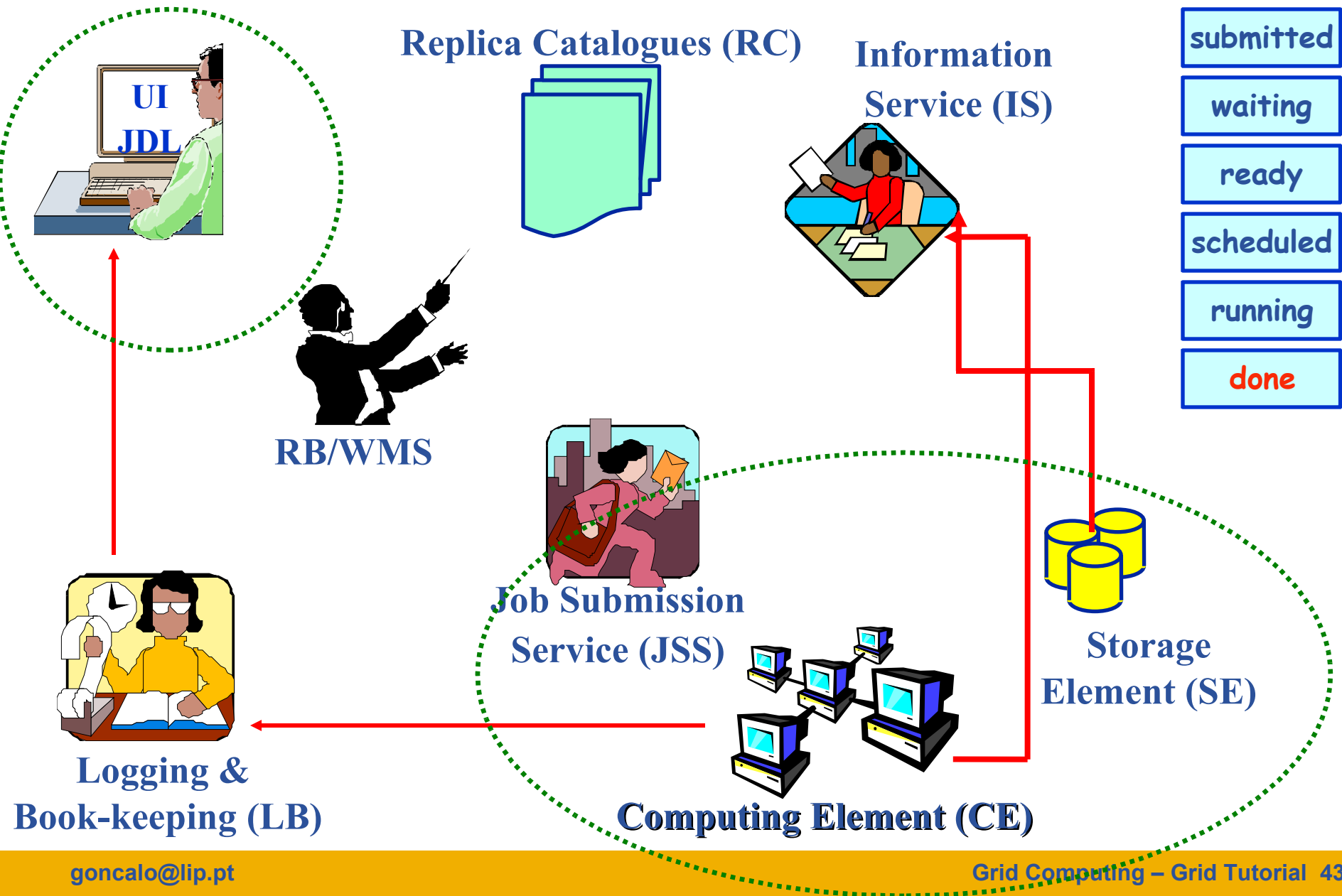


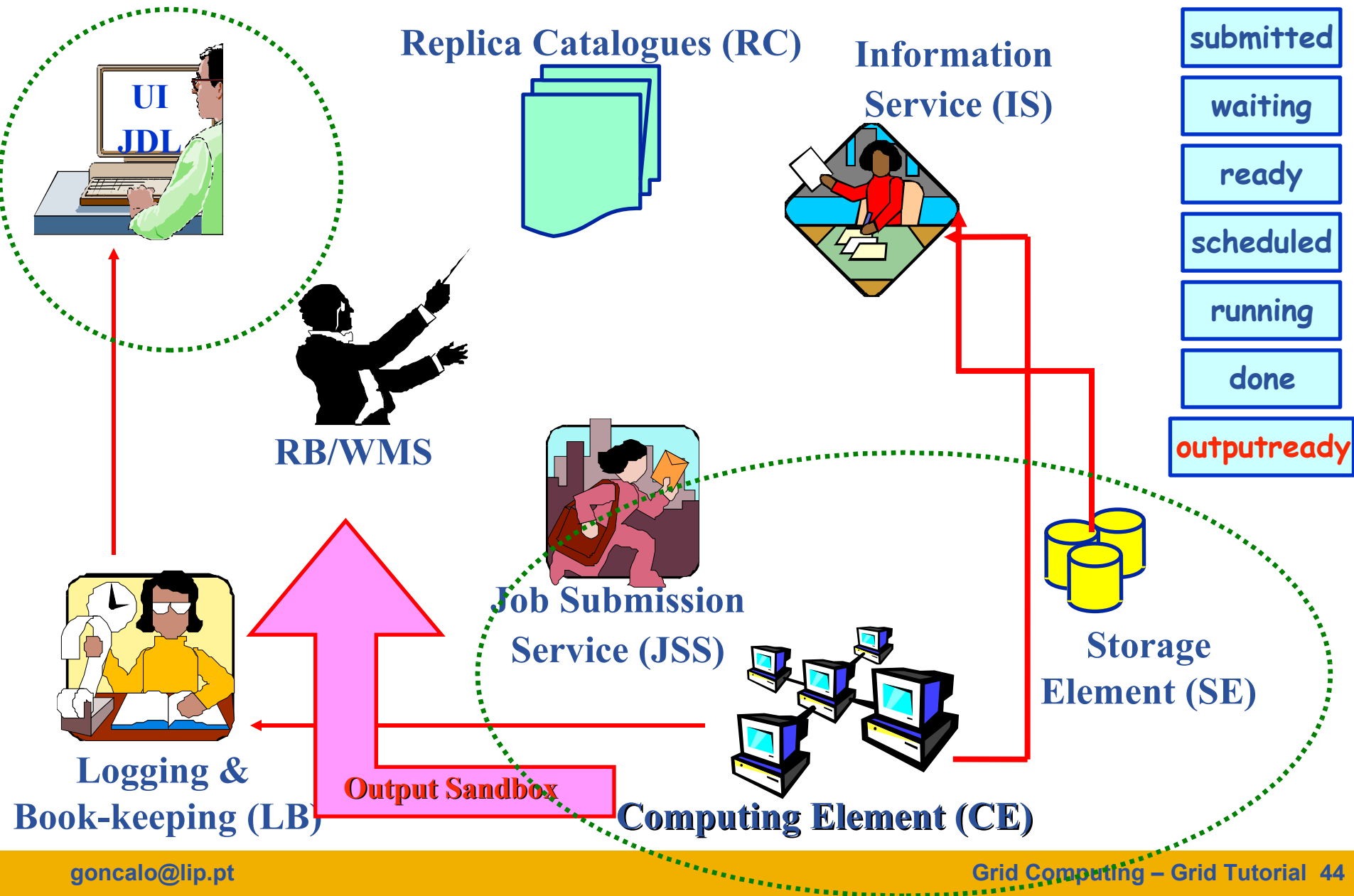


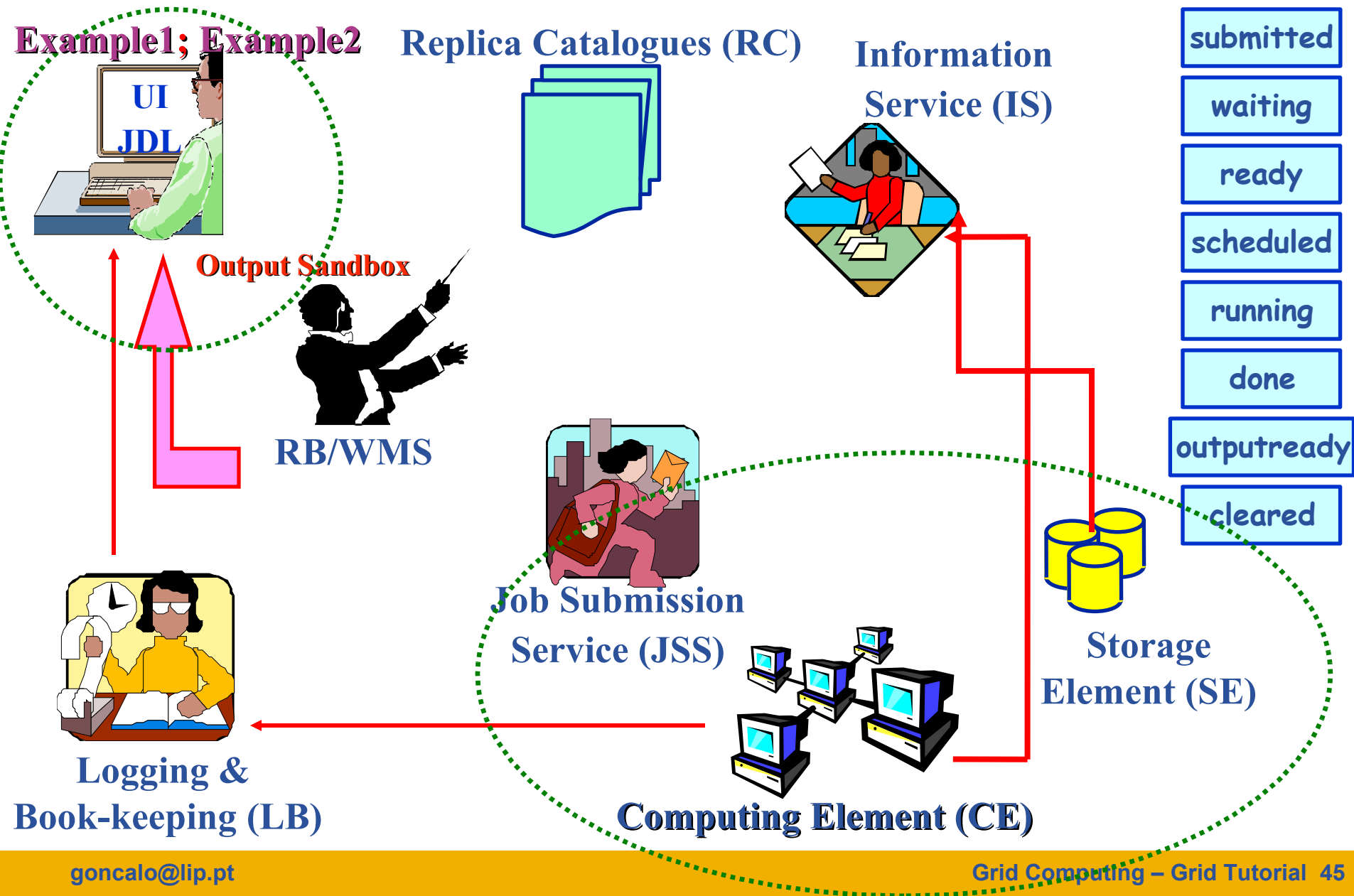












■ LIP

- <http://www.lip.pt>

■ The EGEE Project

- <http://www.eu-egee.org>

■ The LCG Project

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

■ The gLite middleware

- <http://www.glite.org>

■ The Globus Project

- <http://www.globus.org>

Security

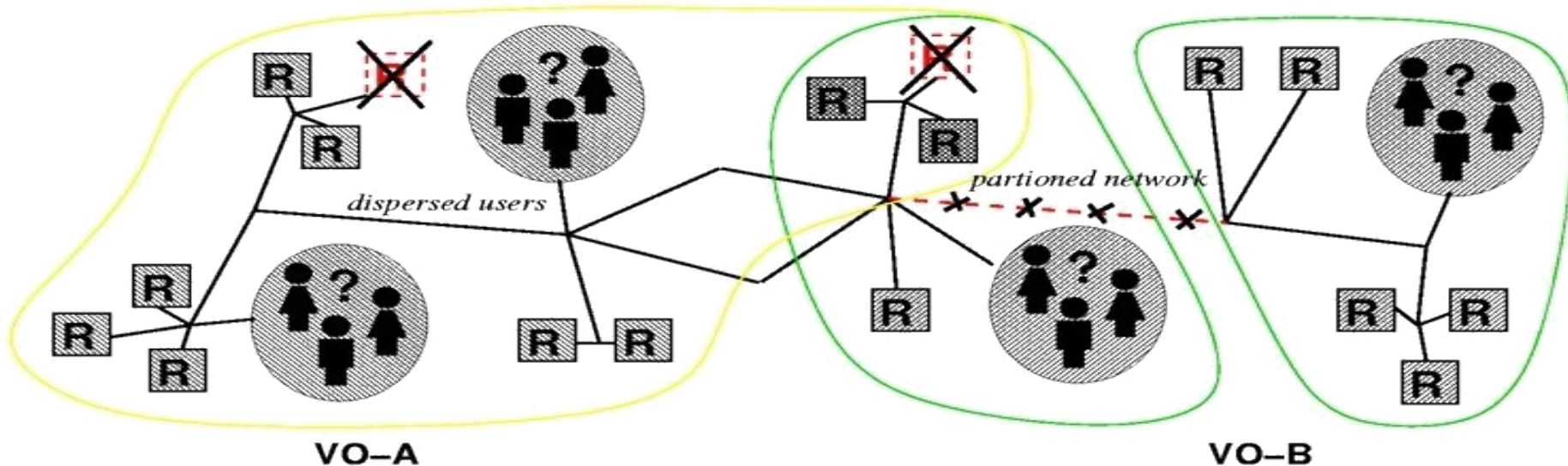
authentication, authorization and
access control in gLite

Mario David, Gonalo Borges, Jorge Gomes

LIP

***With thanks for some slides to EGEE, int.eu.grid and Globus
colleagues***

- **Overview**
- **Authentication**
- **Authorization**
- **Certificates and CA's**
- **VO's, how to “subscribe” to a Virtual Organization**

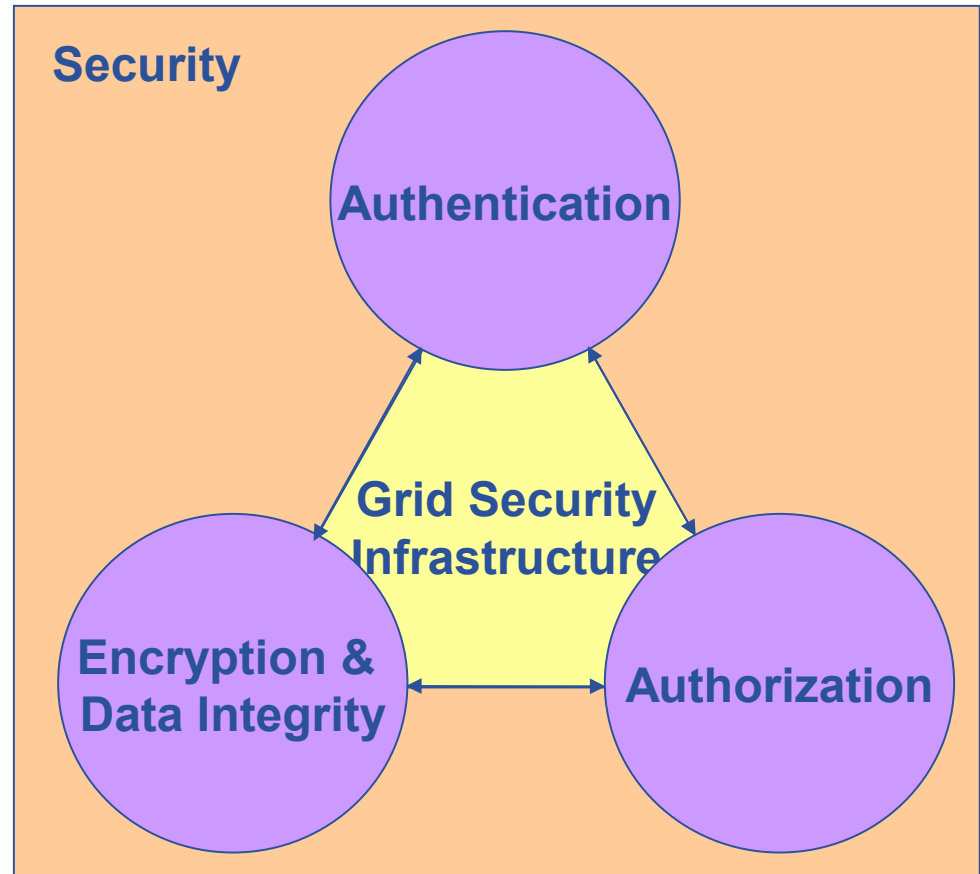


- Authorization based on the virtual organizations (VO) concept
- VO is a collection of:
 - Users with their roles in the VO
 - Resources (computing, storage ...)
- The VOs are based around users that share common goals
- VO for each: application, scientific area, experiment ...
- VO contains resources for a particular use and set of users
- Each VO should bring their own resources !!!

- **The challenge ...**
- **How to share resources**
 - to support multiple user communities with members from many different organizations with common goals but with
 - complex internal organization
 - different roles and responsibilities
 - sometimes with sub communities
 - desiring different access policies within the VOs
 - across highly distributed computing infrastructures
 - geographically and politically
 - across multiple administrative domains
 - crossing firewalls and different security policies

- How can the members of the VO be identified?
- Who does belong to a VO? Who does not?
- How does a machine identifies its client?
- How are access rights controlled?
- How does a user access a VO resource without having an user account on the machines in between or even on the resource?
- How can we still enable the sites to have control on their resources usage ?
- ...

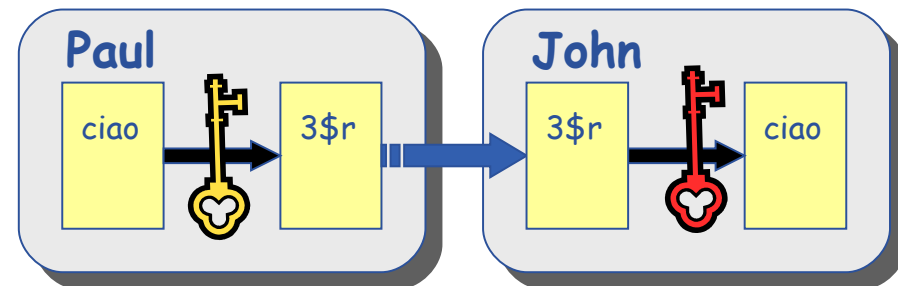
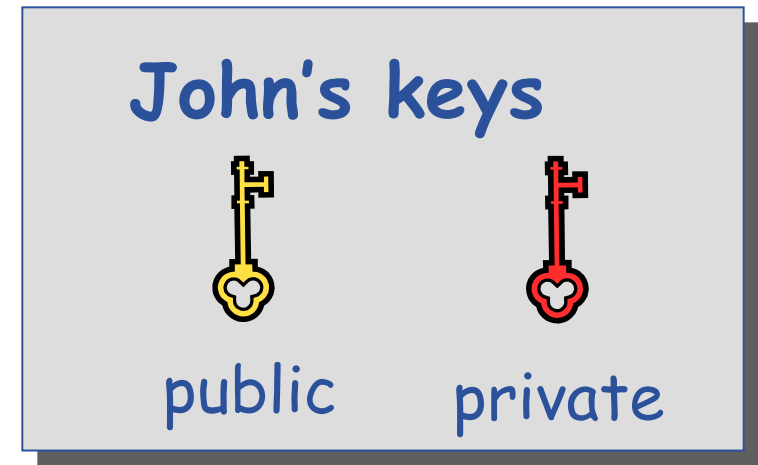
- **Authentication:** how is identity of user/site communicated?
- **Authorisation:** what can a user do?
- **Encryption:** encrypted messages
- **Integrity:** unchanged messages



GSI uses public key cryptography (also known as asymmetric cryptography) as the basis for its functionality.

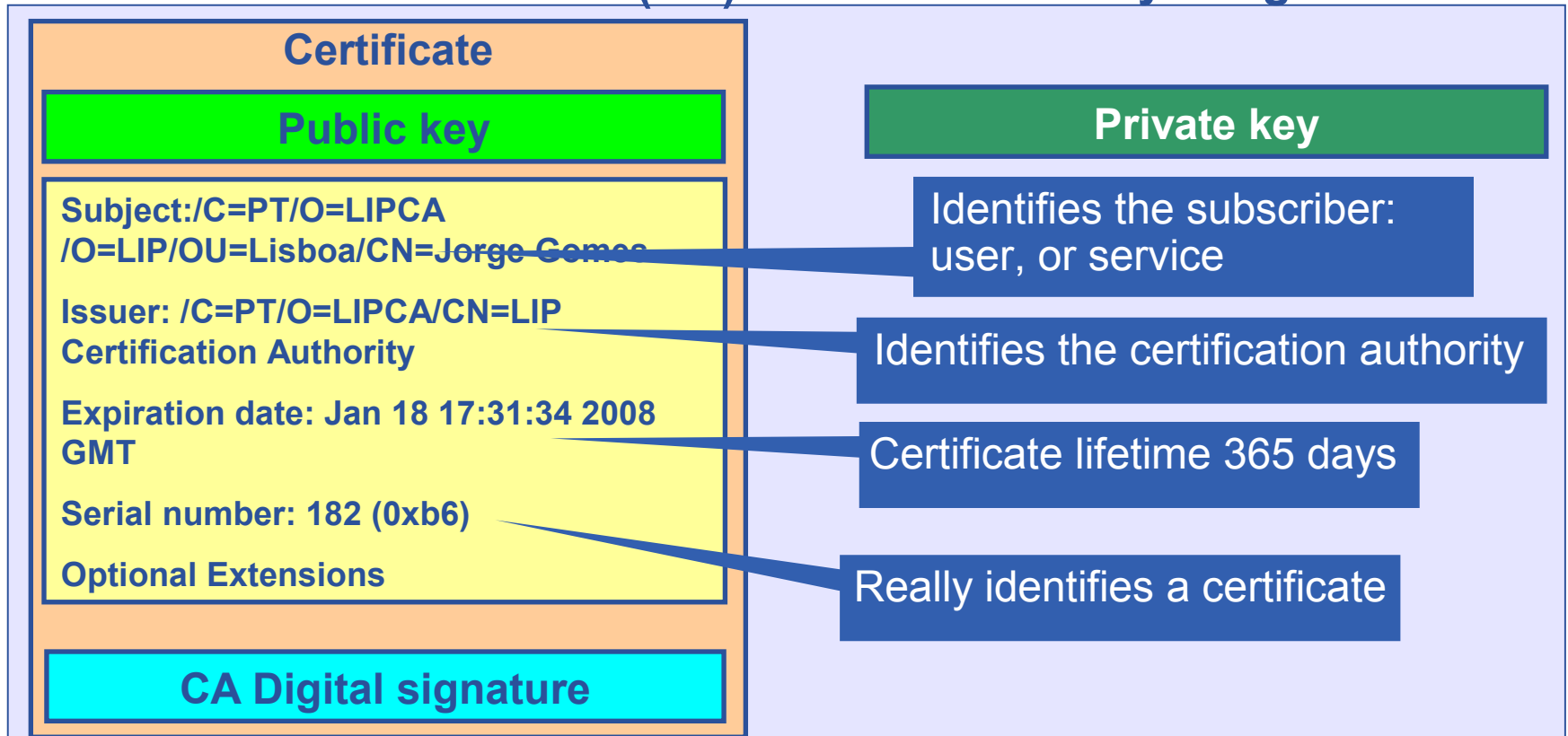
A central concept in GSI authentication is the certificate. Every user and service on the Grid is identified via a certificate.

- Every entity that wants to join a VO (user/machine/software) has two keys: one *private* and one *public*:
 - it is *impossible* to derive the private key from the public one;
 - a message encrypted by one key can be decrypted **only** by the other one.
- **Concept (simplified version):**
 - Public keys are exchanged
 - The sender encrypts using receiver's public key
 - The receiver decrypts using their private key;



- Public key is wrapped into a “certificate file”
- Certificate files are created by trusted third parties: Grid Certification Authorities (CA)

- Private key is stored in encrypted file – protected by a passphrase
- Private and public keys are created by the grid user



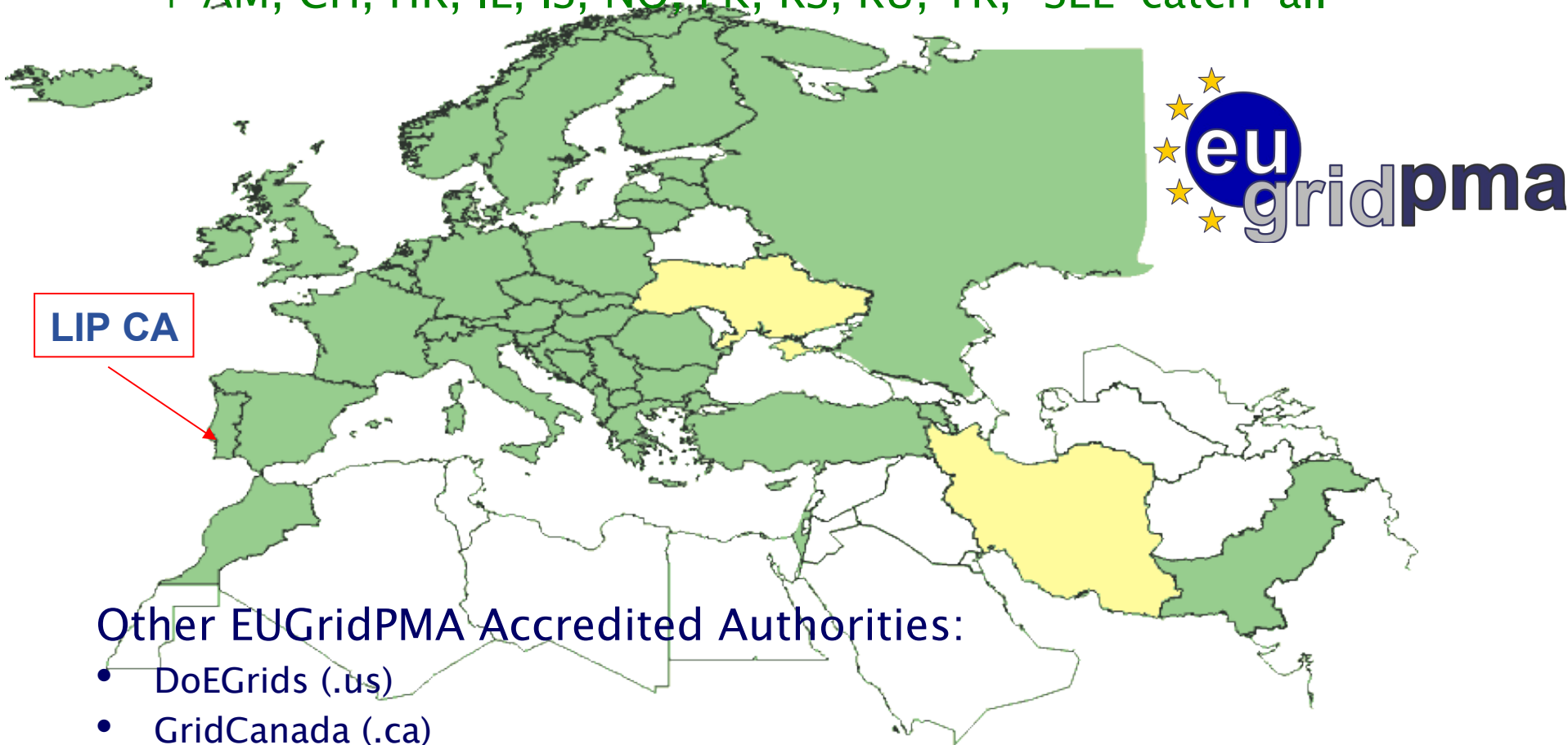
- The user's identity and its bound to a key pair has to be certified by someone
- The third party that identifies the users and issues them the key pairs is called a *Certification Authorities (CA)*
- CA's are independent bodies that issue authentication credentials (CERTIFICATES)
 - Usually they are not associated with projects or infrastructures
- A grid CA issues certificates for Grid Computing to:
 - End-users
 - Grid resources and services
- The CAs assert that:
 - the certificate requester corresponds:
 - User certificate: to the identity data in the issued certificate
 - System: to the person responsible for the system in the identity data
 - the data in the certificate was correct at the time of issuance

- There is usually one CA per country or very large organization
- Each CA issues certificates for grid users and services within its geographical or administrative scope
- Many grid CAs do exist worldwide
- To establish global grids a common trust domain had to be established
- The International Grid Trust Federation (IGTF) is the body that manages a global trust domain for grid computing serving the biggest grid infrastructures worldwide
- The IGTF is split in three regional Policy Management Authorities
 - EUgridPMA => Europe
 - APgridPMA => Asia Pacific
 - TAGPMA => Americas
- The three PMAs make use the Open Grid Forum to establish community policy and best practices for all PMAs
- The IGTF trust domain contains
 - Around 71 CAs operated by 51 entities



Green: EMEA countries with an Accredited Authority

- 23 of 25 EU member states (all except LU, MT)
- + AM, CH, HR, IL, IS, NO, PK, RS, RU, TR, “SEE-catch-all”



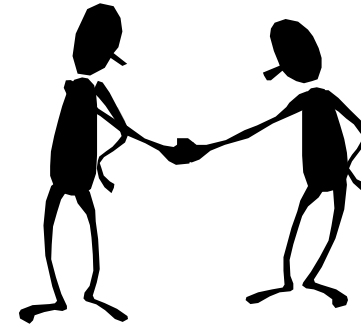
Other EUGridPMA Accredited Authorities:

- DoEGrids (.us)
- GridCanada (.ca)
- CERN

Issuing a certificate



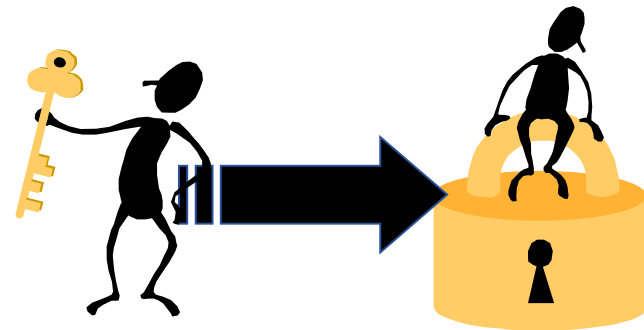
User makes a
certificate request
using its web browser



User identity is confirmed by the
Registration Authority / Certification Authority



The certificate is issued by the CA
and can be downloaded via web



The certificate can be
used for user authentication



LIPCA LIP AUTORIDADE DE CERTIFICAÇÃO

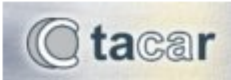
POLITICA
VALIDADE E RESPONSABILIDADE
CERTIFICADO CA ROOT
DESCARREGAR A CRL
OBTER UM CERTIFICADO
CERTIFICADOS E GLOBUS
WEB BROWSERS
CERTIFICADOS EMITIDOS
OUTRAS INFORMAÇÕES
FORMULÁRIOS
ENGLISH VERSION 
PÁGINA INICIAL

A autoridade de certificação do LIP fornece certificados x509.

20 Junho 2003
O LIP Autoridade de Certificação CP/CPS para a comunidade académica Portuguesa foi aprovada pelas autoridades de certificação numa reunião no CERN.

30 Junho 2004
O novo LIPCA para a comunidade académica Portuguesa está online.

O LIPCA é membro do EUGridPMA 

O LIPCA está listado no TERENA TACAR repository  , pode usar esse repositório para validar a integridade do nosso certificado root.

LIP
LIP Autoridade de Certificação
Av. Elias Garcia 14, 1º
1000-149 Lisboa, Portugal

- **The LIP CA is the IGTF grid CA for Portugal**
 - It is supported by all IGTF relying parties
- <http://ca.lip.pt> <https://ca.lip.pt>

Autoridade de Registo	Administrador	Operador
Centro de Física de Plasmas	Paulo Abreu	
Centro de Sistemas Inteligentes (UALG)	Margarida Madeira e Moura	Margarida Madeira e Moura
FCT/UNL	Pedro A. Duarte Medeiros	Vitor M. Alves Duarte
Instituto de Engenharia Eletrónica e Telemática de Aveiro	André Zúquete	-
LIP Coimbra	João Afonso Ferreira Bastos	João Afonso Ferreira Bastos
LIP Lisboa	Nuno Dias	Nuno Dias
Universidade Lusíada de Vila Nova de Famalicão	Rui Lima	Rui Lima, Élio Costa
Universidade Autónoma de Lisboa	Paulo Cabrita	Nuno Gameiro, Renato Bileu
Universidade do Minho	Paulo Jorge Valverde Vargas Costa	José Manuel Araújo Martins Gonçalves, José António Seara Salgado Ramada
Universidade do Porto	Gil Coutinho	Rui Manuel dos Santos Ramos, Sérgio Nuno Figueiredo da Cruz Afonso

Select one of the RAs the one that matches your organization

If none matches your organization a new RA will have to be established

- The LIP CA has started a network of registration authorities
- We welcome more registration authorities !
- The CA manager (Nuno Dias) can be reached at ca@lip.pt

General
CA Infos
User

Request a Certificate
Get Requested Certificate
Test Certificate
Revoke Certificate

Request a certificate

To request a certificate use one of this links. You will be asked to fill in a form and to confirm inserted data. After having completed the request you will have to go to the chosen RA for request approval.

Request a certificate with automatic browserdetection
[Use this link if you don't know what to do]

Netscape's Request
[User's Browser Request - SPKAC]

Internet Explorer Request
[User's Browser Request - Microsoft]

CA management software based on OpenCA modified to fit IGTF and CP/CPS requirements

- Just click on “automatic browser detection”
- Netscape, mozilla, firefox, epiphany, IE 7 do work
- Windows Vista is not supported

General | CA Infos | **User**

Request a Certificate | Get Requested Certificate | Test Certificate | Revoke Certificate

Certificate Data

E-Mail

Name (For User certificate insert Full Name as in Identity Card, for Host Certificate insert Service/Hostname, do not use accents or cedilla)

DNS name (Only fill if you are requesting a server certificate)

User Data

Name (first and Last name)

Email

Department

Telephone

Role

Registration Authority chose the RA where you will be authenticated.

PIN [used to verify the certification request, min 10 chars (please write it down for later usage)]

Re-type your PIN for confirmation

Choose a keysize

Continue

Only these fields are used in the certificate

The DNS name is for server certificates

Enter you full name as in identity card

Information for administrative purposes

Select user or web server for user or server certificates

Enter a passphrase to recover your certificate

- **Please notice:**

- The key pair (private + public key) is generated inside the web browser
- The private key never leaves the browser
- You need the private key installed to recover the signed certificate
- Once the signed certificate is retrieved from the CA it will be stored in the web browser certificate store
- Never request a certificate from an untrusted workstation
- Always protect the certificates stored in the browser (or elsewhere) with a password !!!!!

The screenshot shows the 'User' tab of the LIP Certification Authority interface. At the top, there are tabs for 'General', 'CA Infos', and 'User', with 'User' being the active tab. Below the tabs is a navigation bar with four links: 'Request a Certificate', 'Get Requested Certificate', 'Test Certificate', and 'Revoke Certificate'. The main content area is titled 'Get Additional Parameters' and contains a message: 'You need to enter some additional parameters for the requested functionality.' Below this message is a text box with instructions: 'In the e-mail you should have received from us that states the certificate issuing process has been completed, it is reported a serial number that must be used at this time. It is necessary that you proceed from the same computer from which has been generated the certification request. Please fill in the form with the serial number you received and click on the 'Continue' button.' The form has two fields: 'Serial Number' with a text input box, and 'Type of Serial' with a dropdown menu showing 'Certificate's Serial'. At the bottom of the form are two buttons: 'OK' and 'Reset'.

General | CA Infos | **User**

Request a Certificate | Get Requested Certificate | Test Certificate | Revoke Certificate

Get Additional Parameters

You need to enter some additional parameters for the requested functionality.

In the e-mail you should have received from us that states the certificate issuing process has been completed, it is reported a serial number that must be used at this time. It is necessary that you proceed from the same computer from which has been generated the certification request. Please fill in the form with the serial number you received and click on the 'Continue' button.

Serial Number

Type of Serial

- Remember you need to download the certificate from the same machine used to request it !

- **To use the gLite middleware you will likely need to:**
 - extract the certificate from the browser
 - install it in a Linux system with the gLite user middleware installed
- **Exporting depends on the browser**
 - find the browser certificate management interface on Firefox do:
 - Tools->Options->Advanced->View Certificates->Export
 - Follow the wizard:
 - *Answer **Yes** export the private key*
 - *Enter a passphrase and its confirmation*
 - *Destination filename*
 - It will produce a file containing both the public and private keys in pkcs12 format
- **The grid uses the PEM format where the private key and public key are stored in two different files**
 - Therefore the PKCS12 file must be converted to PEM format
 - Use the openssl command available in Linux and UNIX systems

```
openssl pkcs12 -nocerts -in usercert.p12 -out $HOME/.globus/userkey.pem  
openssl pkcs12 -clcerts -nokeys -in usercert.p12 -out $HOME/.globus/usercert.pem  
chmod 400 $HOME/.globus/userkey.pem
```

- **Keep your private key secure**

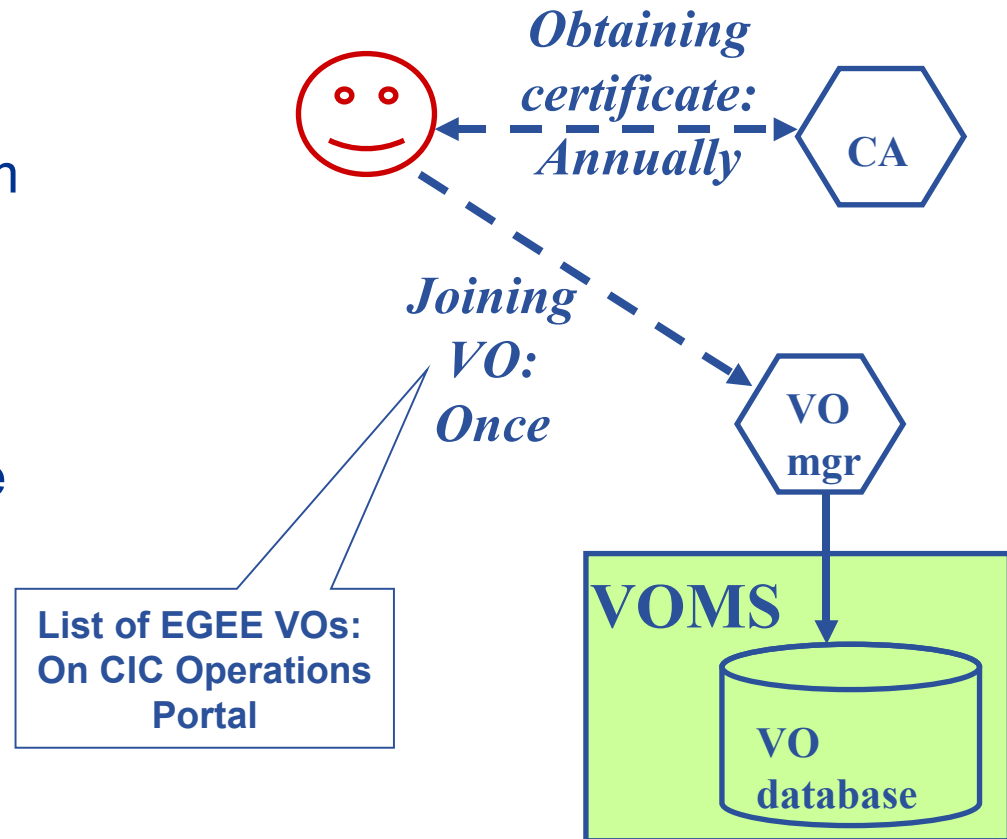
**Remember certificates are valid for one year
They have to be renewed yearly !**

**DO IT ONE MONTH PRIOR TO EXPIRATION
Otherwise you will have to pass
by the identity checks again !**

**If your certificate is used by someone other than
you, it cannot be proven that it was not you.**

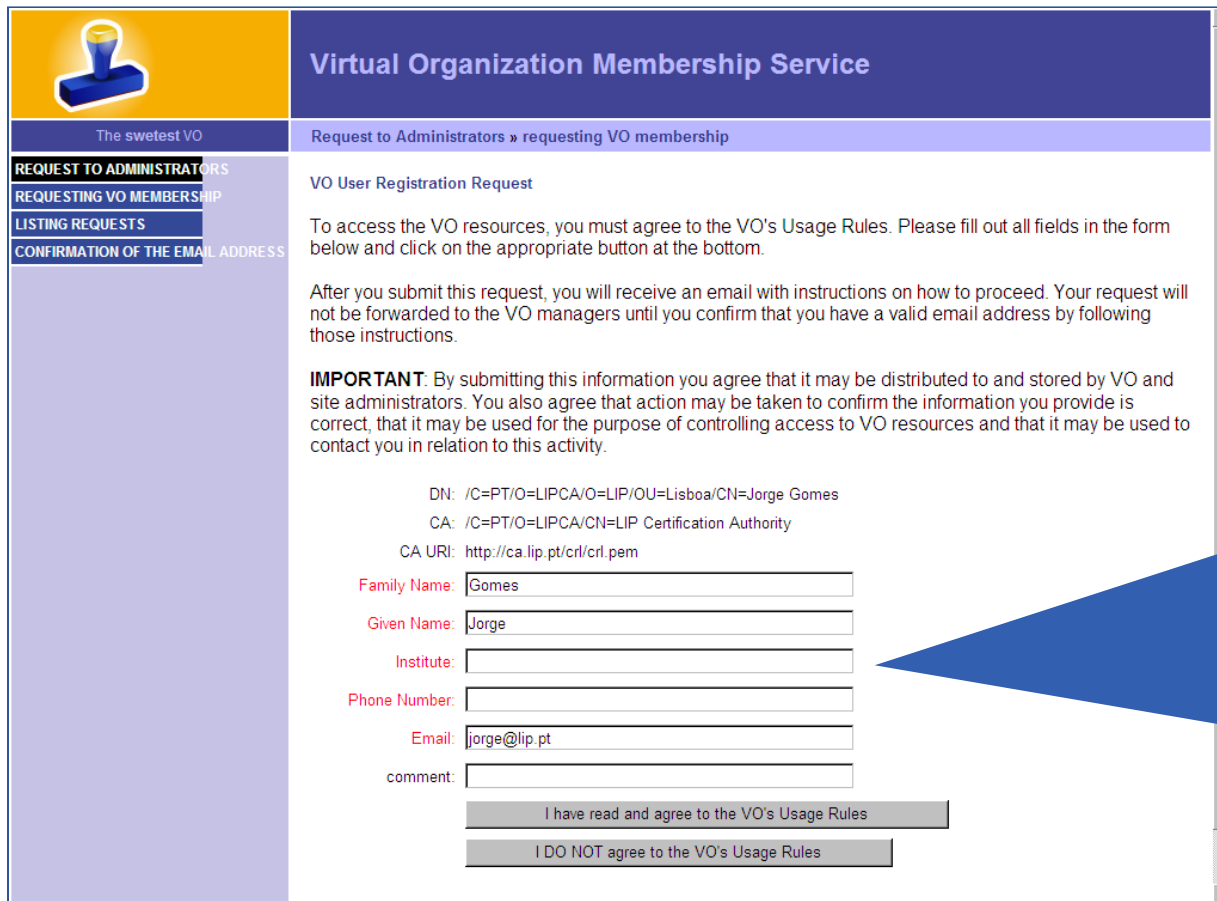
• Steps

- User obtains certificate from Certification Authority
- User registers at the VO
 - via VOMS
- VO manager authorizes the user
 - via VOMS



User's identity in the Grid = Subject of certificate:

/C=PT/O=LIPCA/OU=Lisbon/CN=Jorge Gomes



Need to have the Certificate loaded in the browser.

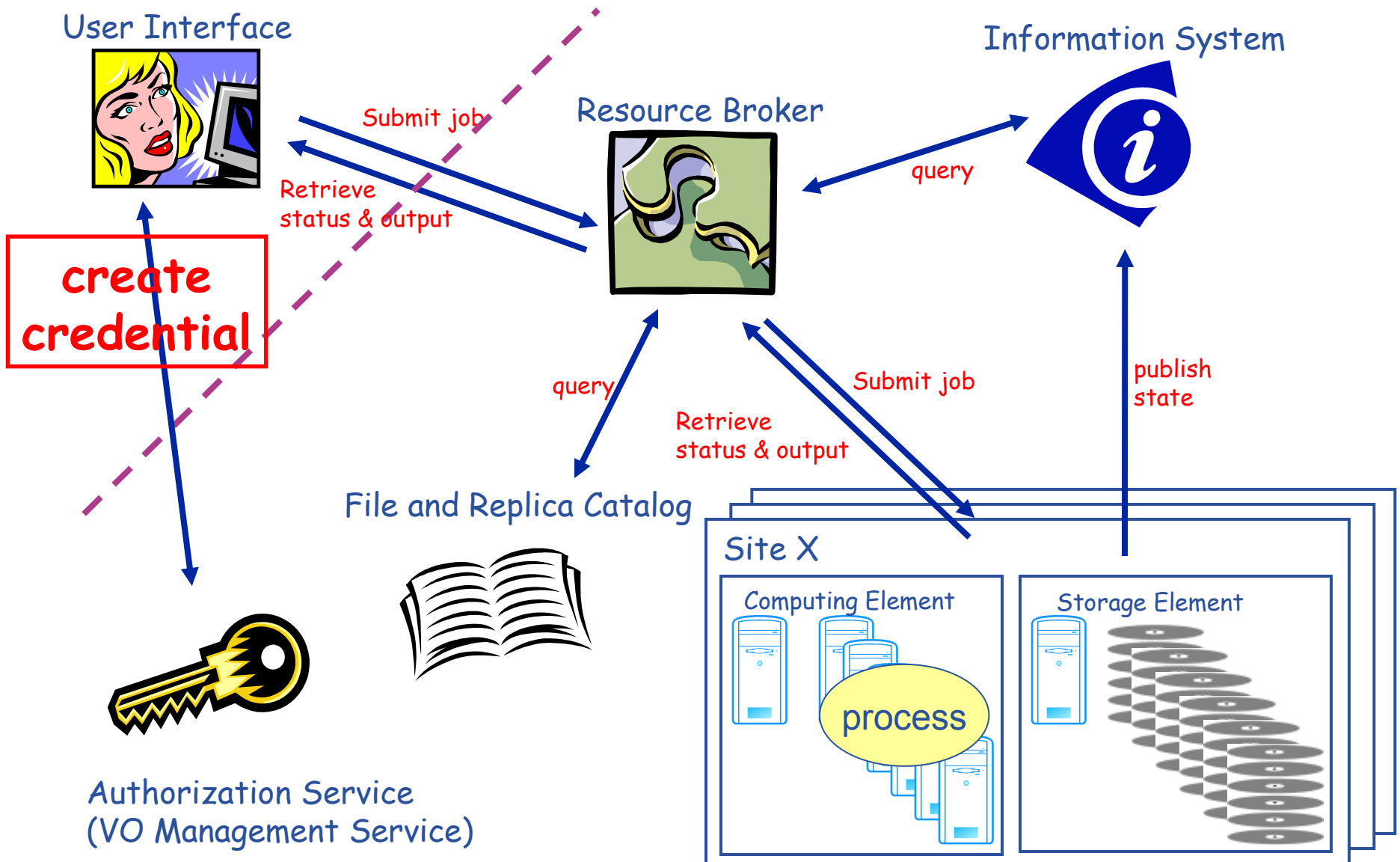
Some information is filled in automatically from the certificate information

Read the VO guidelines before accepting to join

The VO manager will contact you to know who you are and check if you can join the VO

- VOMS is a service to manage virtual organization user memberships
- See CIC portals for joining EGEE VOs

<http://cic.in2p3.fr/>



```
[jorge]$ voms-proxy-init -voms dteam
Your identity: /C=PT/O=LIPCA/O=LIP/OU=Lisboa/CN=Jorge Gomes
Enter GRID pass phrase:
Creating temporary proxy ..... Done
Contacting lcg-voms.cern.ch:15004 [/DC=ch/DC=cern/OU=computers/CN=lcg-
voms.cern.ch] "dteam" Done
Creating proxy ..... Done
Your proxy is valid until Tue Oct 23 06:37:52 2007
```

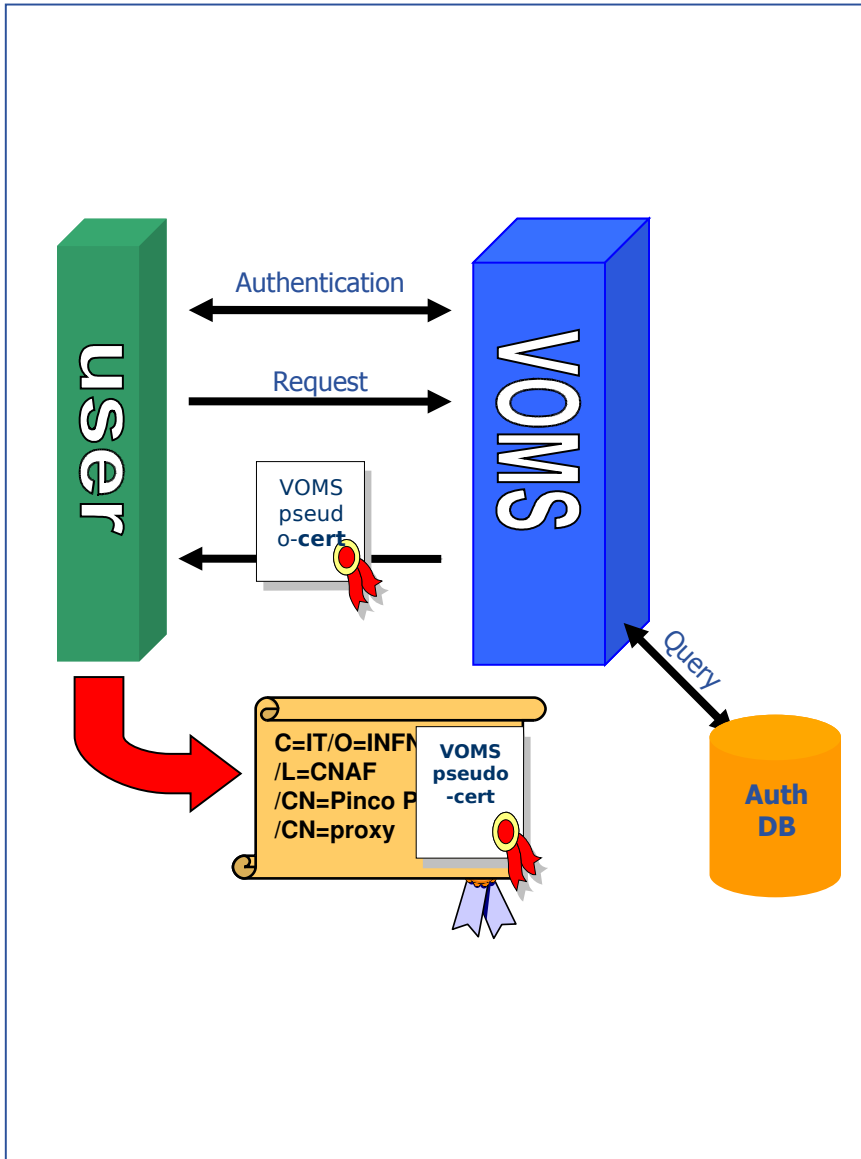
% voms-proxy-init → login to the Grid

Enter GRID pass phrase: ***** → private key is protected by a password

— Options for voms-proxy-init:

- VO name
- -hours <lifetime of new credential>
- -bits <length of key>
- -help

- **voms-proxy-init** $\equiv$ “login to the Grid”
- **Before “logout” you have to destroy your proxy:**
 - **voms-proxy-destroy**
 - This does *NOT* destroy any proxies that were delegated from this proxy.
 - You cannot revoke a remote proxy
 - Usually create proxies with short lifetimes
- **To gather information about your proxy:**
 - **voms-proxy-info**
 - Options for printing proxy information
 - subject -issuer
 - type -timeleft
 - strength -help



- VOMS: VO Management Service
 - VO level service
 - Database of user roles
- voms-proxy-init
 - Creates a proxy locally
 - Contacts the VOMS server and extends the proxy with a role

voms-proxy-init **–voms swetest**

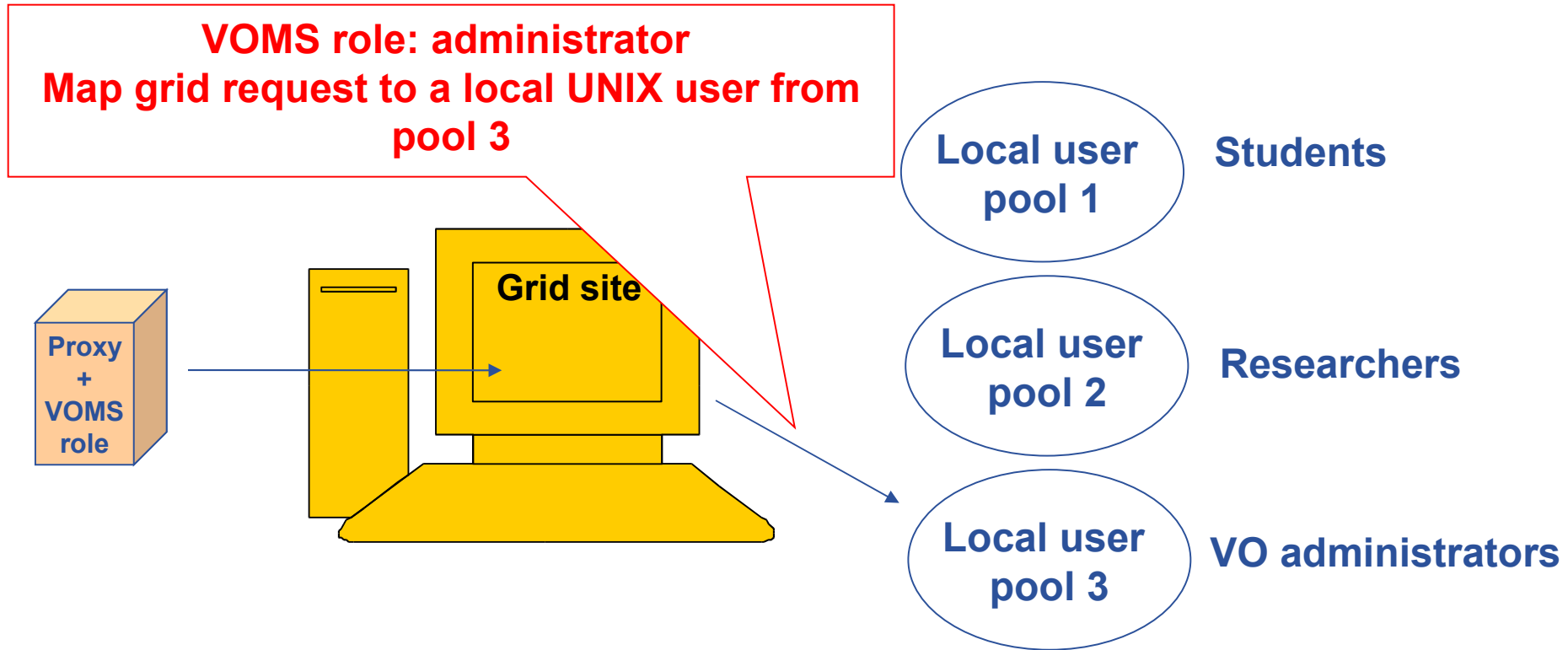
- Allows VOs to centrally manage user roles

Before VOMS

- All VO members have same rights
- Grid user identities are mapped onto local user accounts statically
- User is authorised as a member of a single VO (no aggregation of roles)
- grid-proxy-init

VOMS

- VO can have groups
 - Different rights for each
 - Different groups of experimentalists
 - ...
 - Nested groups
- VOMS has roles
 - Assigned to specific purposes
 - E.g. system admin
 - When assume this role
- User can be in multiple VOs
 - Aggregate roles
- Proxy certificate carries the additional attributes
- voms-proxy-init



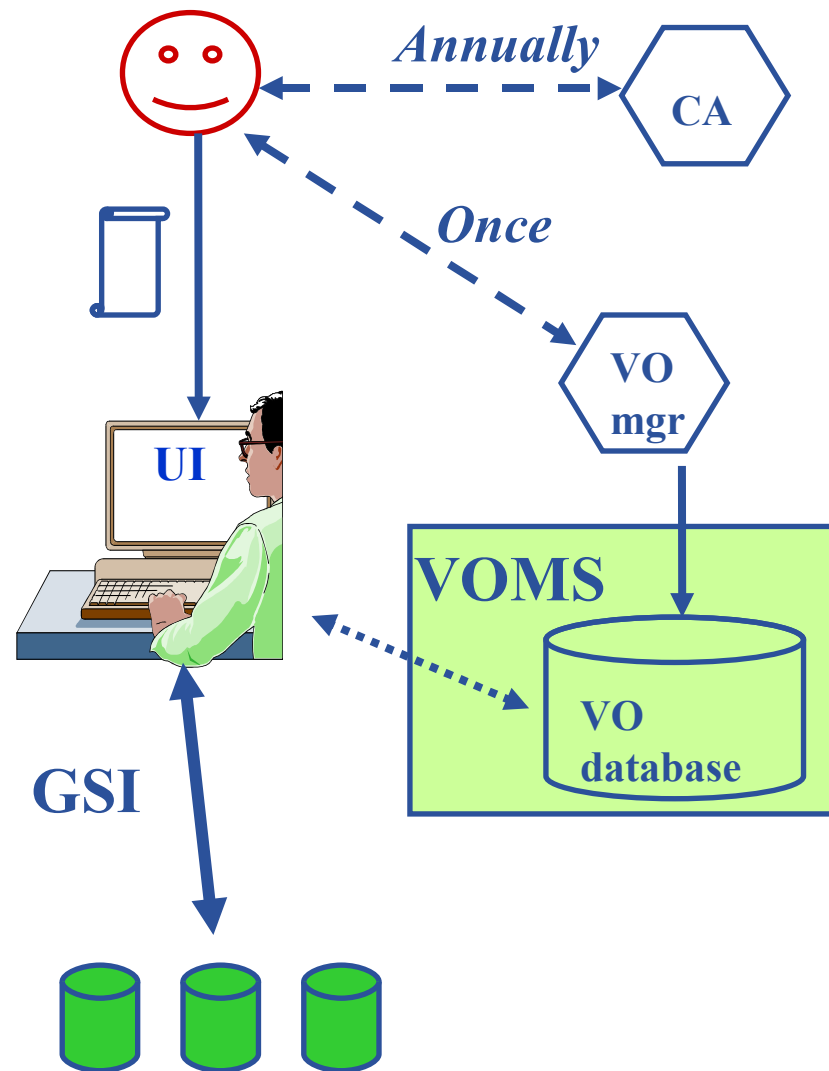
**The grid user can perform those actions on the site
that any user account from pool 3 is allowed to**

- **Authentication**

- User obtains certificate from Certificate Authority
- Connects to UI by ssh and uploads certificate to UI
or
Login to a portal and use MyProxy
- Single logon to the Grid - create proxy
- then **Grid Security Infrastructure uses proxies**

- **Authorisation**

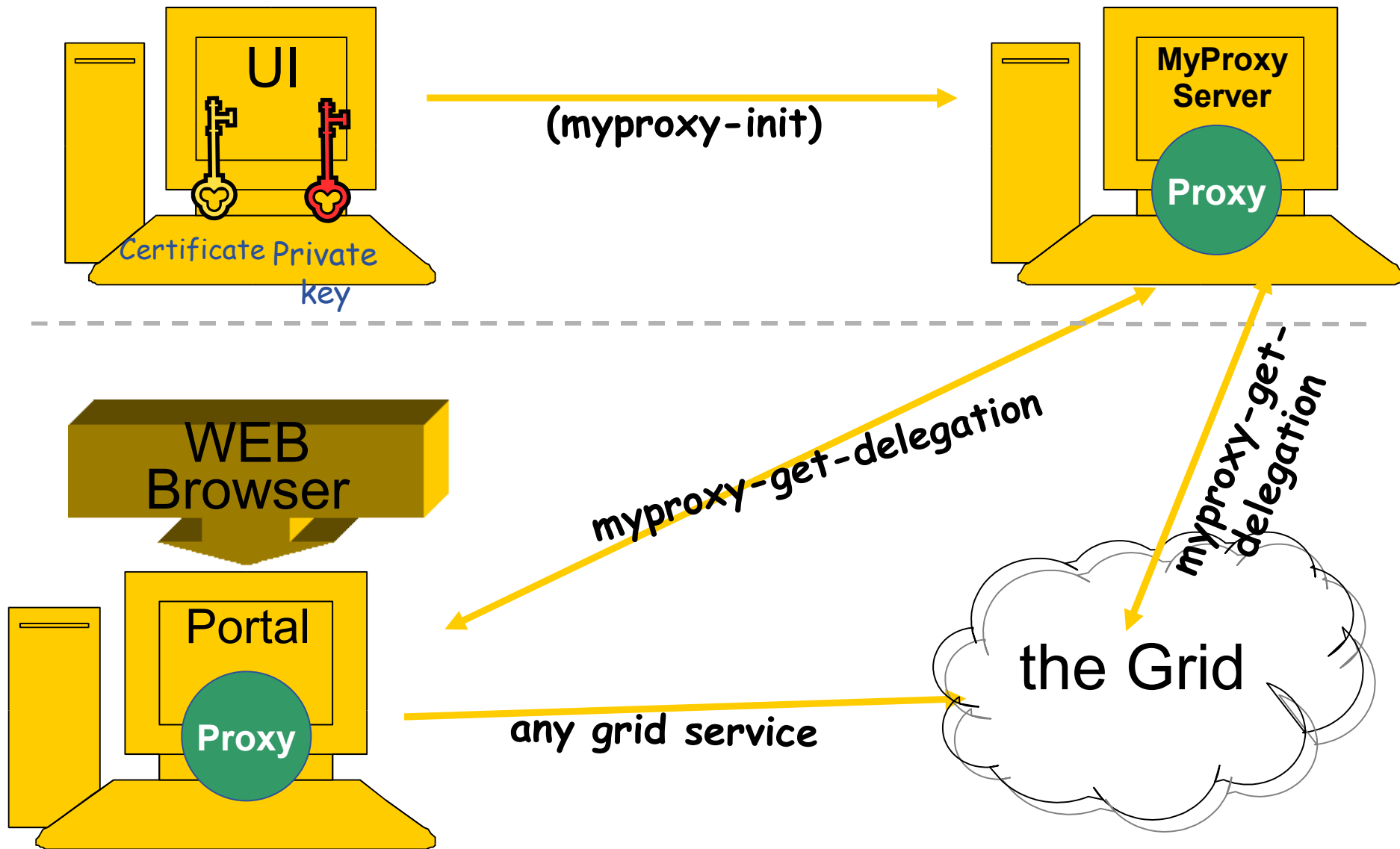
- User joins Virtual Organisation
- VO manager updates VOMS DB
- Capabilities added to proxy by VOMS



- Do not launch a delegation service for longer than your current task needs !!!!

If your certificate *or delegated service* is used by someone other than you, it cannot be proven that it was not you.

- **You may need:**
 - To interact with a grid from many machines
 - And you realise that you must NOT, EVER leave your certificate where anyone can find and use it....
- **Solution: you can store a proxy in a “MyProxy server” and derive a proxy certificate when needed**
- **MyProxy ~ storage server for proxy files**



- **Obtain a certificate from a recognized CA:**
 - www.gridpma.org → 1 year long, renewable certificates, accepted in every EGEE VO
 - For Portugal this is the LIPCA <http://ca.lip.pt>
- **Find and register at a VO**
 - EGEE NA4 - CIC Operations portal: <http://cic.gridops.org/>
- **Use the grid:**
 - **command line clients installed on the User Interface server**
(UI is a machine maintained by: the VO / your institute / you)
 - **voms-proxy-init –voms VONAME**
 - **voms-proxy-destroy**

<http://www.globus.org/security/overview.html>
<http://www.gridpma.org/>

<http://ca.lip.pt/>
<http://ca.lip.pt/index.php?link=info>

Information System

EGEE & Int.EU.Grid Tutorial
Porto, 15th May 2008

Gonçalo Borges, Jorge Gomes, Mário David
LIP Lisboa

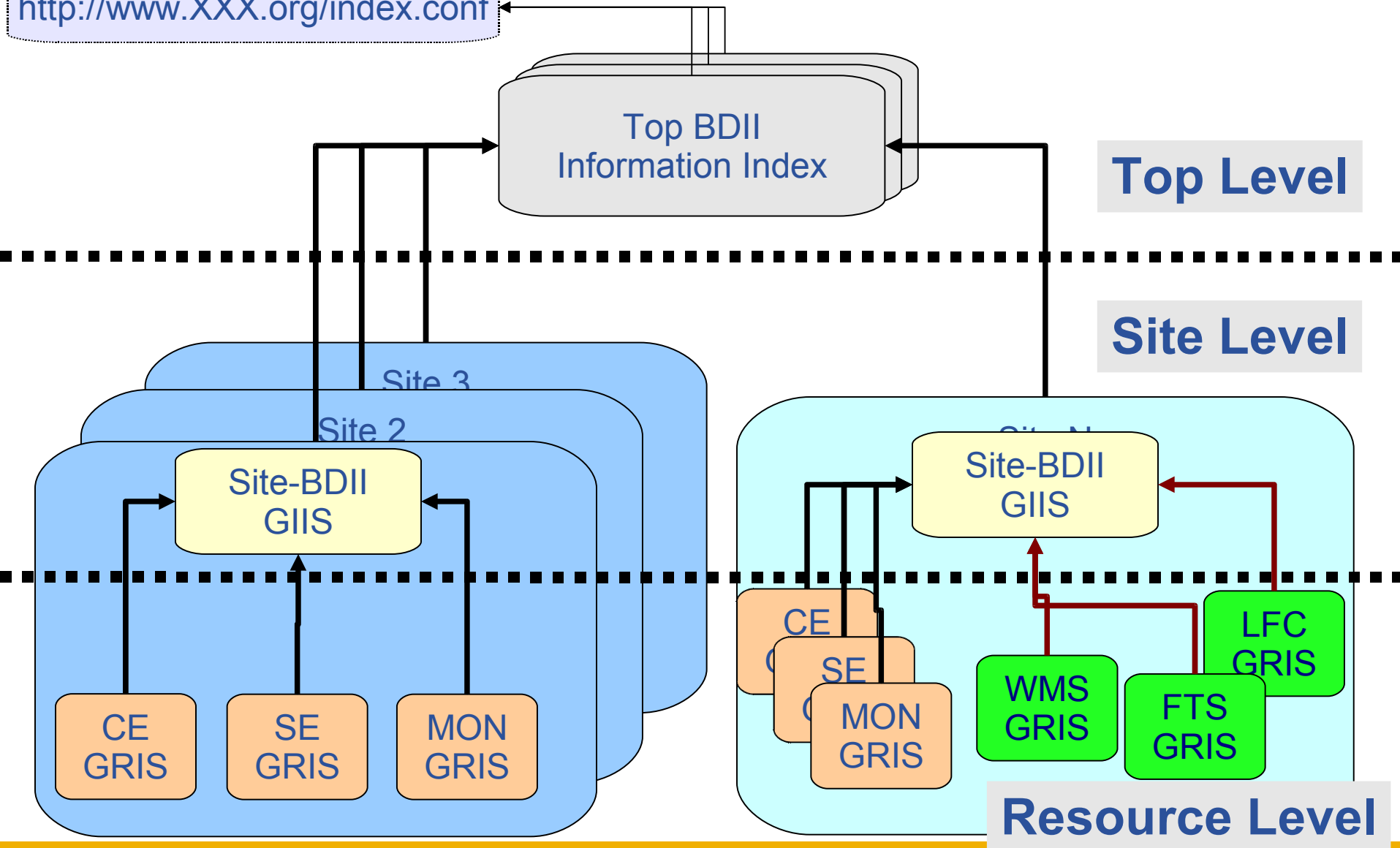


www.eu-egee.org

- **Grid Information Systems Overview**
- **Architectures:**
 - LCG Information System
 - Relational Grid Monitoring Architecture (**R-GMA**)
 - Will not be covered in this tutorial
- **Information data model**
 - *Grid Laboratory Uniform Environment* (**GLUE**) Schema
- **Resource information and Monitoring**

- **Collect information of/from Grid resources:**
 - Discovering new added resources
 - Monitoring resource load and health status
- **Publish these information:**
 - Grid resources are dynamic “by nature”.
 - Periodically updated.
 - Well known/standard data model: The GLUE schema.
- **Used by:**
 - Users searching a concrete resource.
 - RB/WMS allocating and managing jobs.
 - Other monitoring services.

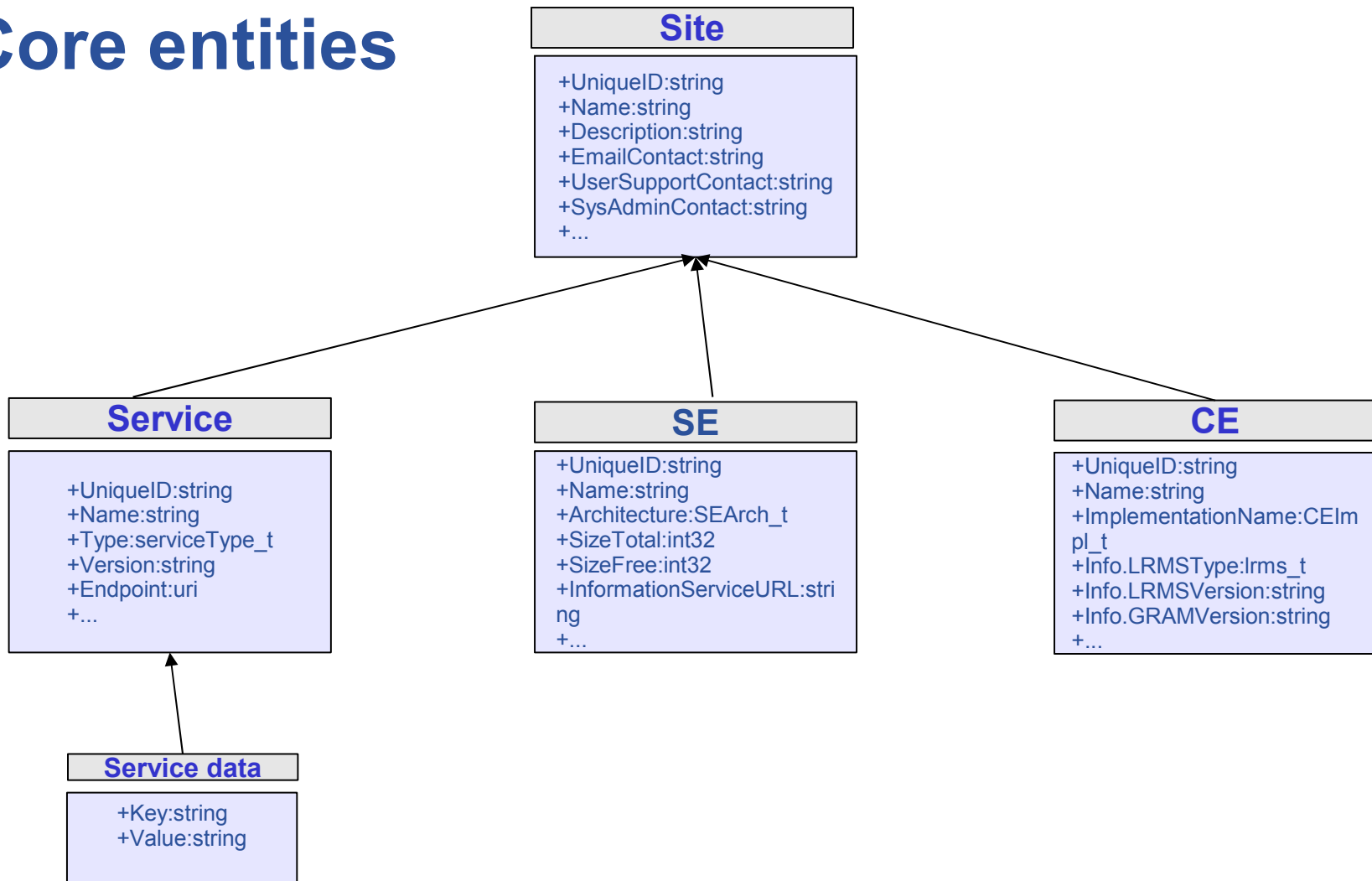
<http://www.XXX.org/index.conf>



- **Resource level: Grid Resource Information Server (GRIS)**
 - One GRIS running on each CE, SE, RB, MyProxy, etc..
 - Plugins collect static and dynamic information about the specific resource, and makes it available to be published by the GRIS.
- **Site level: Grid Index Information Server (GIIS)**
 - Collects the information of **all** GRIS's in a site.
 - Stores this information on a Berkeley DB.
 - Makes it available to the Top level Information Index.
 - Called the **site BDII**.
- **Top level: Berkeley DB Information Index (BDII)**
 - Collects the information of **all** GIIS's.
 - Stores this information on a Berkeley DB.
 - Only queries sites that are included in a configuration file
 - available through http.

- **GLUE Schema** is an abstract model for Grid resources
 - Maps resource properties to a concrete schema that can be used in Grid Information Services
 - “Common language” for all Grid Resources
 - Describes static and dynamic objects.
 - Presents a hierarchical representation.
 - Independent of the framework (LDAP, XML, SQL...).
- **Present release (1.3) is mapped into**
 - LDAP, XML, ClassAd(vertisement), used by Condor Matchmaking

Core entities



https://forge.cnaif.infn.it/plugins/scmsvn/viewcvs.php/*checkout*/v_1_3/spec/draft-3/pdf/GLUESchema.pdf?rev=33&root=glueschema

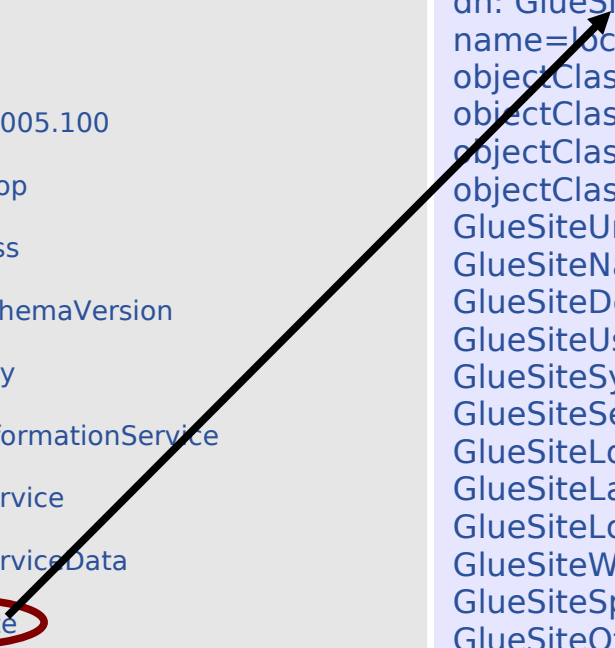
- **GLUE schema** implemented via LDAP servers
 - Service which published information available through a given port

Glue-CORE.schema

```
# OID Structure
#
# Top
# |
# ---- GlueTop 1.3.6.1.4.1.8005.100
# |
# ---- .1. GlueGeneralTop
# |
# ---- .1. ObjectClass
# |
# ---- .1 GlueSchemaVersion
# |
# ---- .4 GlueKey
# |
# ---- .5 GlueInformationService
# |
# ---- .6 GlueService
# |
# ---- .7 GlueServiceData
# |
# ---- .8 GlueSite
# |
# ---- .2. Attributes
# |
# ---- .1. Attributes for GlueSchemaVersion
#
```

static-file-Site.ldif

```
dn: GlueSiteUniqueID=LIP-Lisbon,mds-vo-
name=local,o=grid
objectClass: GlueTop
objectClass: GlueSite
objectClass: GlueKey
objectClass: GlueSchemaVersion
GlueSiteUniqueID: LIP-Lisbon
GlueSiteName: LIP-Lisbon
GlueSiteDescription: LCG Site
GlueSiteUserSupportContact: mailto:grid.admin@lip.pt
GlueSiteSysAdminContact: mailto:grid.admin@lip.pt
GlueSiteSecurityContact: mailto:grid.admin@lip.pt
GlueSiteLocation: Lisboa, Portugal
GlueSiteLatitude: 38.739925290125484
GlueSiteLongitude: -9.143403768539429
GlueSiteWeb: http://www.lip.pt
GlueSiteSponsor: none
GlueSiteOtherInfo: TIER-2
GlueSiteOtherInfo: lip.pt
GlueForeignKey: GlueSiteUniqueID=LIP-Lisbon
GlueSchemaVersionMajor: 1
GlueSchemaVersionMinor: 2
```



- **Generic Information Provider (GIP):**
 - Configurable information provider that makes a separation between **static** and **dynamic** information.
 - Produces “***Idif***” files and publishes in **LDAP servers**.
 - Information can be retrieved contacting a given port

BDII

```
ldapsearch -x -H ldap://i2g-ce01.lip.pt:2170 -b mds-vo-name=resource,o=grid
```

Site BDII

```
ldapsearch -x -H ldap://i2g-ce01.lip.pt:2170 -b mds-vo-name=LIP2G-Lisbon,o=grid
```

Top BDII

```
ldapsearch -x -H ldap://i2g-ii01.lip.pt:2170 -b mds-vo-name=local,o=grid
```

```
[ui03] # > ldapsearch -x -H i2g-ii01.lip.pt:2170 -b mds-vo-name=local,o=grid
(...)
# ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-imain, BIFI, local, grid
dn: GlueCEUniqueID=ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-imain,mds-vo-n
ame=BIFI,mds-vo-name=local,o=grid
(...)
GlueCEHostingCluster: ce-ieg.bifi.unizar.es
GlueCEName: imain
GlueCEUniqueID: ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-imain
GlueCEInfoGatekeeperPort: 2119
GlueCEInfoHostName: ce-ieg.bifi.unizar.es
GlueCEInfoLRMSType: pbs
GlueCEInfoLRMSVersion: 2.1.9
GlueCEInfoTotalCPUs: 22
GlueCEInfoJobManager: lcgpbs
GlueCEInfoContactString: ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-imain
GlueCEInfoApplicationDir: /opt/exp_soft
GlueCEInfoDataDir: unset
GlueCEInfoDefaultSE: se-ieg.bifi.unizar.es
GlueCEStateEstimatedResponseTime: 0
GlueCEStateFreeCPUs: 22
GlueCEStateRunningJobs: 0
GlueCEStateStatus: Production
GlueCEStateTotalJobs: 0
GlueCEStateWaitingJobs: 0
(...)
```

Top BDII:
ii02.lip.pt

Port:
2170

Base string:
mds-vo-name=
local,o=grid

Enabling Grids for E-science

Applications Places System

GQ

File Filters Help

Search Browse Schema

localhost
ii02.lip.pt
o=grid
mds-vo-name=local,o=grid
Mds-Vo-name=LIP-Coimbra
Mds-Vo-name=LIP-Lisbon
GlueCESEBindGroupCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-atlasgrid
GlueCESEBindGroupCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-augergrid
GlueCESEBindGroupCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-biomedgrid
GlueCESEBindGroupCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-cmsgrid
GlueCESEBindGroupCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-compassgrid
GlueCESEBindGroupCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-dteamgrid
GlueCESEBindGroupCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-edteamgrid
GlueCESEBindGroupCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-eelagrid
GlueCESEBindGroupCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-opsgrid
GlueCESEBindGroupCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-swetestgrid
GlueCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-atlasgrid
GlueCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-augergrid
GlueCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-biomedgrid
GlueCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-cmsgrid
GlueCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-compassgrid
GlueCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-dteamgrid
GlueCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-edteamgrid
GlueCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-eelagrid
GlueCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-opsgrid
GlueCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsgge-swetestgrid
GlueClusterUniqueID=ce02.lip.pt
GlueSEUniqueID=dcache01.lip.pt
GlueServiceUniqueID=gsissh://phedex.lip.pt:1975
GlueServiceUniqueID=http://lfc01.lip.pt:8085/
GlueServiceUniqueID=https://dcache01.lip.pt:8443/srm/managerv1
GlueServiceUniqueID=lfc01.lip.pt
GlueServiceUniqueID=mon01.lip.pt:2136
GlueServiceUniqueID=px01.lip.pt:7512
GlueServiceUniqueID=rb02.lip.pt:7772
GlueServiceUniqueID=wms01.lip.pt:7772
GlueSiteUniqueID=LIP-Lisbon
Mds-Vo-name=T2_Estonia
Mds-Vo-name=UPorto
mds-vo-name=AEGIS01-PHY-SCL
mds-vo-name=ALBERTA-LCG2
mds-vo-name=AMD64.PSNC.PL
mds-vo-name=AUVERGRID
mds-vo-name=Australia-ATLAS
mds-vo-name=Australia-UNIMELB-LCG2
mds-vo-name=BEIJING-LCG2
mds-vo-name=BEgrid-KULeuven
mds-vo-name=BEgrid-UGent
mds-vo-name=BEgrid-ULB-VUB
mds-vo-name=BG-INRNE

GlueCEUniqueID

GlueCEName
GlueCEHostingCluster
GlueCEInfoTotalCPUs
GlueCEInfoLRMSType
GlueCEInfoLRMSVersion
GlueCEInfoHostName
GlueCEInfoGatekeeperPort
GlueCEInfoContactString
GlueCEInfoJobManager
GlueCEInfoApplicationDir
GlueCEInfoDataDir
GlueCEInfoDefaultSE
GlueCEStateRunningJobs
GlueCEStateWaitingJobs
GlueCEStateTotalJobs
GlueCEStateStatus
GlueCEStateWorstResponseTime
GlueCEStateEstimatedResponseTime
GlueCEStateFreeCpus
GlueCEStateFreeJobSlots
GlueCEPolicyPriority
GlueCEPolicyMaxRunningJobs
GlueCEPolicyMaxTotalJobs
GlueCEPolicyMaxCPUTime
GlueCEPolicyMaxWallClockTime
GlueCEPolicyAssignedJobSlots
GlueCEAccessControlBaseRule
GlueInformationServiceURL
GlueSchemaVersionMajor
GlueSchemaVersionMinor
GlueForeignKey

ce02.lip.pt:2119/jobmanager-lcgsgge-atlasgrid
atlasgrid
ce02.lip.pt
48 (int)
sge
6.0u7
ce02.lip.pt
2119
ce02.lip.pt:2119/jobmanager-lcgsgge-atlasgrid
lcgsgge
/exper-sw
unset
dcache01.lip.pt
16 (int)
0 (int)
16 (int)
Production
0 (int)
0 (int)
3 (int)
3 (int)
1 (int)
48 (int)
1000 (int)
4320 (int)
9000 (int)
0 (int)
VO:atlas
ldap://ce02.lip.pt:2135/mds-vo-name
1 (int)
2 (int)
GlueClusterUniqueID=ce02.lip.pt

Apply Add as new Refresh

The GStat tool

- a monitoring tool checking the availability/health of all sites GII's
- The results are available in a web portal and updated every 5 minutes (<http://goc.grid.sinica.edu.tw/gstat/>)

Color Legend																	
GSTAT	OK	INFO	NOTE	WARN	ERROR	CRIT	MAINT	OFF									
SFT	OK			WARN	ERROR	CRIT	SchedDown										
No	Site Reports	GIIS Host	bnode	cernse	gperf	sanity	serv	serEntry	version	sclust	totalCPU	freeCPU	runJob	waitJob	seAvail TB	seUsed TB	max
1	BIFI	ce-egEE.bifi.unizar.es	ok	ok	ok	ok	ok	ok	GLITE-3 0 0	Scientific Linux CERN 3.0.6	22	22	0	0	0.12	0.07	24
2	CESGA-EGEE	mon.egEE.cesga.es	ok	ok	ok	ok	ok	ok	GLITE-3 0 2	Scientific Linux 3.0.7	111	50	61	8	0.40	0.83	179
3	CIEMAT-LCG2	lcg03.ciemat.es	ok	ok	ok	ok	ok	ok	GLITE-3 0 0	Scientific Linux CERN 3.0.6	192	83	0	0	33.65	31.56	196
4	CNB-LCG2	mallarme.cnb.uam.es	ok	ok	ok	ok	ok	ok	GLITE-3 0 0	Scientific Linux 3.0.5	16	16	0	0	0.00	0.12	16
5	IFCA-LCG2	egEEi01.ifca.es	ok	note	ok	ok	ok	ok	GLITE-3 0 0	Scientific Linux 3.0.8	322	292	29	238	35.70	12.92	322
6	IFIC-LCG2	lcg2ce.ific.uv.es	ok	note	ok	ok	ok	ok	GLITE-3 0 0	Scientific Linux CERN 3.0.6	186	0	119	425	1895.16	932.80	186
7	UAM-LCG2	grid003.ft.uam.es	ok	ok	ok	ok	ok	ok	GLITE-3 0 2	Scientific Linux 3.0.5	134	2	128	7	25.53	13.74	134
8	UB-LCG2	lcg-ce.ecm.ub.es	ok	ok	ok	ok	ok	ok	GLITE-3 1 0	ScientificCERNSLC 4.5	165	145	0	0	0.05	0.00	165
9	UPV-GRYCAP	ramses.dsic.upv.es	ok	ok	ok	error	error	na							20	17	ok
10	USC-LCG2	lcg-ce.usc.cesga.es	ok	ok	ok	ok	ok	ok	GLITE-3 0 2	Scientific Linux CERN 3.0.8	154	154	0	0	0.02	0.03	158
11	ifae	ifaece01.pic.es	ok	ok	ok	ok	ok	ok	GLITE-3 0 0	Scientific Linux 3.0.6	16	1	15	277	13.69	6.72	16
12	pic	site-bdii.pic.es	ok	note	ok	ok	ok	ok	GLITE-3 1 0	ScientificCERNSLC 4.4	247	24	0	0	1951.94	1116.69	258
13	ESA-ESAC	esacgrid09.esac.esa.int	ok	ok	ok	ok	ok	ok	GLITE-3 0 2	Scientific Linux 3.0.7	10	6	4	0	0.38	0.00	10

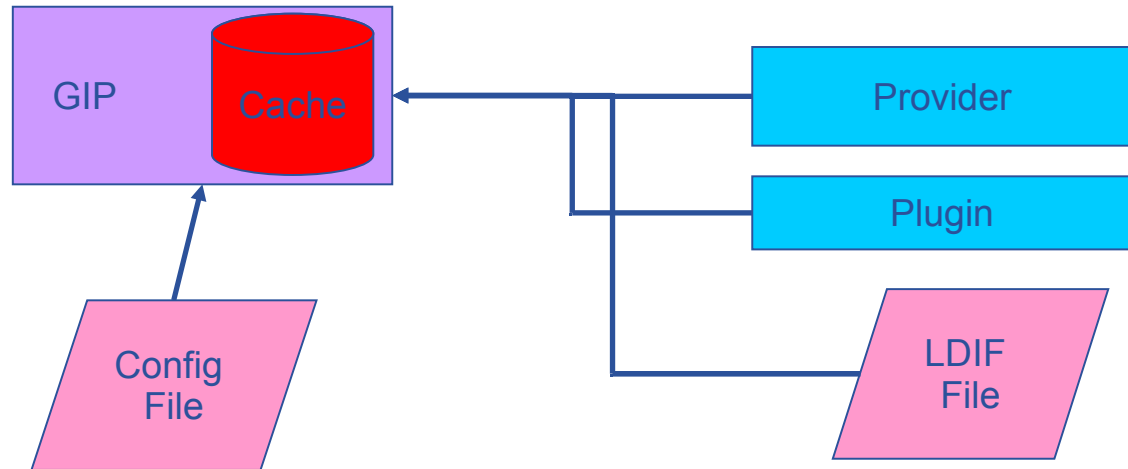
The RB contacts the top-BDII via ldap for resource discovering

- The top-BDII returns all available information collected from all site GIIS's
- The RB parses that information for matchmaking

- **Generic Information Provider (GIP):**
 - Configurable information provider that makes a separation between **static** and **dynamic** information.
 - It can be used to produce any kind of information:
 - In gLite the output format is the “*ldif*” for use with LDAP.
- **Publishing the GRIS: *BDII* or *globus-mds***
 - In the past all the Information System was based on the globus-mds package.
 - LCG/EGEE projects started to move away from the globus-mds, by using a Berkeley Database to cache the information.
 - Improvement of the robustness, stability and scalability of the system.
 - The format (“*ldif*”) and service (“*LDAP*”) of the Information System did not change.
 - The resource BDII's are updated every 30 seconds (site BDII's also updated every 30 seconds).

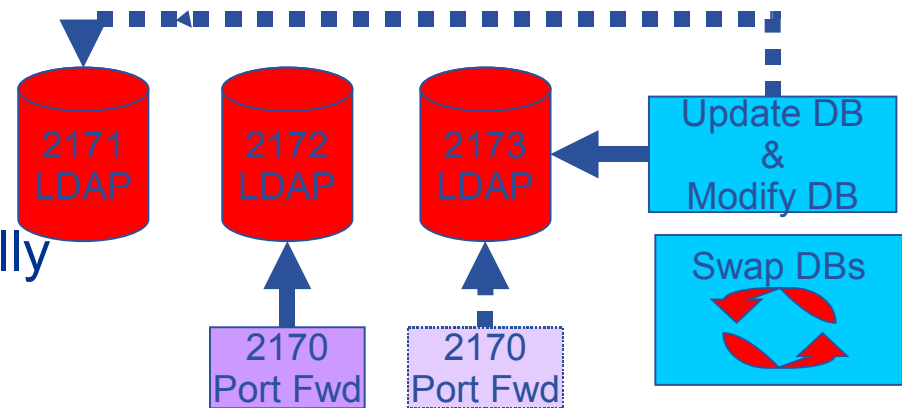
- **Generic Information Provider (GIP)**

- Provides LDIF information about a grid service in accordance to the **GLUE** Schema



- **BDII: Information system in gLite 3.1**

- LDAP database that is updated by a process
- More than one DBs is used separate read and write
- A port forwarder is used internally to select the correct DB



- **Now using version 1.3 of the GLUE schema**
 - Aim to contribute and adopt the GLUE 2.0 schema as defined by the new GLUE Working Group at OGF
- **Access to the Information System via the Service Discovery**
 - gLite Service Discovery currently supports R-GMA, BDII and XML files back ends
 - Working on a SAGA-compliant interface
- **EGEE is using the BDII as Service Discovery back-end**
 - Based on an LDAP database
 - Adequate performance to address the infrastructure needs
 - Up to 2 million queries/day served (over 20 Hz)
- **gLite 3.1 R-GMA will have authorization, Virtual DB support and schema replication**



Interactive European Grid

Information System: Hands On

EGEE & Int.EU.Grid Tutorial
Porto, 15th May 2008

Gonçalo Borges, Jorge Gomes, Mário David
LIP Lisboa



- ❑ In order to query directly the IS two high level tools are provided:
 - ▶ **lcg-infosites**
 - ▶ **lcg-info**
- ❑ These tools should be OK for common user needs
- ❑ These tools hide the complexity of raw LDAP queries (**ldapsearch**)

□ The lcg-infosites command

- ▶ used to obtain VO-specific information on Grid resources

```
lcg-infosites --vo <vo> <option> -v <verbosity> -f <site> --is <bdi>
```

--vo <vo>: the name of the VO to which the information to print is related (mandatory);

<option>: specifies what information has to be printed. It can take the following values: **ce**, **se**, **all**, **closeSE**, **tag**, **lfc**, **lfcLocal**, **rb**, **dli**, **dliLocal**, **vobox**, **fts**

--is <bdi>: the BDII to query. If not specified, the BDII defined in the environment variable **LCG_GFAL_INFOSYS** will be queried.

-f <site>: restricts the information printed to the specified site (it applies only to options **rb**, **dli**, **vobox** and **fts**).

Obtaining information about CE's

```
[ui03] /home/liplisbon/goncalo > export LCG_GFAL_INFOSYS=i2g-ii01.lip.pt:2170
```

```
[ui03] /home/liplisbon/goncalo > echo $LCG_GFAL_INFOSYS
i2g-ii01.lip.pt:2170
```

```
[ui03] /home/liplisbon/goncalo > lcg-infosites --vo itut ce (--is i2g-ii01.lip.pt)
```

#CPU	Free	Total Jobs	Running	Waiting	ComputingElement
22	22	0	0	0	ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-itut
19	11	0	0	0	ce.i2g.cesga.es:2119/jobmanager-lcgpbs-itutgrid
20	7	0	0	0	ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs-itut
60	40	0	0	0	i2g-ce01.lip.pt:2119/jobmanager-lcgsge-itutgridsdj

Try with options: **-v 0** , **-v 1**, **-v 2**

Obtaining information about SE's

```
[ui03] /home/liplisbon/goncalo > lcg-infosites --vo itut se
```

Avail Space(Kb)	Used Space(Kb)	Type	SEs
105347036	4905124	n.a	se-ieg.bifi.unizar.es
3358560	7163032	n.a	se.i2g.cesga.es
13570000000	16420000000	n.a	dpm.cyf-kr.edu.pl
1125000376	6577596	n.a	i2gse01.ifca.es
491960000	n.a	n.a	i2gse.ui.savba.sk
1369635410	202322620	n.a	dcache01.lip.pt
1369635410	202322620	n.a	dcache01.lip.pt
72955452	59984	n.a	i2g-se01.lip.pt
4378670000	301310000	n.a	se1.egee.man.poznan.pl

lcg-infosites: Exercise 3

Obtaining information about services

```
[ui03] /home/liplisbon/goncalo > lcg-infosites --vo itut rb --is i2g-ii01.lip.pt  
i2g-rb02.lip.pt:7772
```

```
[ui03] /home/liplisbon/goncalo > lcg-infosites --vo itut lfc --is i2g-ii01.lip.pt  
lfc01.lip.pt
```

Obtaining information about software tags

```
[ui03] /home/liplisbon/goncalo > lcg-infosites --vo atlas tag --is ii02.lip.pt  
[....]  
Name of the CE: ce02.lip.pt  
V0-atlas-production-12.0.31  
V0-atlas-production-12.0.6  
V0-atlas-production-12.0.7  
V0-atlas-production-12.0.95  
V0-atlas-production-13.0.20  
V0-atlas-production-13.0.30  
V0-atlas-production-13.0.30.1
```

- ❑ Query the BDII, in a deeper way, the BDII
 - ▶ It requires to set **LCG_GFAL_INFOSYS** to the value of the BDII to be used:
 - ▶ **export LCG_GFAL_INFOSYS=i2g-ii01.lip.pt:2170**

- ❑ Prints the list of the **CE's**, **SE's**, **services** or **sites** satisfying a given query

```
lcg-info --list-ce [--bdii bdii] [--vo vo] [--sed] [--debug] [--query query] [--attrs list]
--list-se
--list-service
--list-site

lcg-info -list-attrs
lcg-info -help
```

The query syntax is the following:

attr1 op1 value1, ... , attrN opN valueN

where **attrN** is an attribute name, **op** is **=**, **>=** or **<=**, and the cuts are AND'ed. The cuts are comma-separated and spaces are not allowed.

List all the attributes

```
[ui03] /home/liplisbon/goncalo > lcg-info -list-attrs
```

Attribute name	Glue object class	Glue attribute name
[.....]		
TotalCPUs	GlueCE	GlueCEInfoTotalCPUs
MaxRunningJobs	GlueCE	GlueCEPolicyMaxRunningJobs
MaxCPUTime	GlueCE	GlueCEPolicyMaxCPUTime
[.....]		
Accesspoint	GlueCESEBind	GlueCESEBindCEAccesspoint
Root	GlueSA	GlueSARoot
AvailableSpace	GlueSA	GlueSAStateAvailableSpace
[.....]		
ArchType	GlueSL	GlueSLArchitectureType
ServiceID	GlueService	GlueServiceUniqueID
ServiceOwner	GlueService	GlueServiceOwner
ServiceURI	GlueService	GlueServiceURI
ServiceType	GlueService	GlueServiceType
[.....]		
SiteSupport	GlueSite	GlueSiteUserSupportContact
SiteName	GlueSite	GlueSiteName
SiteAdmin	GlueSite	GlueSiteSysAdminContact
[.....]		
SMPSize	GlueSubCluster	GlueHostArchitectureSMPSize
Processor	GlueSubCluster	GlueHostProcessorModel
OSVersion	GlueSubCluster	GlueHostOperatingSystemVersion
[.....]		
VORunningJobs	GlueVOView	GlueCEStateRunningJobs
VOTotalJobs	GlueVOView	GlueCEStateTotalJobs
VOWorstRespTime	GlueVOView	GlueCEStateWorstResponseTime

Listing some CE attributes: **TotalCPUs,FreeCPUs,OS**

```
[ui03] /home/liplisbon/goncalo > lcg-info -list-ce --attrs TotalCPUs,FreeCPUs,OS
[.....]
- CE: ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs-itut
  - TotalCPUs      20
  - FreeCPUs       13
  - OS             Scientific Linux CERN

- CE: ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs-iusct
  - TotalCPUs      20
  - FreeCPUs       13
  - OS             Scientific Linux CERN

- CE: ce02.lip.pt:2119/jobmanager-lcgsgs-atlasgrid
  - TotalCPUs      48
  - FreeCPUs       6
  - OS             ScientificCERNSLC
```

Listing some SE attributes with a filter

```
[ui03] /home/liplisbon/goncalo > lcg-info -list-se --attrs SE,SESite,AvailableSpace,UsedSpace
--vo itut --query 'SE=*lip.pt*'
- SE: dcache01.lip.pt
  - SE                dcache01.lip.pt
  - SESite            LIP-Lisbon
  - AvailableSpace    1369635410
  - UsedSpace         202322620
- SE: i2g-se01.lip.pt
  - SE                i2g-se01.lip.pt
  - SESite            LIPI2G-Lisbon
  - AvailableSpace    72955452
  - UsedSpace         59984
```

List all sites

```
[ui03] /home/liplisbon/goncalo > lcg-info -list-site --attrs SiteLocation
- Site: BIFI
  - SiteLocation      Zaragoza, Spain
- Site: CESGA-I2G
  - SiteLocation      Santiago de Compostela, Spain
- Site: CYFRONET-I2G
  - SiteLocation      Cracow, Poland
- Site: FZK-I2G
  - SiteLocation      Karlsruhe, Germany
- Site: ICM
  - SiteLocation      Warsaw, Poland
- Site: IFCA-I2G
  - SiteLocation      Santander, Spain
- Site: IISAS-I2G
  - SiteLocation      Bratislava, Slovakia
- Site: LIP-Lisbon
  - SiteLocation      Lisboa, Portugal
- Site: LIPI2G-Lisbon
  - SiteLocation      Lisboa, Portugal
- Site: PSNC-I2G
  - SiteLocation      Poznan, Poland
```

List services of a given type: Resource Broker

```
[ui03] /home/liplisbon/goncalo > lcg-info -list-service --attrs ServiceName,ServiceEndpoint  
--query 'ServiceType=ResourceBroker'
```

```
- Service: i2g-rb02.lip.pt:7772  
- ServiceName      LIPI2G-Lisbon-rb  
- ServiceEndpoint  i2g-rb02.lip.pt:7772  
  
- Service: rb02.lip.pt:7772  
- ServiceName      LIP-Lisbon-rb  
- ServiceEndpoint  rb02.lip.pt:7772
```

- ❑ The LDAP servers can be queried directly with the **ldapsearch** command.
- ❑ For the LDAP implementation of the GLUE schema, the root of the Directory Information Tree (DIT) is always: **o=grid**.
- ❑ At the GRIS level the next entry is either:
 - ▶ **mds-vo-name=resource** (for BDII service port 2170).
- ❑ At the GIIS (site) level:
 - ▶ **mds-vo-name=<SITE NAME>** (BDII service port 2170)
- ❑ At top level:
 - ▶ **mds-vo-name=local** (BDII service port 2170)

ldapsearch: Exercise 1

Information published by a GRIS

```
[ui03] /home/liplisbon/goncalo > ldapsearch -x -H ldap://ce02.lip.pt:2170 -b mds-vo-  
name=resource,o=grid  
version: 2  
#  
# filter: (objectclass=*)  
# requesting: ALL  
#  
# ce02.lip.pt:2119/jobmanager-lcgsge-atlasgrid, local, grid  
dn: GlueCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsge-atlasgrid,mds-vo-name=lo  
cal,o=grid  
objectClass: GlueCETop  
objectClass: GlueCE  
objectClass: GlueSchemaVersion  
objectClass: GlueCEAccessControlBase  
objectClass: GlueCEInfo  
objectClass: GlueCEPolicy  
objectClass: GlueCEState  
objectClass: GlueInformationService  
objectClass: GlueKey  
GlueCEHostingCluster: ce02.lip.pt  
GlueCEName: atlasgrid  
GlueCEUniqueID: ce02.lip.pt:2119/jobmanager-lcgsge-atlasgrid  
GlueCEInfoGatekeeperPort: 2119  
GlueCEInfoHostName: ce02.lip.pt  
GlueCEInfoLRMSType: sge  
[.....]
```

ldapsearch: Exercise 2

Information published by a GRIS: BDII

```
[david@ui01 david]$ ldapsearch -x -H ldap://pprod06.lip.pt:2170 -b mds-vo-name=resource,o=grid  
version: 2
```

```
#  
# filter: (objectclass=*)  
# requesting: ALL  
#  
# pprod06.lip.pt:2170/bdii-top, resource, grid  
dn: GlueServiceUniqueID=pprod06.lip.pt:2170/bdii-top,Mds-Vo-name=resource,o=gr  
id  
objectClass: GlueTop  
objectClass: GlueService  
objectClass: GlueKey  
objectClass: GlueSchemaVersion  
GlueServiceUniqueID: pprod06.lip.pt:2170/bdii-top  
GlueServiceName: pprod06.lip.pt-top  
GlueServiceType: bdii_top  
GlueServiceVersion: 3.9.0  
GlueServiceEndpoint: ldap://pprod06.lip.pt:2170/mds-vo-name=local,o=grid  
GlueServiceURI: pprod06.lip.pt:2170/mds-vo-name=local,o=grid  
GlueServiceAccessPointURL: ldap://pprod06.lip.pt:2170/mds-vo-name=local,o=grid  
GlueServiceStatus: OK  
[.....]
```

Ldapsearch: Exercise 2

Information published by a site GII: BDII

```
[ui03] /home/liplisbon/goncalo > ldapsearch -x -H ldap://site-bdii.lip.pt:2170 -b mds-vo-  
name=LIP-Lisbon,o=grid  
version: 2  
#  
# filter: (objectclass=*)  
# requesting: ALL  
#  
# LIP-Lisbon, grid  
dn: Mds-Vo-name=LIP-Lisbon,o=grid  
objectClass: GlueTop  
  
# ce02.lip.pt:2119/jobmanager-lcgsge-atlasgrid, LIP-Lisbon, grid  
dn: GlueCEUniqueID=ce02.lip.pt:2119/jobmanager-lcgsge-atlasgrid,Mds-Vo-name=LI  
P-Lisbon,o=grid  
objectClass: GlueCETop  
objectClass: GlueCE  
objectClass: GlueSchemaVersion  
objectClass: GlueCEAccessControlBase  
objectClass: GlueCEInfo  
objectClass: GlueCEPolicy  
objectClass: GlueCEState  
objectClass: GlueInformationService  
objectClass: GlueKey  
GlueCEHostingCluster: ce02.lip.pt  
GlueCEName: atlasgrid  
[.....]
```

Information published by a Top level BDII

```
[ui03] /home/liplisbon/goncalo > ldapsearch -x -H ldap:///i2g-ii01.lip.pt:2170 -b mds-vo-  
name=local,o=grid  
version: 2  
  
#  
# filter: (objectclass=*)  
# requesting: ALL  
#  
  
# local, grid  
dn: mds-vo-name=local,o=grid  
objectClass: GlueTop  
  
# BIFI, local, grid  
dn: mds-vo-name=BIFI,mds-vo-name=local,o=grid  
objectClass: GlueTop  
  
# ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-imaing, BIFI, local, grid  
dn: GlueCEUniqueID=ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-imaing,mds-vo-n  
ame=BIFI,mds-vo-name=local,o=grid  
[.....]
```



Job management

Introduction to job management

Gonçalo Borges, (Jorge Gomes), Mario David, LIP

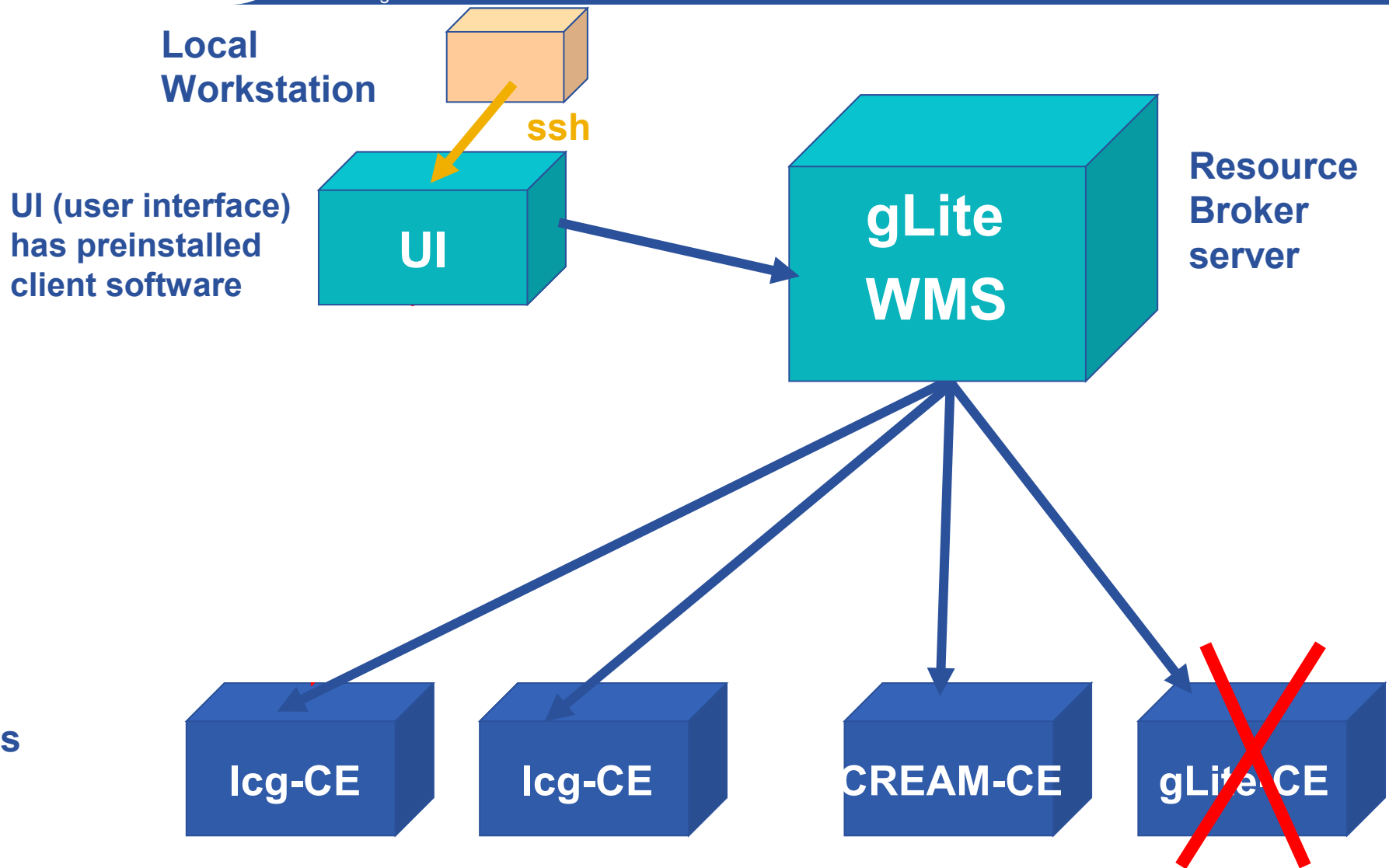
With thanks for some slides to EGEE, int.eu.grid and Globus colleagues

- Simple direct job submission
- What is the Workload Management System and Resource Brokers?
- How do you use it?
- Further information
- Practicals



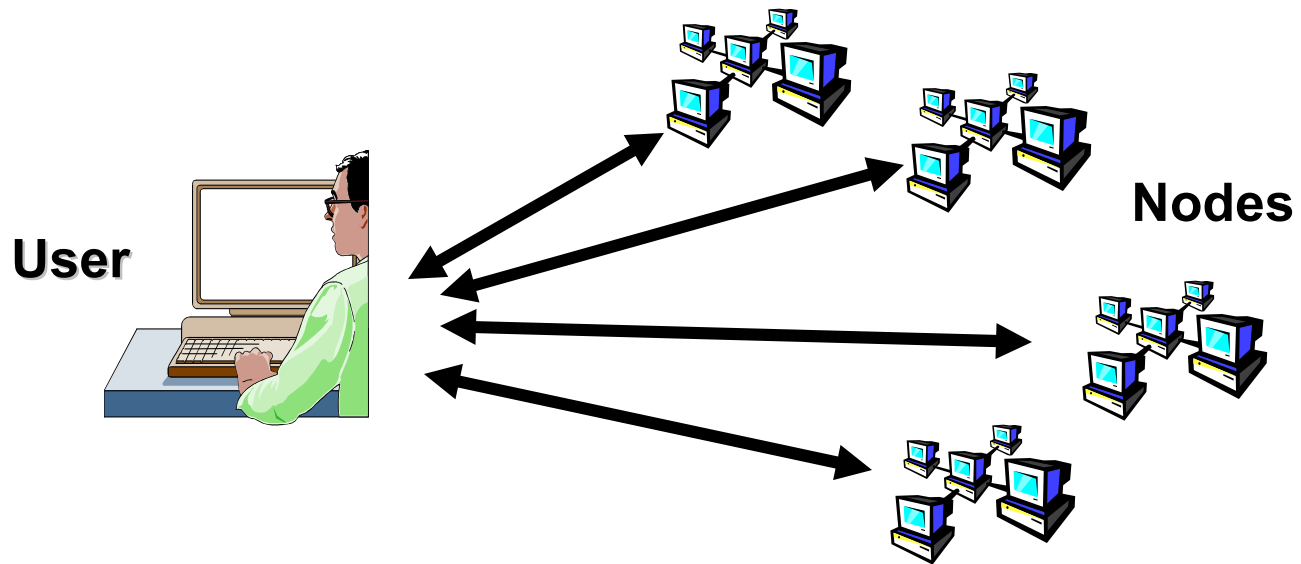
- **The most deployed type of Computing Element in EGEE is the so called lcg-CE**
 - Based on the globus gatekeeper
 - Enhanced by the CERN LCG project
 - Further enhanced and maintained by EGEE and included in gLite
- **Other computing elements include**
 - CREAM a web services based CE fully independent from globus
- **It is possible to submit jobs directly to the lcg-CE using the globus utilities available in the User Interface**





- **Direct job submission to the CE in gLite infrastructures may not be support in the future**
- **CLI tools for direct job submission to CREAM CEs will be provided as well (ICE) → ICE-CREAM**
- **Direct job submission does not fit in the gLite job management philosophy**
 - gLite philosophy is based on job submission through the Resource Brokers or WMS
- **But it may be useful for management and test purposes**
 - It is faster
 - Uses less levels of middleware services

- Uses the Globus GRAM services only
- Steps:
 - Find a computing element matching your needs and a local queue were to submit the job
 - Make sure you have a valid proxy
 - Use the **globus-job-run** for synchronous job execution or the **globus-job-submit** to get the prompt back
 - When using the **globus-job-submit** you get back a job identifier that can be used to:
 - check the job status with **globus-job-status**
 - cancel the job with **globus-job-cancel**
 - retrieve the job output with **globus-job-get-output**



- **Without a RB/WMS, need direct interaction with nodes**
 - Need to know resource addresses, capabilities, availability ...
- *Usually* want a higher level abstraction – submit a job to a Grid not to a CE

- **Why do the RB/WMS exist?**
- **Grids have**
 - Many users
 - Running many jobs – a “job” = an executable you want to run
 - Where many compute nodes are available
 - Different types of Computing Elements
 - Users need to specify complex job requirements
 - It is a software service that makes running jobs easier for the user
- **It builds on the basic grid services**
 - E.g. Authorisation, Authentication, Security, Information Systems, Job submission
- **Terminology: “Compute element”: defined as a batch queue - One cluster can have many queues**

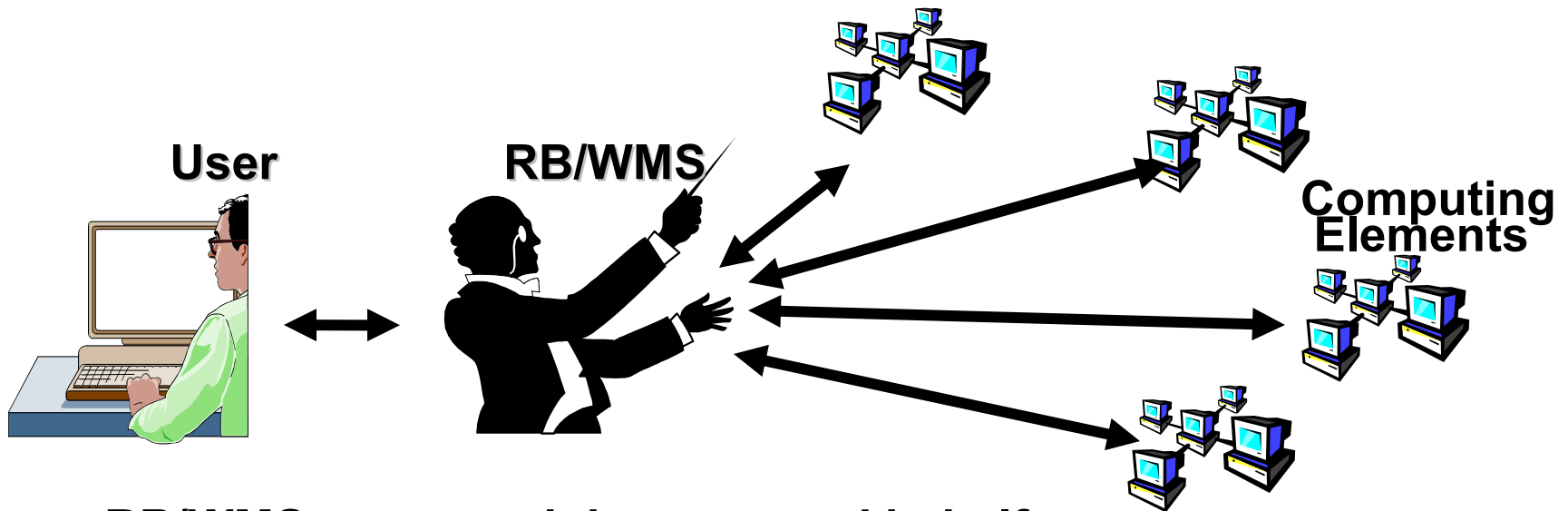
Which CE do you want to use?

- Without an RB/WMS, you have to use the Information System to see what's available, then choose...

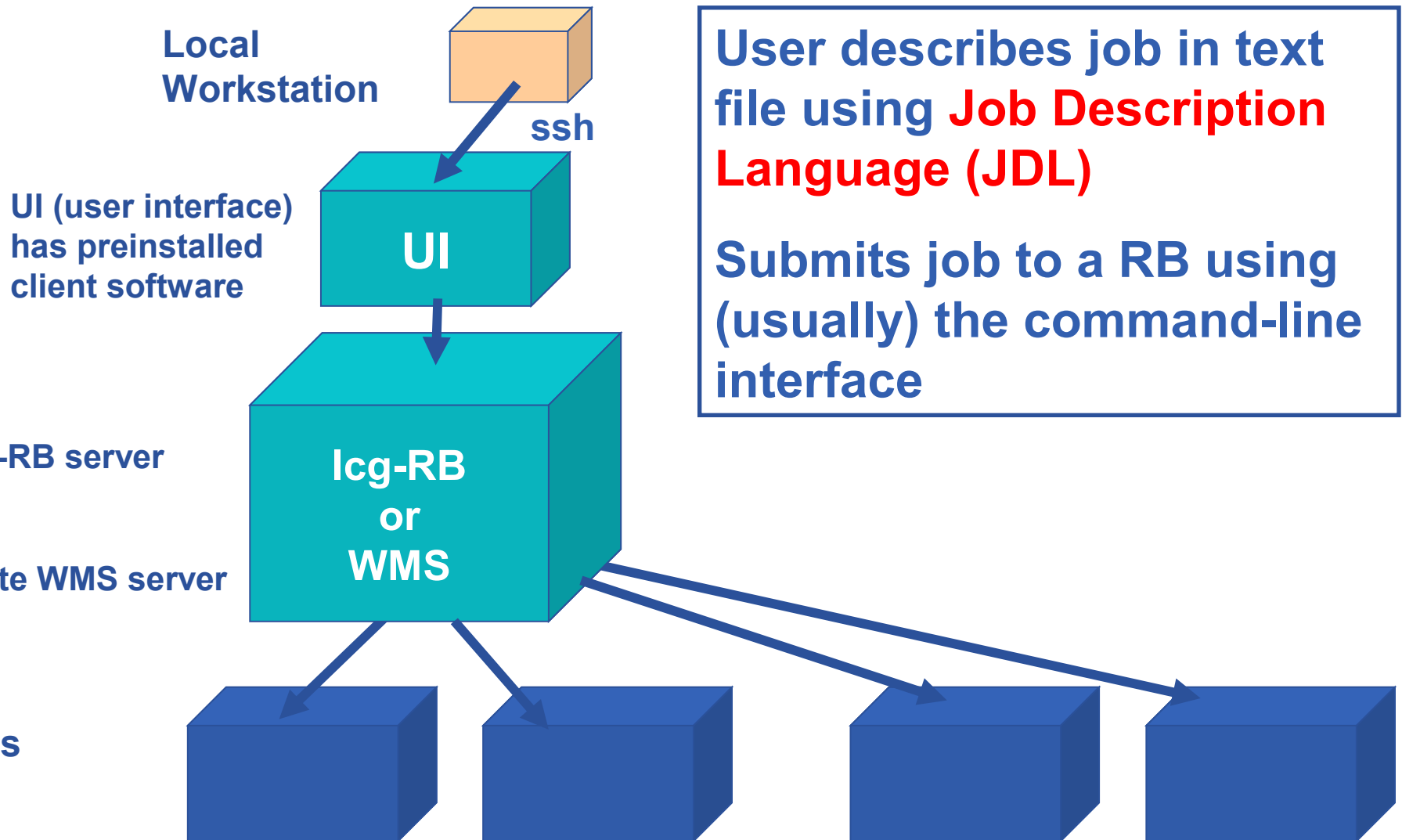
lcg-infosites --vo itut ce

#CPU	Free	Total	Jobs	Running	Waiting	Computing	Element
22	22	0	0	0			ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-itut
20	20	0	0	0			ce.i2g.cesga.es:2119/jobmanager-lcgpbs-itutgrid
20	11	0	0	0			ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs-itut
350	293	0	0	0			i2gce01.ifca.es:2119/jobmanager-lcgpbs-itut
32	32	0	0	0			i2gce.ui.savba.sk:2119/jobmanager-pbs-itut
60	22	0	0	0			i2g-ce01.lip.pt:2119/jobmanager-lcgsge-itutgridsdj

- The RB/WMS does this for you!
 - chooses CE for each job, balances workload, manages jobs and their files



- **RB/WMS manages jobs on users' behalf**
 - User doesn't decide where jobs are run
 - User defines the job and its requirements, RB/WMS matches this with available CEs
- **Effect:**
 - Easier submission
 - Users insulated from change in Compute elements
 - RB/WMS – can optimise your jobs – e.g. which CE?



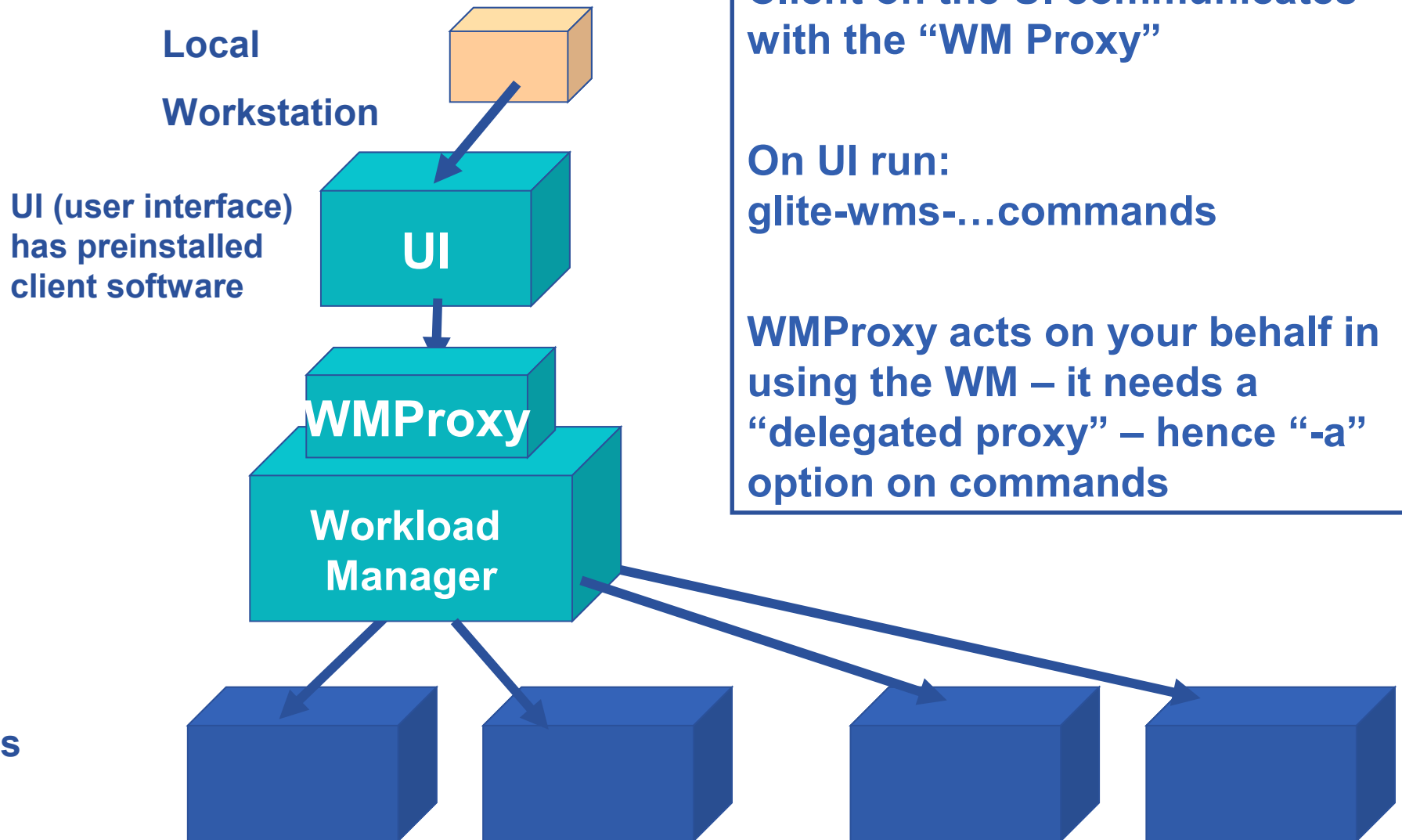
- Steps in running a job on a gLite grid with a lcg-RB:
- Create a text file in “Job Description Language”
- Optional check: list the compute elements that match your requirements (“**edg-job-list-match myfile.jdl**” command)
- Submit the job ~ “**edg-job-submit myfile.jdl**”
Non-blocking - Each job is given an id.
- Occasionally check the status of your job (“**edg-job-status**” command)
- When “Done” retrieve output (“**edg-job-get-output**” command)
- Or just cancel the job (“**edg-job-cancel**” command)

- Steps in running a job on a gLite grid with WMS:
- Create a text file in “Job Description Language”
- Optional check: list the compute elements that match your requirements (“**glite-wms-job-list-match myfile.jdl**” command)
- Submit the job ~ “**glite-wms-job-submit myfile.jdl**”
Non-blocking - Each job is given an id.
- Occasionally check the status of your job (“**glite-wms-job-status**” command)
- When “Done” retrieve output (“**glite-wms-job-output**” command)
- Or just cancel the job (“**glite-wms-job-cancel**” command)

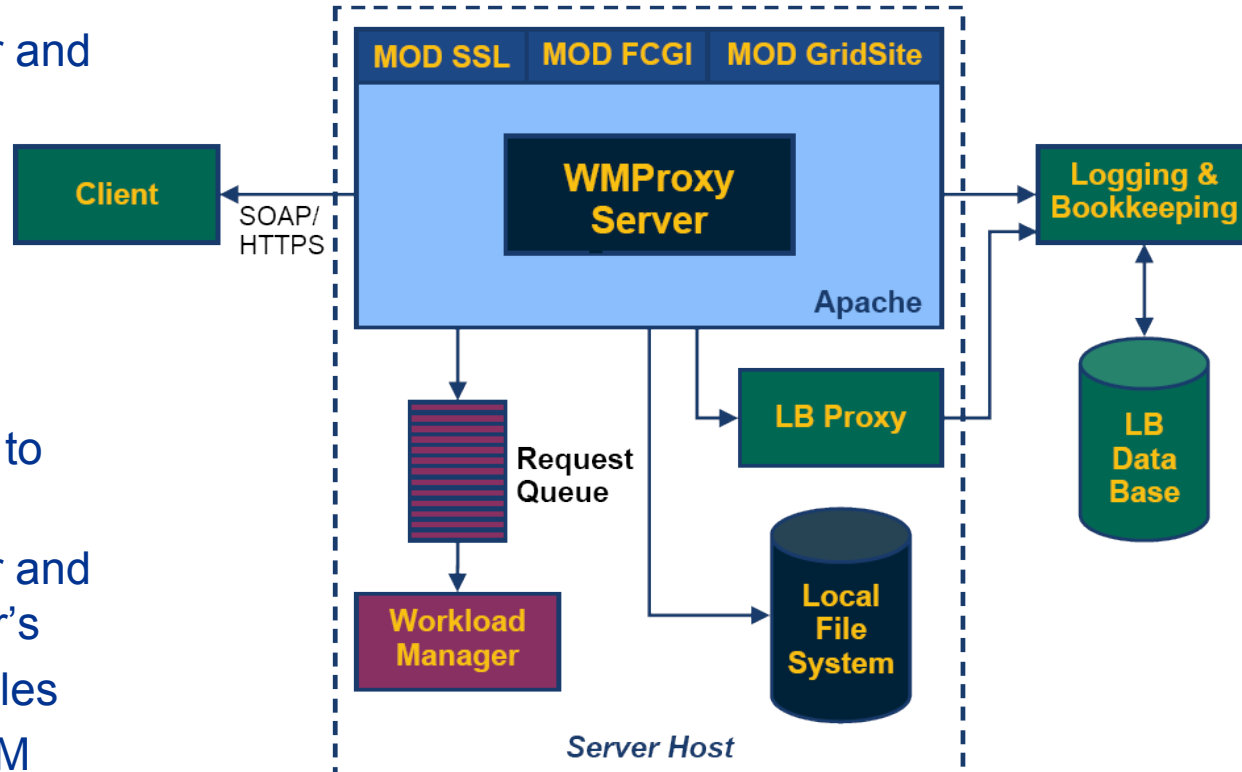
- **Executable** – sets the name of the executable file;
- **Arguments** – command line arguments of the program;
- **StdOutput, StdError** - files for storing the standard output and error messages output;
- **InputSandbox** – set of input files needed by the program, including the executable;
- **OutputSandbox** – set of output files which will be written during the execution, including standard output and standard error output; these are sent from the CE to the WMS for you to retrieve
- **ShallowRetryCount** – in case of grid error, retry job this many times (“Shallow”: before job is running)

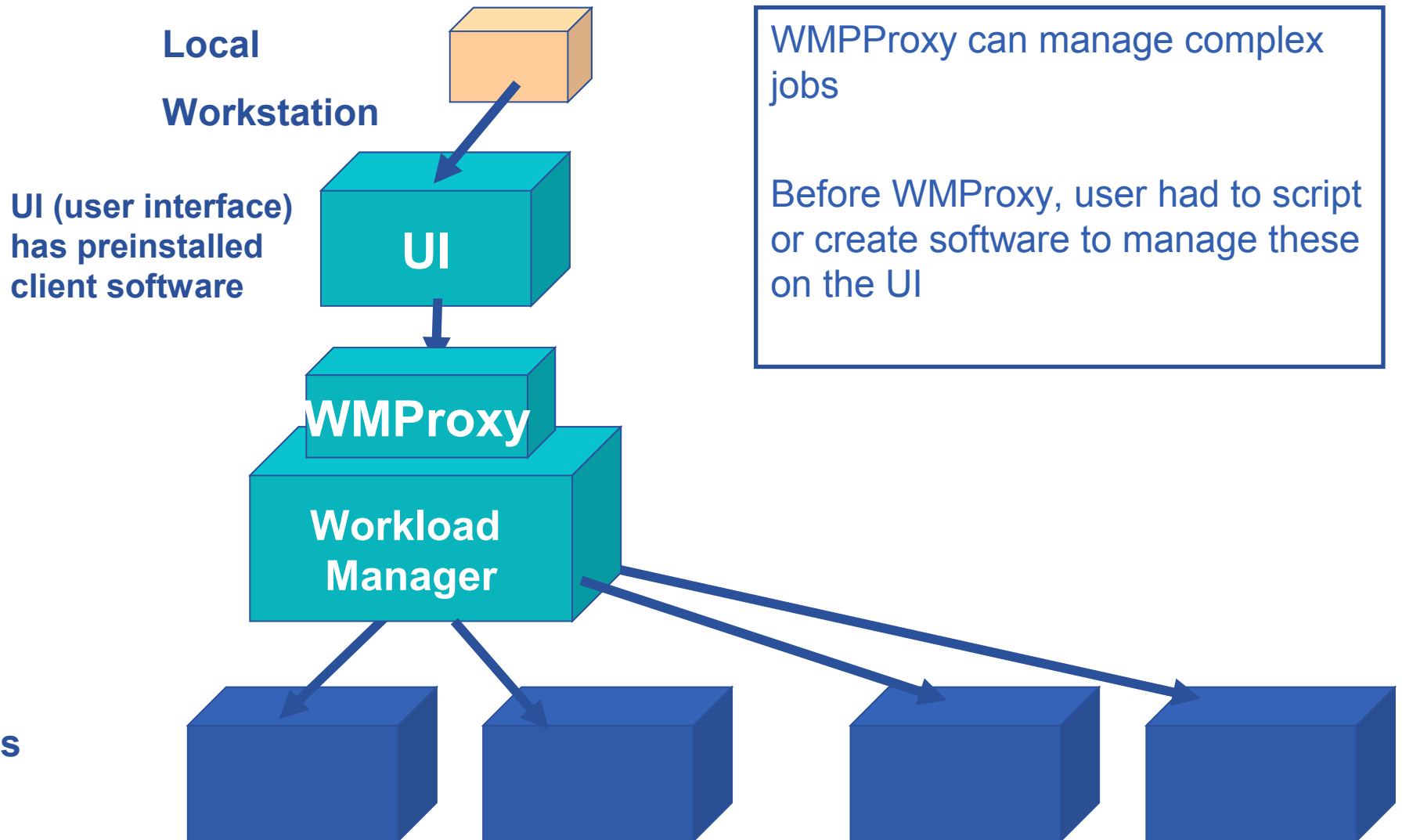
```
Executable = “gridTest”;  
StdError = “stderr.log”;  
StdOutput = “stdout.log”;  
InputSandbox = {“/home/joda/test/gridTest”};  
OutputSandbox = {“stderr.log”, “stdout.log”};  
Requirements = other.GlueCEPolicyMaxCPUTime > 480;  
RetryCount = 3;
```

Flag	Meaning
SUBMITTED	submission logged in the Logging & Bookkeeping service
WAIT	job match making for resources
READY	job being sent to executing CE
SCHEDULED	job scheduled in the CE queue manager
RUNNING	job executing on a Worker Node of the selected CE queue
DONE	job terminated without grid errors
CLEARED	job output retrieved
ABORT	job aborted by middleware, check <i>reason</i>

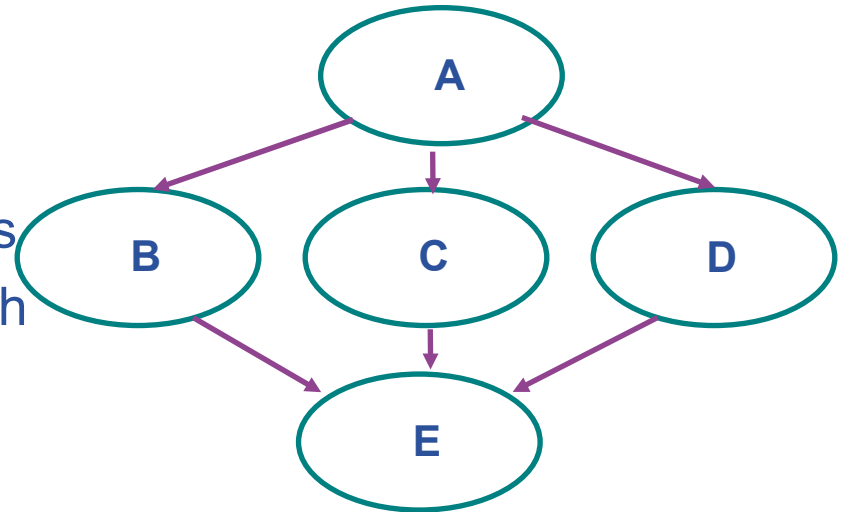


- **WMPProxy is a SOAP Web service providing access to the Workload Management System (WMS)**
- **Job characteristics specified via JDL**
 - **jobRegister**
 - create id
 - map to local user and create job dir
 - register to L&B
 - return id to user
 - **input files transfer**
 - **jobStart**
 - register sub-jobs to L&B
 - map to local user and create sub-job dir's
 - unpack sub-job files
 - deliver jobs to WM

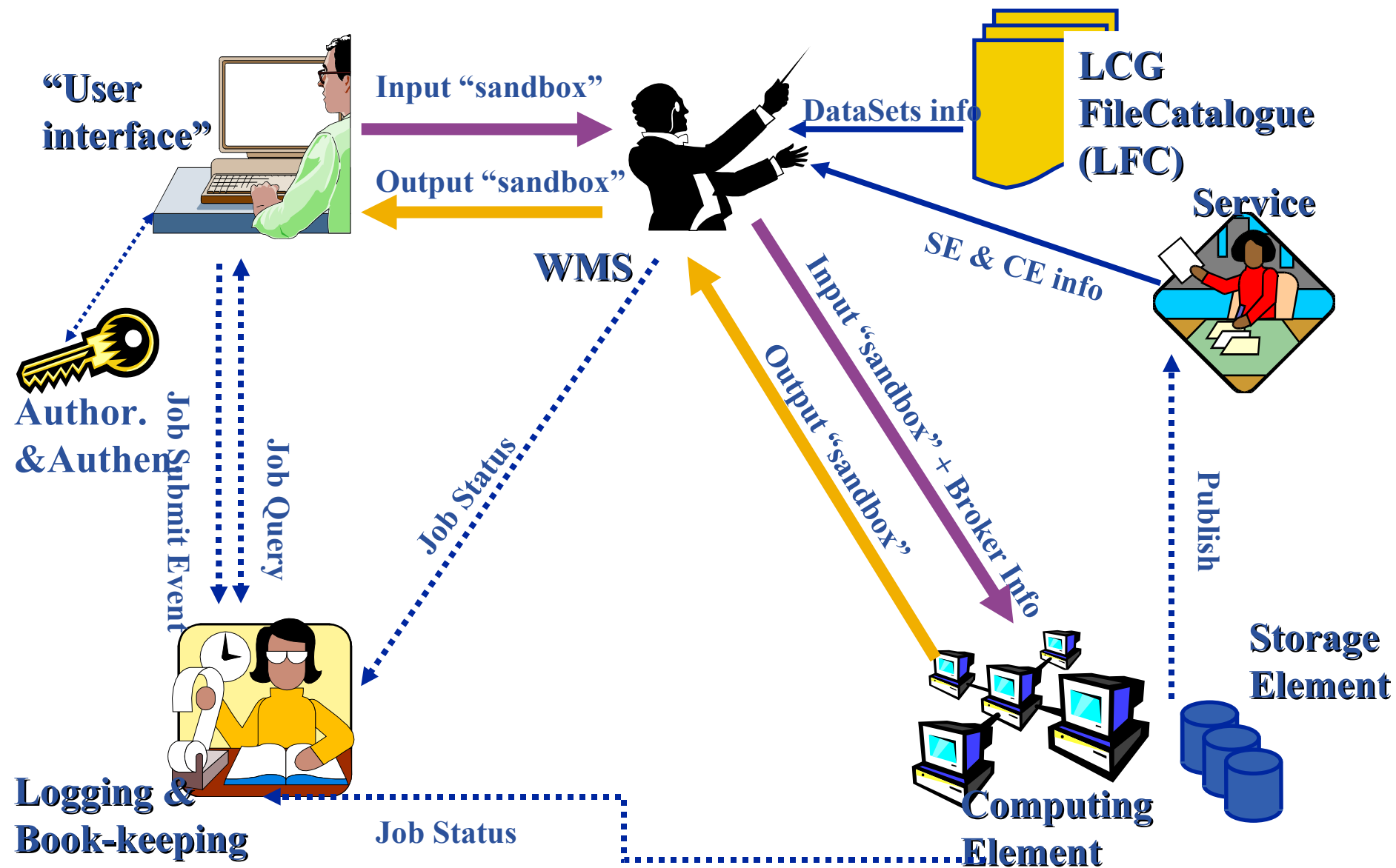




- **Direct Acyclic Graph (DAG)** is a set of jobs where the input, output, or execution of one or more jobs depends on one or more other jobs
- **A Collection** is a group of jobs with no dependencies
 - basically a collection of JDL's
 - Can have common sandbox
- **A Parametric job** is a job having one or more attributes in the JDL that vary their values according to parameters
- **It is possible to have one shot submission of a (possibly very large, up to thousands) group of jobs**
 - Submission time reduction
 - Single call to WMPProxy server
 - Single Authentication and Authorization process
 - Sharing of files between jobs
 - Availability of both a single Job Id to manage the group as a whole and an Id for each single job in the group



- For simple jobs: **glite-wms-...** becoming the recommended way to use the WMS
- History:
 - Before the **glite-wms-** commands we had **glite-** commands
 - used with the old lcg-RB (without WMPProxy)
 - Before the **glite-** commands we had
 - **edg-** commands (edg-job-submit....)
 - *European Data Grid – project before EGEE*
 - Used the “resource broker”
 - Still very widely used
 - You might see these commands still in use.
- Status
 - Complex jobs with WMPProxy: not yet in routine production use
 - Watch for news!



- **gLite Users Guide**
 - Follow <http://www.glite.org> and “Documentation”
- **About JDL and grid**
 - <https://edms.cern.ch/file/555796/1/EGEE-JRA1-TEC-555796-JDL>
 - <http://www.grid.org.tr/servisler/dokumanlar/DataGrid-JDL-HowTo>.
- **GILDA wiki**
 - We are using some of these pages
 - <https://grid.ct.infn.it/twiki/bin/view/GILDA/>
- **EGEE Digital Library** <http://egee.lib.ed.ac.uk/>



Interactive European Grid

Job Management Exercises

Gonçalo Borges, (Jorge Gomes), Mario David
LIP

I2G training, Porto, 15th-May-2008

- ❑ The Int.Eu.Grid (I2G) job management is based on the LCG Resource Broker (CrossBroker) with enhancements
 - ▶ Support for MPI jobs inside clusters (OpenMPI)
 - ▶ Support for MPI jobs across clusters (PACX-MPI)
 - ▶ Support for interactivity
- ❑ During this tutorial the Int.Eu.Grid infrastructure will be used for the exercises including the I2G CrossBroker
- ❑ The Int.Eu.Grid User Interface is fully compatible with the gLite UI however the job submission commands start with the prefix **i2g**
 - ▶ **i2g-job-submit, i2g-job-status, i2g-job-cancel, i2g-job-get-output, i2g-job-list-match**

Exercise 1

- ❑ Simple remote execution with Globus
- ❑ Select a Computing Element

lcg-infosites -vo itut ce

#CPU	Free	Total	Jobs	Running	Waiting	ComputingElement
22	22	0	0	0		ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-itut
19	13	0	0	0		ce.i2g.cesga.es:2119/jobmanager-lcgpbs-itutgrid
20	11	0	0	0		ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs-itut
350	279	0	0	0		i2gce01.ifca.es:2119/jobmanager-lcgpbs-itut
32	30	0	0	0		i2gce.ui.savba.sk:2119/jobmanager-pbs-itut
60	16	0	0	0		i2g-ce01.lip.pt:2119/jobmanager-lcgsge-itutgridsdj

- ❑ Use the command globus-job-run

**globus-job-run **
**ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs **
**-q itut **
/bin/uname -a

- This command is actually a wrapper that
 - ▶ produces a globus RSL job description
 - ▶ submits it using other globus commands

```
globus-job-run -dumprsl \  
ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs \  
-q itut \  
/bin/uname -a
```

```
&(executable="/bin/uname")  
(queue="itut")  
(arguments= "-a")
```

❑ Job execution with globus

```
globus-job-submit \  
    ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs \  
    -q itut \  
    /bin/uname -a
```

❑ Returns

<https://ce.i2g.cyf-kr.edu.pl:20005/19369/1193331791/>

❑ Check the job status with

```
globus-job-status https://ce.i2g.cyf-kr.edu.pl:20005/19369/1193331791/
```

DONE

❑ Get the job output with

```
globus-job-get-output \  
    https://ce.i2g.cyf-kr.edu.pl:20005/19369/1193331791/
```

Linux i2gwn16.ui.savba.sk 2.4.21-47.EL.cernsmp #1 SMP Mon Jul 24 15:33:59 CES

- ❑ The previous job could have been cancelled with the command

```
globus-job-cancel https://ce.i2g.cyf-kr.edu.pl:20005/19369/1193331791/
```

- ❑ In the end you should cleanup the job files at the remote end with

```
globus-job-clean https://ce.i2g.cyf-kr.edu.pl:20005/19369/1193331791/
```

More about globus jobs

- ❑ Other examples using globus

```
globus-job-run ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs \  
-q itut /bin/sh -c "cd /tmp; pwd"
```

- ❑ Stage and execute a script or binary file

```
globus-job-run ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs \  
-q itut -s myshellscript.sh
```

- ❑ What globus actually does

```
globusrun -s -r ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs  
'&(executable=$(GLOBUSRUN_GASS_URL) # "/home/tutorial/  
user01/myshellscript.sh")(queue="itut")'
```

- ❑ Simple job submission through a resource broker
- ❑ Need to create a file containing the job description using JDL (Job Description Language) language
- ❑ Create a file named **e3_1.jdl** with the following content

```
Type = "Job";  
JobType = "Normal";  
Executable = "/bin/hostname";  
StdOutput = "hostname.out";  
StdError = "hostname.err";  
OutputSandbox = {"hostname.err", "hostname.out"};  
Arguments = "-f";
```

❑ Check matching resources:

```
i2g-job-list-match -vo itut e3_1.jdl
```

Selected Virtual Organisation name (from proxy certificate extension): itut

Connecting to host i2g-rb01.lip.pt, port 7772

COMPUTING ELEMENT IDs LIST

The following CE(s) matching your job requirements have been found:

CEId

i2gce01.ifca.es:2119/jobmanager-lcgpbs-itut/itut

i2gce.ui.savba.sk:2119/jobmanager-pbs-itut/itut

i2g-ce01.lip.pt:2119/jobmanager-lcgsge-itutgridsdj/itut

ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs-itut/itut

ce.i2g.cesga.es:2119/jobmanager-lcgpbs-itutgrid/itut

ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-itut/itut

□ Submit the job:

```
i2g-job-submit e3_1.jdl
```

Selected Virtual Organisation name (from proxy certificate extension): itut

Connecting to host i2g-rb01.lip.pt, port 7772

Logging to host i2g-rb01.lip.pt, port 9002

JOB SUBMIT OUTCOME

The job has been successfully submitted to the Network Server.

Use i2g-job-status command to check job current status. Your job identifier (edg_jobId) is:

- <https://i2g-rb01.lip.pt:9000/Heot5Ro-5qI5jXdowY0ygw>

☐ Check job status:

i2g-job-status <https://i2g-rb01.lip.pt:9000/Heot5Ro-5ql5jXdowY0ygw>

BOOKKEEPING INFORMATION:

Status info for the Job : <https://i2g-rb01.lip.pt:9000/Heot5Ro-5ql5jXdowY0ygw>

Current Status: Scheduled

Status Reason: Job successfully submitted to Globus

Destination: i2g-ce01.lip.pt:2119/jobmanager-lcgsge-itutgridsdj

reached on: Wed Oct 24 16:00:01 2007

☐ You could cancel the job with: (don't do it !!!!)

i2g-job-cancel <https://i2g-rb01.lip.pt:9000/Heot5Ro-5ql5jXdowY0ygw>

- ❑ Check job status until it finishes (done state):

i2g-job-status <https://i2g-rb01.lip.pt:9000/Heot5Ro-5qI5jXdowY0ygw>

BOOKKEEPING INFORMATION:

Status info for the Job : <https://i2g-rb01.lip.pt:9000/Heot5Ro-5qI5jXdowY0ygw>

Current Status: **Done (Success)**

Exit code: 0

Status Reason: Job terminated successfully

Destination: i2g-ce01.lip.pt:2119/jobmanager-lcgsge-itutgridsdj

reached on: Wed Oct 24 16:06:18 2007

□ Get job output:

```
i2g-job-get-output https://i2g-rb01.lip.pt:9000/Heot5Ro-5ql5jXdowY0ygw
```

Retrieving files from host: i2g-rb01.lip.pt (for https://i2g-rb01.lip.pt:9000/Heot5Ro-5ql5jXdowY0ygw)

JOB GET OUTPUT OUTCOME

Output sandbox files for the job:

- https://i2g-rb01.lip.pt:9000/Heot5Ro-5ql5jXdowY0ygw

have been successfully retrieved and stored in the directory:

/tmp/jobOutput/jorge_Heot5Ro-5ql5jXdowY0ygw

```
ls /tmp/jobOutput/jorge_Heot5Ro-5ql5jXdowY0ygw
```

- ❑ Submit directly to a CE via the RB
- ❑ Bypasses the matchmaking
- ❑ Can be used for debugging or management purposes

```
i2g-job-submit -vo itut \  
    -r i2g-ce01.lip.pt:2119/jobmanager-lcgsgsge-itutgridsdj \  
    e3_1.jdl
```

- ❑ Remaining steps are the same

```
i2g-job-status jobid
```

```
i2g-job-get-output jobid
```

Exercise 5

- ❑ Submit a shell script and add requirements to the job
- ❑ Create a file named **e5_1.jdl** with the following content

```
Type = "Job";
JobType = "Normal";
Executable = "/bin/bash";
StdOutput = "e5.out";
StdError = "e5.err";
OutputSandbox = {"e5.err", "e5.out"};
InputSandbox = "e5uname.sh";
Arguments = "e5uname.sh";
Requirements = (
  other.GlueCEUniqueID == "i2g-ce01.lip.pt:2119/jobmanager-lcgsgc-itutgridsdj"
);
```

- ❑ Create a script named **e5uname.sh** with something like

```
#!/bin/bash
uname -a
```

❑ Check matching resources:

```
i2g-job-list-match -vo itut e5_1.jdl
```

Selected Virtual Organisation name (from proxy certificate extension): itut
Connecting to host i2g-rb01.lip.pt, port 7772

COMPUTING ELEMENT IDs LIST

The following CE(s) matching your job requirements have been found:

CEId

i2g-ce01.lip.pt:2119/jobmanager-lcgsgc-itutgridsdj

□ Submit the job:

```
i2g-job-submit -o jobid.dat e5_1.jdl
```

```
i2g-job-status -i jobid.dat
```

```
i2g-job-get-output -i jobid.dat
```

```
cat /tmp/jobOutput/jorge_iI0hZGXmcsRn-7wMCvjhwA/e2.out
```

Linux wn013.i2g.cesga.es 2.4.21-47.0.1.ELsmp #1 SMP Thu Oct 19 10:46:05 CDT 2006 i686 i686 i386
GNU/Linux

❑ JDL requirements can be powerful expressions

```
Type = "Job";
JobType = "Normal";
Executable = "/bin/bash";
StdOutput = "e5.out";
StdError = "e5.err";
OutputSandbox = {"e5.err","e5.out"};
InputSandbox = "e5uname.sh";
Arguments = "e5uname.sh";
requirements = (
  RegExp("lip.pt",other.GlueCEUniqueId) &&
  Member("GLITE-3_0_2",other.GlueHostApplicationSoftwareRunTimeEnvironment));
```

❑ The SoftwareRunTimeEnvironment contains tags of installed software or run time capabilities

```
ldapsearch -x -b mds-vo-name=local,o=grid -H ldap://i2g-ce01:2170 \
  GlueSubClusterUniqueID=i2g-ce01.lip.pt \
  'GlueHostApplicationSoftwareRunTimeEnvironment'
```

❑ JDL requirements can be powerful expressions

```
Executable = "/bin/date";  
StdOutput = "e5_6.out"; StdError = "e5_6.err";  
OutputSandbox = {"e5_6.err", "e5_6.out"};  
Requirements = !(RegExp("lip.pt", other.GlueCEUniqueID));
```

❑ Test it with

i2g-job-list-match e5_6.jdl

COMPUTING ELEMENT IDs LIST

The following CE(s) matching your job requirements have been found:

CEId

ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-itut
ce.i2g.cesga.es:2119/jobmanager-lcgpbs-itutgrid
ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs-itut
i2gce.ui.savba.sk:2119/jobmanager-pbs-itut
i2gce01.ifca.es:2119/jobmanager-lcgpbs-itut

❑ Sort matching nodes based on published information

```
Executable = "/bin/date";
StdOutput = "e5_7.out"; StdError = "e5_7.err";
OutputSandbox = {"e5_7.err", "e5_7.out"};
Rank = other. GlueCEStateFreeCPUs;
Requirements = other.GlueCEInfoLRMSType == "torque";
RetryCount = 7;
```

❑ Test it with

i2g-job-list-match e5_7.jdl
lcg-infosites --vo itut ce

i2gce01.ifca.es:2119/jobmanager-lcgpbs-itut
ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-itut
ce.i2g.cesga.es:2119/jobmanager-lcgpbs-itutgrid
ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs-itut

#CPU	Free	Total	Jobs	Running	Waiting	ComputingElement
22	22	0	0	0		ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-itut
20	20	0	0	0		ce.i2g.cesga.es:2119/jobmanager-lcgpbs-itutgrid
20	12	0	0	0		ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs-itut
350	210	0	0	0		i2gce01.ifca.es:2119/jobmanager-lcgpbs-itut
32	30	0	0	0		i2gce.ui.savba.sk:2119/jobmanager-pbs-itut
60	15	0	0	0		i2g-ce01.lip.pt:2119/jobmanager-lcgsgs-itutgridsdj

❑ Running my own executable and setting environment variables

```
Executable = "e5uname.sh";
StdOutput = "e5_8.out"; StdError = "e5_8.err";
OutputSandbox = {"e5_8.err", "e5_8.out"};
InputSandbox = "e5uname.sh";
Environment = {"JAVA_INSTALL_PATH=/usr/java/j2sdk1.4.2_11",
               "VO_ITUT_DEFAULT_SE=i2g-se01.lip.pt"};
Rank = ( other.GlueCEStateWaitingJobs == 0 ?
         other.GlueCEStateFreeCPUs : -other.GlueCEStateWaitingJobs);
Requirements =
    anyMatch(other.storage.CloseSEs, target.GlueSAStateAvailableSpace > 501000000);
```

- ❑ Notice that the executable bit is set by default by the jobwrapper
- ❑ Try to increase the storage requirement and do a list match to see what sites match the required storage



Interactive European Grid

Advanced job submission

Jobs with data requirements

- ❑ Requirements on data location can be defined for matchmaking purposes
 - ▶ Submit a job only to the sites where a certain file replica is stored at the “near SE”
- ❑ The JDL attributes for data requirements have changed from the LCG RB to the gLite WMS
 - ▶ Compatibility will be maintained for some time
- ❑ The examples here provided are based on the LCG RB notation
- ❑ Requirements are specified:
 - ▶ Based on files stored in SEs
 - ▶ Files must be registered on data catalogues

□ Lets create the LFC data catalogue entry

```
export LFC_HOST=`lcg-infosites --vo itut lfc|uniq`
```

```
lfc-mkdir /grid/itut/tut-14-11-07/${USER}
```

```
lcg-cr --vo itut -d i2g-se01.lip.pt \  
-l lfn:/grid/itut/tut-14-11-07/${USER}/a1mytest \  
file://`pwd`/a1mytest.dat
```

```
lcg-lr lfn:/grid/itut/tut-14-11-07/${USER}/a1mytest
```

□ Job with input data requirements

```
Type = "Job";  
JobType = "Normal";  
Executable = "a1.sh";  
StdOutput = "a1.out";  
StdError = "a1.err";  
OutputSandbox = {"a1.err","a1.out"};  
InputSandbox = "a1.sh";  
InputData = { "lfn:/grid/itut/tut-14-11-07/user01/a1mytest" };  
DataAccessProtocol = "gsiftp";
```

□ To verify the effect of the InputData requirement

i2g-job-list-match a1_2.jdl

- ❑ The “a1.sh” script should contain code to retrieve the InputData

```
#!/bin/sh
echo =getCE=; /opt/glite/bin/glite-brokerinfo getCE
echo =getDataAccessProtocol=; PROTO=`/opt/glite/bin/glite-brokerinfo getDataAccessProtocol`
echo $PROTO
echo =getInputData=; LFN=`/opt/glite/bin/glite-brokerinfo getInputData`
echo $LFN
echo =getSEs=; /opt/glite/bin/glite-brokerinfo getSEs
echo =getCloseSEs=; CLOSESE=`/opt/glite/bin/glite-brokerinfo getCloseSEs`
echo $CLOSESE
echo =getSEMountPoint=; /opt/glite/bin/glite-brokerinfo getSEMountPoint $CLOSESE
echo =getSEFreeSpace=; /opt/glite/bin/glite-brokerinfo getSEFreeSpace $CLOSESE
echo =getLFN2SFN=; /opt/glite/bin/glite-brokerinfo getLFN2SFN $LFN
echo =getSEProtocols=; /opt/glite/bin/glite-brokerinfo getSEProtocols $CLOSESE

echo COPY DATA FILE
lcg-cp $LFN file://`pwd`/myinputfile.dat
echo CAT DATA FILE
cat myinputfile.dat
```

□ State of the art MPI implementation

- ▶ Full support of the MPI-2 standard
- ▶ Full thread support
- ▶ Avoidance of old legacy code
- ▶ Profit from long experience in MPI implementations
- ▶ Avoiding the “forking” problem
- ▶ Community / 3rd party involvement
- ▶ Production-quality research platform
- ▶ Rapid deployment for new platforms

□ Compile with Open MPI

▶ Compilers location is:

■ /opt/i2g/openmpi/bin

```
/opt/i2g/openmpi/bin/mpicc -c cpip.c
```

```
/opt/i2g/openmpi/bin/mpicc -o cpip cpip.o -lm
```

□ Write a JDL file named “a2_2.jdl” containing

```
JobType           = "parallel";
SubJobType        = "openmpi";
NodeNumber        = 2;
VirtualOrganisation = "itut";
Executable        = "cpip";
StdOutput         = "cpip.out";
StdError          = "cpip.err";
InputSandbox      = {"cpip"};
OutputSandbox     = {"cpip.out","cpip.err"};
```

Exercise 2

- ❑ Submit the job with:

```
i2g-job-submit -o cpip.jobid a2_2.jdl
```

- ❑ Check the status until DONE

```
i2g-job-status -i cpip.jobid
```

- ❑ Get the job output

```
i2g-job-get-output -i cpip.jobid
```

- ❑ Check the output

```
cat /tmp/jobOutput/jorge_ruwKQvDW62LSBkVNCD0x2A/cpip.out
```

```
pi is approximately 3.1416009869231241, Error is 0.00000833333333309  
wall clock time = 0.010742
```

```
cat /tmp/jobOutput/jorge_ruwKQvDW62LSBkVNCD0x2A/cpip.err
```

```
Process 1 on i2gwn15.ui.savba.sk
```

```
Process 0 on i2gwn16.ui.savba.sk
```

- Using hooks to define pre and post execution tasks

```
JobType = "Parallel";  
SubJobType = "openmpi";  
NodeNumber = 8;  
VirtualOrganisation = "itut";  
Executable = "cpip";  
StdOutput = "std.out";  
StdError = "std.err";  
InputSandbox = {"cpip", "o3_hooks.sh", "input.8"};  
OutputSandbox = {"std.out", "std.err"};  
Environment = {"I2G_MPI_PRE_RUN_HOOK=./o3_hooks.sh",  
               "I2G_MPI_POST_RUN_HOOK=./o3_hooks.sh"};
```

□ o3_hooks.sh 1/2

```
#!/bin/sh
export OUTPUT_PATTERN=L8
export OUTPUT_ARCHIVE=output.tar.gz
export OUTPUT_HOST=iwrse2.fzk.de
export OUTPUT_SE=lfm:/grid/imain/sven
export OUTPUT_VO=imain

pre_run_hook () {
}

copy_from_remote_node() {
    if [[ $1 == `hostname` || $1 == 'hostname -f' || $1 == "localhost" ]]; then
        echo "skip local host"
        return 1
    fi
    # pack data
    CMD="scp -r $1:\ "$PWD/$OUTPUT_PATTERN" ."
    echo $CMD
    $CMD
}
```

□ o3_hooks.sh 2/2

```
...
post_run_hook () {
    echo "post_run_hook called"
    if [ "x$MPI_START_SHARED_FS" == "x0" ] ; then
        echo "gather output from remote hosts"
        mpi_start_foreach_host copy_from_remote_node
    fi
    ls al
    echo "pack the data"
    tar cvzf $OUTPUT_ARCHIVE $OUTPUT_PATTERN
    echo "upload the data"
    lcg-cr -vo $OUTPUT_VO -d $OUTPUT_HOST -I $OUTPUT_SE/$OUTPUT_ARCHIVE \
    file://$PWD/$OUTPUT_ARCHIVE
    return 0
}
```

- Debugging can be activated through environment variables

```
JobType           = "parallel";
SubJobType        = "openmpi";
NodeNumber        = 2;
VirtualOrganisation = "itut";
Executable        = "cpip";
StdOutput         = "cpip.out";
StdError          = "cpip.err";
InputSandbox      = {"cpip"};
OutputSandbox     = {"cpip.out", "cpip.err"};
Environment = {"MPI_START_VERBOSE=1",
               "MPI_START_DEBUG=1",
               "MPI_START_TRACE=1"};
```

□ Debugging support

▶ I2G_MPI_START_VERBOSE

- If set to 1 only very basic information is produced

▶ I2G_MPI_START_DEBUG

- If set to 1 information about the internal flow is produced

▶ I2G_MPI_START_TRACE

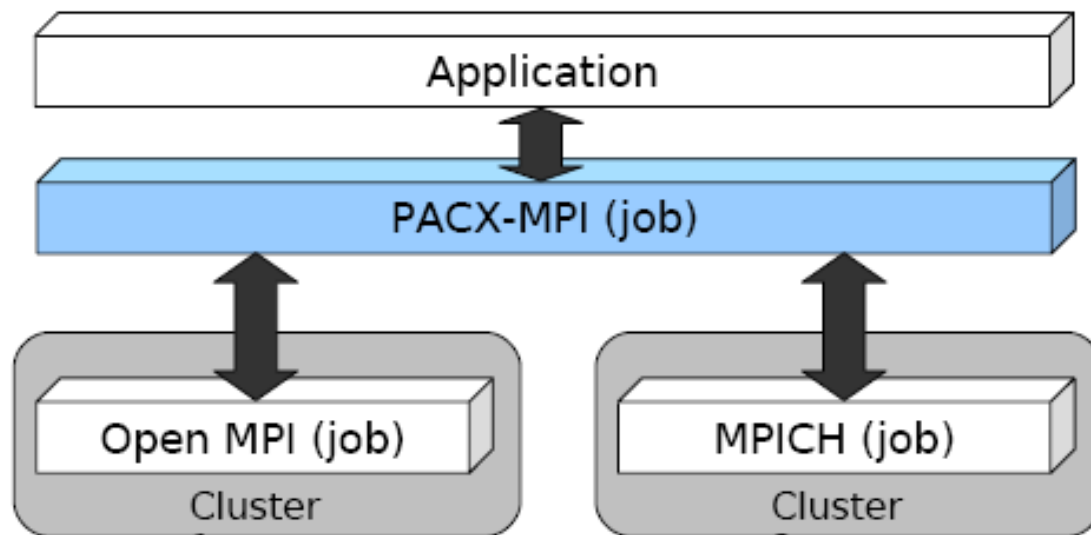
- If set to 1 that “set -x” is enabled at the beginning.

Parallel jobs using PACX-MPI

- ❑ A middleware for seamlessly running an MPI-application on a network of parallel computers
- ❑ originally developed in 1995 to connect Vector+MPP
- ❑ PACX-MPI is an optimized standard conforming MPI- implementation
- ❑ application just needs to be **recompiled**
- ❑ PACX-MPI uses locally installed, optimized vendor implementations for cluster inter communication

Parallel jobs using PACX-MPI

- ❑ PACX-MPI starts an MPI job in each cluster
- ❑ PACX-MPI “merges/manages” these MPI jobs internally and emulate transparently a bigger MPI job to the application



□ Compile with Open MPI

▶ Compilers location is:

■ /opt/i2g/pacx-openmpi/bin

```
/opt/i2g/pacx-openmpi/bin/ppacxcc \  
-o cpip-pacx cpip.c -lm
```

□ Write a JDL file “a3_1.jdl” containing

```
JobType           = "parallel";
SubJobType        = "pacx-mpi";
NodeNumber        = 22;
VirtualOrganisation = "itut";
Executable        = "cpip-pacx";
StdOutput         = "cpip-pacx.out";
StdError          = "cpip-pacx.err";
InputSandbox      = {"cpip-pacx"};
OutputSandbox     = {"cpip-pacx.out","cpip-pacx.err"};
```

Test matchmaking with:

i2g-job-list-match a3_1.jdl

GROUPS OF CE IDs LIST

The following groups of CE(s) matching your job requirements have been found:

	Groups with 2 CEs	*TotalCPUs*	*ValidCPUs*
[Rank=0 TotalCPUs=34 ValidCPUs=32]			
ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-itut	22	22	
ce.i2g.cesga.es:2119/jobmanager-lcgpbs-itutgrid	12	10	
[Rank=0 TotalCPUs=42 ValidCPUs=35]			
ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-itut	22	22	
ce.i2g.cyf-kr.edu.pl:2119/jobmanager-pbs-itut	20	13	
[Rank=0 TotalCPUs=82 ValidCPUs=37]			
ce-ieg.bifi.unizar.es:2119/jobmanager-lcgpbs-itut	22	22	
i2g-ce01.lip.pt:2119/jobmanager-lcgsgsge-itutgridsdj	60	15	
[Rank=0 TotalCPUs=72 ValidCPUs=25]			
ce.i2g.cesga.es:2119/jobmanager-lcgpbs-itutgrid	12	10	
i2g-ce01.lip.pt:2119/jobmanager-lcgsgsge-itutgridsdj	60	15	

- ❑ Submit the job with:

```
i2g-job-submit -o cpip-pacx.jobid a3_1.jdl
```

- ❑ Check the status until DONE

```
i2g-job-status -i cpip-pacx.jobid
```

- ❑ Get the job output

```
i2g-job-get-output -i cpip-pacx.jobid
```

□ Contact

- ▶ Dr. Isabel Campos
- ▶ iscampos@ifca.unica.es

□ Further information about I2G

- ▶ <http://dissemination.interactive-grid.eu/>
- ▶ https://wiki.fzk.de/i2g/index.php/Main_Page

□ I2G main VOMS server

- ▶ <https://i2g-voms.lip.pt:8443/vomses/>

Data Management

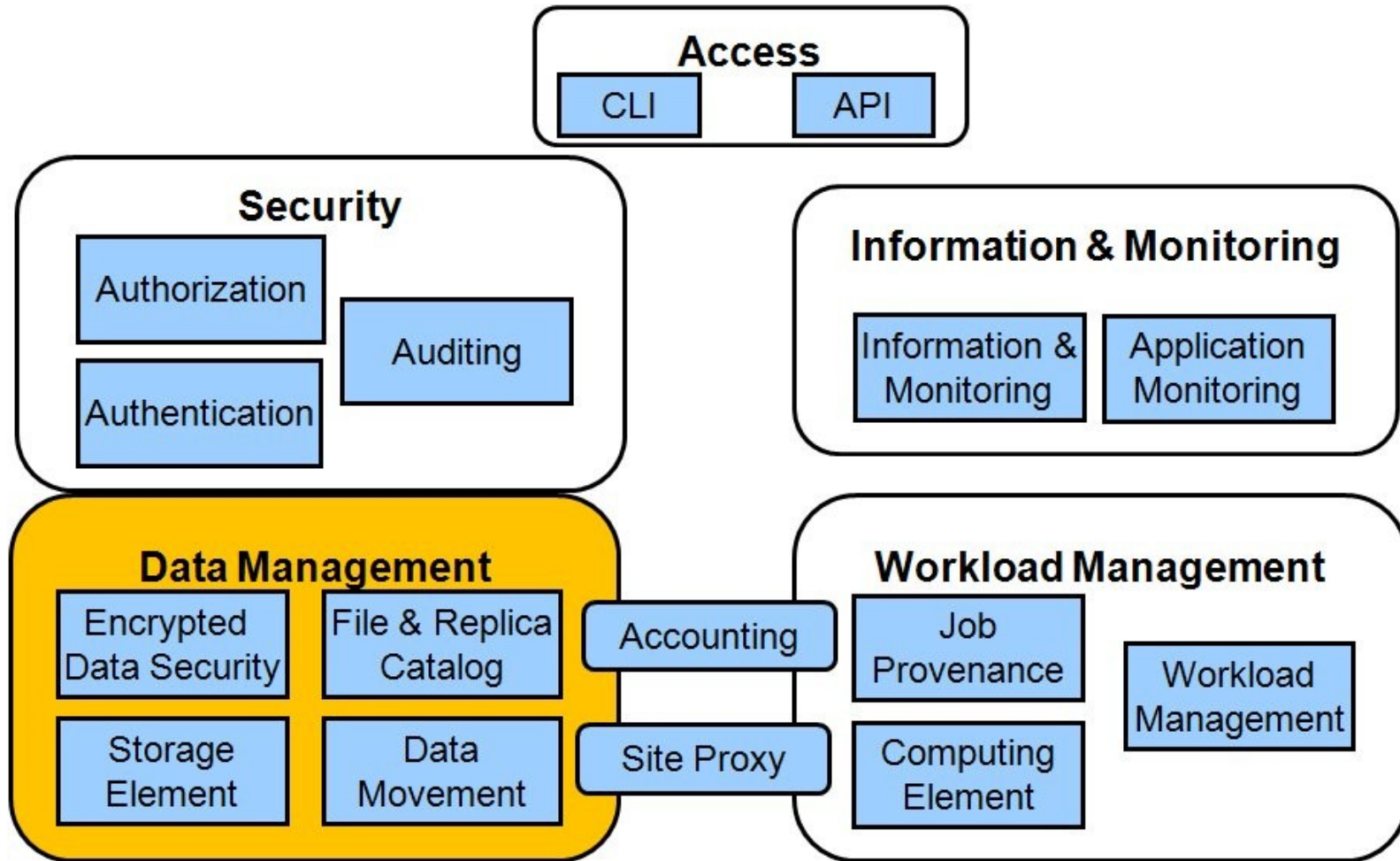
Gonçalo Borges, Jorge Gomes, Mário David
(*goncalo@lip.pt, jorge@lip.pt, david@lip.pt*)

LIP Lisboa

- **Grid Data Management:**
 - Data Management Services
- **Storage Elements and SRM:**
 - Disk Pool Manager (DPM)
 - dCache
 - CASTOR, STORM, BestMan (not covered here)
 - The Storage Resource Manager: *"To bind them all"*
- **File Catalogs and DM tools:**
 - The LCG File Catalog (LFC)
- **File transfer:**
 - Globus (gridftp) tools
 - SRM tools
 - Lcg-utils
 - **File Transfer Service (FTS)** (not covered here).

- **Heterogeneity:**
 - Data are stored on different storage systems using different access technologies.
- **Distribution:**
 - Data are stored in different locations – in most cases there is no shared file system or common namespace.
 - Data need to be moved between different locations.
- **Need common interface to storage resources:**
 - Storage Resource Manager (SRM)
- **Need to keep track where data is stored:**
 - File and Replica Catalogs.
- **Need scheduled, reliable file transfer (FTS).**

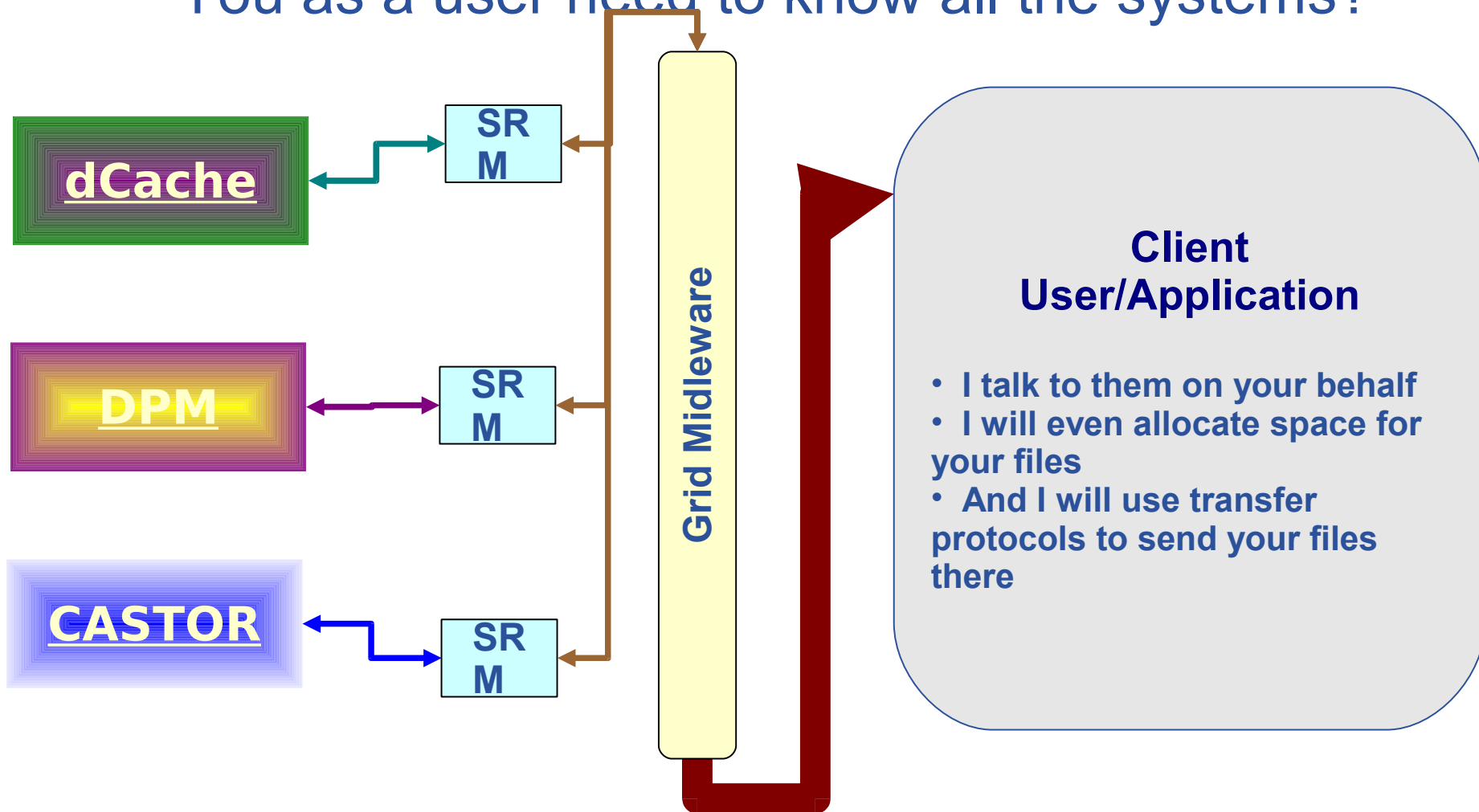
- **Assumptions:**
 - Users and programs produce and require data
 - The lowest granularity of the **data** is on the **file** level (we deal with files rather than data objects or tables): **data = files**
- **Files:**
 - Mostly write once, read many
 - Located in Storage Elements (SE's)
 - Several replicas of one file in different sites
 - Accessible by Grid users and applications from “anywhere”
 - Locatable by the WMS (data requirements in JDL)
- **Also...**
 - WMS can send (small amounts of) data to/from jobs: Input and Output Sandbox
 - Files may be copied from/to local filesystems (WNs, UIs) to the Grid (SE's)



- **4 types of services for DM:**
 - **Storage** (SE's): where files are “physically” located
 - Storage URL or **SURL**:
 - **`srm://dcache01.lip.pt/pnfs/lip.pt/data/itut/test`**
 - **Catalogs**: High level hierarchical namespace, maps the “physical” files to a virtual “logical” filename
 - Logical File Name or **LFN**:
 - **`lfn:/grid/itut/mdavid/itutest`**
 - **Encrypted Storage** (*Not covered in this tutorial*):
 - HYDRA, encrypts files on SE's (used by the biomed community)
 - **Movement**: put/get files into grid SE's, move/replicate files between SE's.
 - File Transfer Service or FTS (*Not covered in this tutorial*)
 - Transport URL or **TURL**:
 - **`gsiftp://se03.lip.pt/pnfs/lip.pt/data/itut/test`**

- **The Storage Element is the service which allow a user or an application to store data for future retrieval**
 - Manage local storage (disks) and interface to Mass Storage Systems(tapes) like: HPSS, CASTOR, ...
 - Most common types of SE's found in the Grid infrastructures:
 - DPM
 - dCache
 - Castor is less common (used at CERN).
 - StoRM and BestMAN.
 - Be able to manage different storage systems uniformly and transparently for the user through a SRM interface
 - Support basic file transfer protocols
 - GSIFTP mandatory
 - Others if available (https, ftp, etc)
 - Support a native I/O (remote file) access protocol
 - POSIX (like) I/O client library for direct access of data (GFAL)

You as a user need to know all the systems?

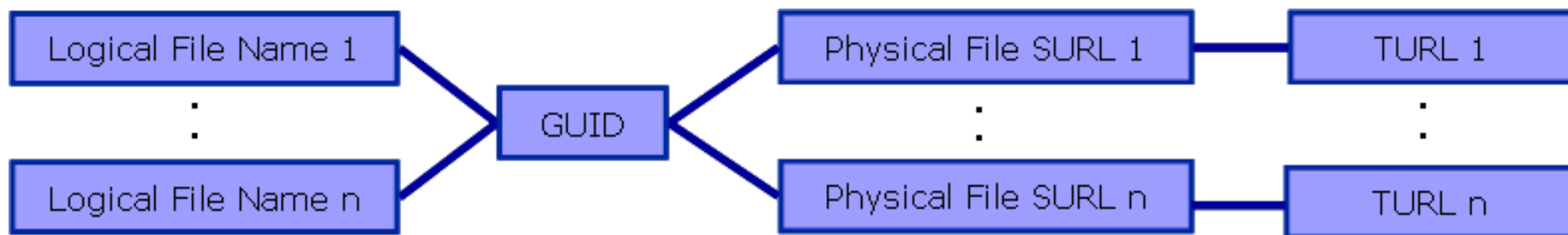


- The Grid Unique Identifier (GUID), which identifies a file uniquely, is of the form:
 - `guid:<36_bytes_unique_string>`
 - `guid:38ed3f60-c402-11d7-a6b0-f53ee5a37e1d`
- The Logical File Name (LFN) or User Alias, which can be used to refer to a file in the place of the GUID (and which should be the normal way for a user to refer to a file), has this format:
 - `lfn:<any_string>`
 - `lfn:/grid/itut/tut-14-11-07/mpi_and_interactivity/BIGFILE.data`
- In the case of the LCG File Catalogue, the LFN's are organized in a hierarchical directory-like structure.

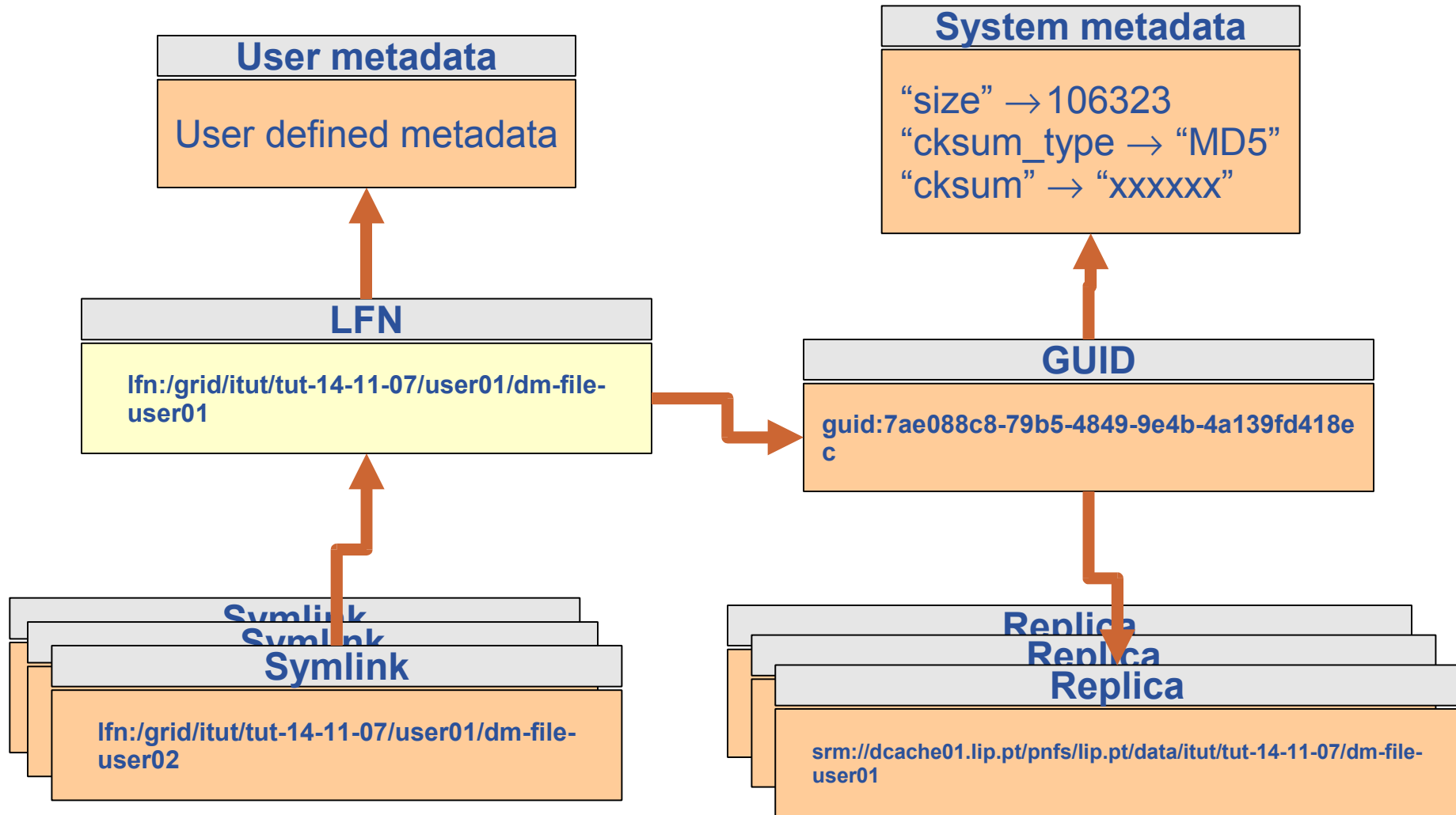
- The Storage URL (SURL), also known as Physical File Name (PFN), which identifies a replica in a SE, is of the general form:
 - `<sfn|srm>://<SE_hostname>/<some_string>`
- The prefix is `sfn` for files located in SE's without a SRM interface and `srm` for SRM-managed SE's.
 - The string after the host name is the path to the location of the file:
 - `sfn://<SE_hostname><SE_Accesspoint><V0_path><filename>`
 - `sfn://i2g-se01.lip.pt/flatfiles/itut/SOME_FILE`
- The SRM-managed SEs can use virtual file systems and the name a file receives may have nothing to do with its physical location:
 - `srm://dcache01.lip.pt/pnfs/lip.pt/data/itut/test`

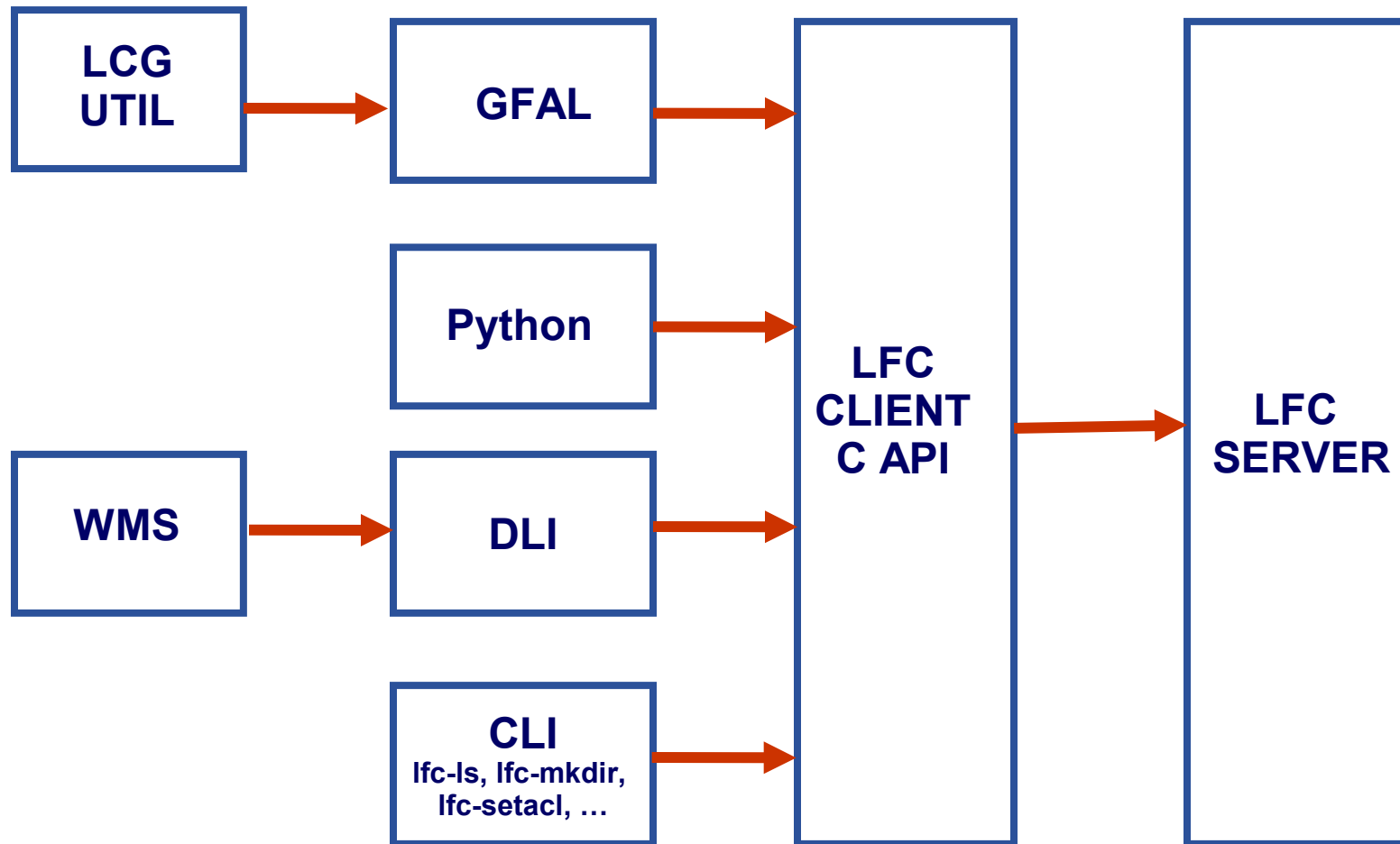
- The Transport URL (TURL), which is a valid URI with the necessary information to access a file in a SE, has the following form:
 - `<protocol>://<some_string>`
 - `gsiftp://se03.lip.pt:2811//pnfs/lip.pt/data/itut/test`
- The `<protocol>` must be a valid protocol (supported by the SE) to access the contents of the file (GSIFTP, RFIO, gsidcap)
- The string after the double slash may have any format that can be understood by the SE serving the file.
- TURL's are obtained dynamically from the SURL through the Information System or the SRM interface.

- Users and applications need to locate files (or replicas) on the Grid. The File Catalogue is the service which maintains mappings between LFN(s), GUID and SURL(s)
- The catalogue publishes its endpoint (service URL) in the Information Service so that it can be discovered by Data Management tools and other services.
- In gLite 3, the adopted file catalog is the **LCG File Catalog** or LFC.
- LFC consists of a unique catalogue, where the LFN is the main key. Further LFN's can be added as symlinks to the main LFN.



- **Timeouts and retries from the client**
- **User exposed transactional API (+ auto rollback on failure)**
- **Hierarchical namespace and namespace operations (for LFN's)**
- **Integrated GSI Authentication + Authorization**
- **Access Control Lists (Unix Permissions and POSIX ACLs)**
- **Checksums**
- **Integration with VOMS (VirtualID and VirtualGID)**



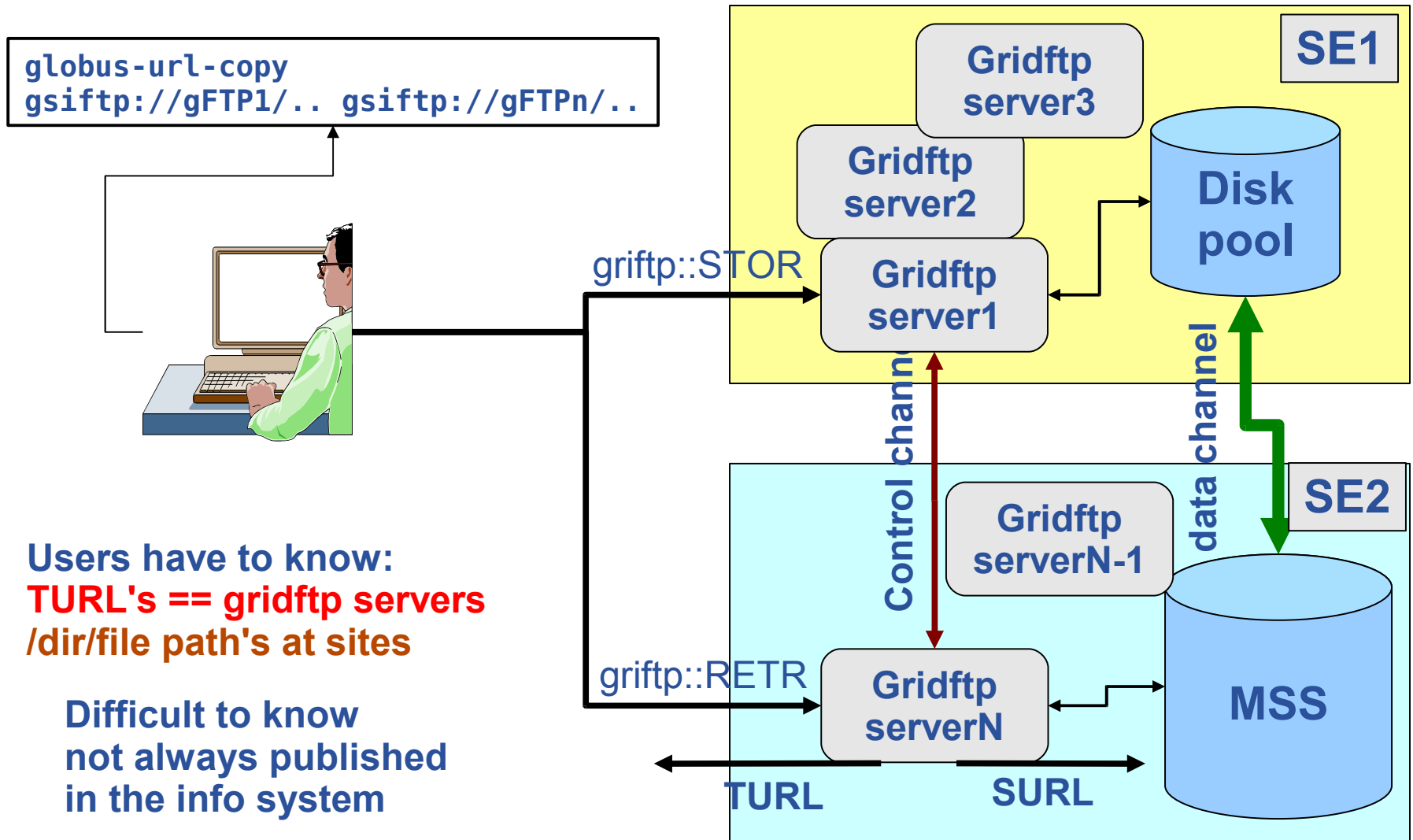


- **LFC client commands (CLI)**
 - Provide administrative functionality
 - Unix-like
 - LFNs seen as a Unix filesystem (/grid/<VO>/ ...)
- **LFC C API**
 - Alternative way to administer the catalog
 - Python wrapper provided
- **Integration with Grid File Access Library (GFAL) and lcg_util API's**
 - lcg-utils access the catalog in a transparent way
- **Integration with the WMS, Data Location Interface (DLI)**
 - The RB/WMS can locate Grid files: allows for data based match-making using the **DLI**.

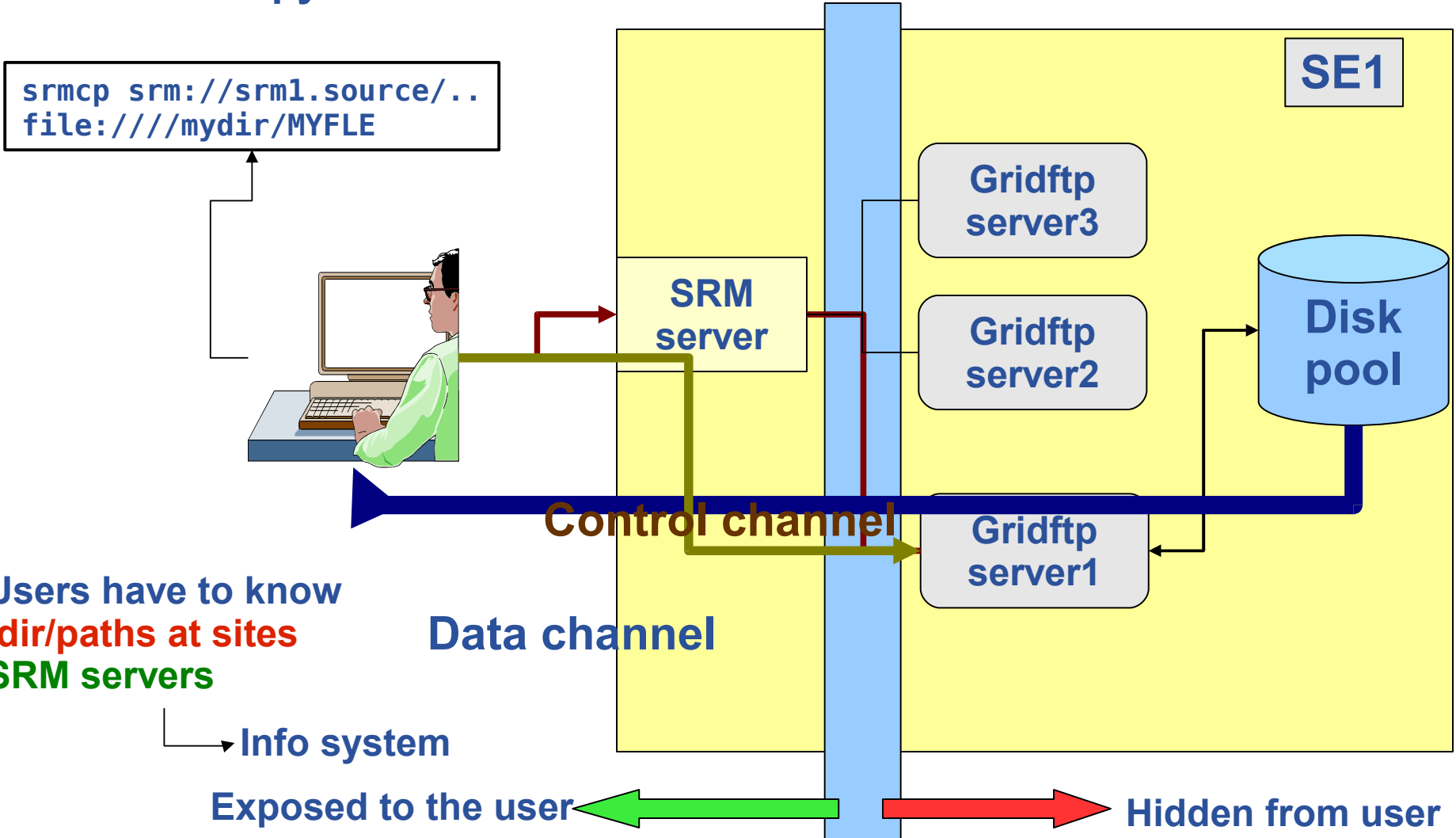
- **Interactions with SE require some components:**
 - File catalog services to locate replicas
 - SRM
 - *File access mechanism to access files from the SE on the WN*
- **What GFAL does:**
 - Hides all these operations
 - *Presents a POSIX like interface for the I/O operations, just prefix `gfal_` to the function name (`open()`, `read()`...)*
 - *Single shared library in threaded and unthreaded versions*
 - **`libgfal.so`, `libgfal_pthr.so`**
 - *Single header file: **`gfal_api.h`***
 - User can create all commands needed for storage management:
 - *It offers as well an interface to SRM*
- **Supported protocols:**
 - *`dcap` and `gsidcap` (dCache access)*
 - *`rfio` and `gsirfio` (dpm and Castor)*

- **A definition first:**
 - *A file is considered to be a **Grid File** if it is both physically present in a SE and registered in the File Catalog*
- **There are three levels of tools to move files around, i.e:**
 - *Put files in SE's from WN's and UI's*
 - *Get files from SE's into WN's and UI's*
 - *Replicate/copy files between SE's*
 - Low level: Globus/gridftp CLI's
 - Medium level: SRM CLI's
 - High level: LCG catalog CLI
- **We recommend that users use the “High level” tools, though any of the other lower level can also be used by users.**

Use case: Copy a file from SE1 to SE2 (also called “**third party transfer**”)



Use case: Copy a file from SE1 to local UI



Use case: Replicate a file to a site SE

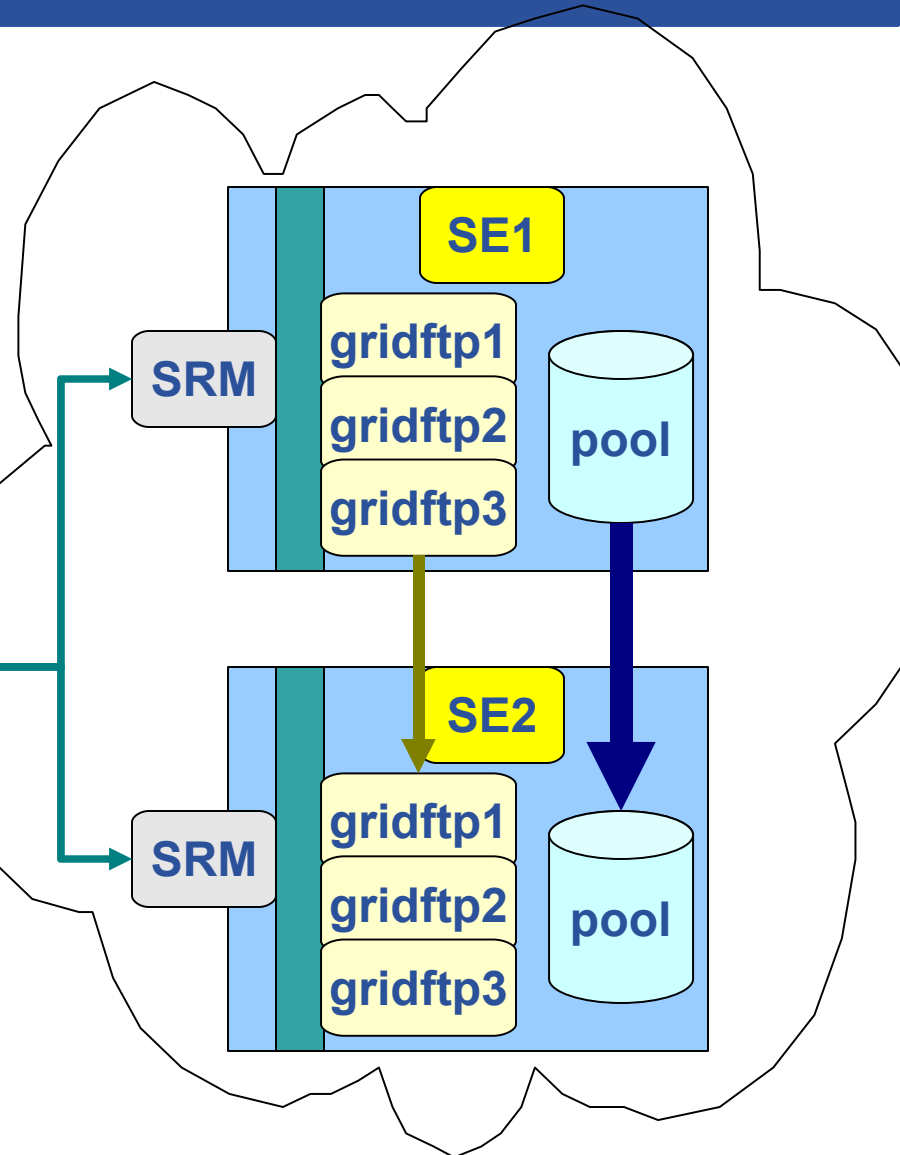
```
lcg-rep -d se.destination  
lfn:/grid/MyV0/myDIR/myFILE
```



LFC server

From the information system you know
the SE's available and their free space

From the LFC server you have your nice
logical file name







Interactive European Grid

Data Management System: Hands On

**Gonçalo Borges, Jorge Gomes, Mário David
(goncalo@lip.pt, jorge@lip.pt, david@lip.pt)**

LIP Lisboa

gLite and INT.EU.GRID training for end-users, FCCN 14 November 2007

- ❑ Data movement:
 - ▶ Low level DM tools: **GSIFTP**
 - ▶ SRM Interface: **srm client tools**
- ❑ File Catalogs:
 - ▶ LFC command line interface: **LFC-client**
- ❑ File and Replica Management tools
 - ▶ LCG data management tools: **lcg_utils**

- ❑ In the exercises, the following file will be used:
 - ▶ `/home/tutorial/userNN/data-manag/dm-file-userNN`
- ❑ The storage directory in the dCache SE system:
 - ▶ `/pnfs/lip.pt/data/itut`
- ❑ The LFC catalog directory:
 - ▶ `/grid/itut/tut-15-05-08`
- ❑ Please export first:
 - ▶ `[ui03] > export LFC_HOST=lfc01.lip.pt`
- ❑ GSIftp servers: `se03.lip.pt` and `i2g-se01.lip.pt`
- ❑ SRM server: `dcache01.lip.pt`

- ❑ The following low level tools can be used to interact with GSIFTP servers running on SE's.

edg-gridftp-exists	TURL	Checks the existence of a file or directory on a SE
Edg-gridftp-ls	TURL	Lists a directory on a SE
edg-gridftp-mkdir	TURL	Creates a directory on a SE
edg-gridftp-rename	sourceTURL destTURL	Renames a file on a SE
edg-gridftp-rm	TURL	Removes a file on a SE
edg-gridftp-rmdir	TURL	Removes a directory on a SE
globus-url-copy	sourceTURL destTURL	Copies files to/from/between SE's

The commands **edg-gridftp-rename**, **edg-gridftp-rm**, and **edg-gridftp-rmdir** should be used with extreme care.

All the **edg-gridftp-*** commands accept **gsiftp** as the only valid protocol for the TURL.

To obtain help on these commands use the option **--usage** or **--help**.

Exercise 1: Globus gridftp

Listing, copying and removing files.

```
[ui03] > edg-gridftp-ls --verbose gsiftp://i2g-se01.lip.pt/flatfiles/itut
total 4
drwxrwxr-x      3 itut          4096 Nov  8 15:06 generated
drwxrwxr-x      2 itut          4096 Nov  8 18:32 tut-14-11-07
```

```
[ui03] > globus-url-copy -vb file:/home/tutorial/user01/data-manag/dm-file-user01
gsiftp://i2g-se01.lip.pt/flatfiles/itut/tut-14-11-07/dm-file-user01
1048576 bytes          329.49 KB/sec avg          329.49 KB/sec inst
```

```
[ui03] > edg-gridftp-ls --verbose gsiftp://i2g-
se01.lip.pt/flatfiles/itut/tut-14-11-07
total 9412
-rw-rw-r--      1 itut        9621413 Nov  8 18:33 dm-file-user01
```

```
[ui03] > edg-gridftp-rm gsiftp://i2g-
se01.lip.pt/flatfiles/itut/tut-14-11-07/dm-file-user01
```

- ❑ The SRM version (1.1) has only three client tools:
 - ▶ **srncmp**: copies files to/from/between SE's.
 - ▶ **srm-advisory-delete**: deletes a file from a SE.
 - ▶ **srm-get-metadata**: prints metadata information for a file (e.g. size, checksum, etc.)
 - ▶ **-h/-help/--help**: gives usage

Exercise 2: SRM

Copying files.

```
[ui03] > srmcp -debug=false file:///home/tutorial/user01/data-manag/dm-file-user01 srm://dcache01.lip.pt:8443/pnfs/lip.pt/data/itut/tut-14-11-07/dm-file-user01
```

Try with **-debug=true** and don't forget to change the destination file name.

```
[ui03] > edg-gridftp-ls --verbose  
gsiftp://se03.lip.pt/pnfs/lip.pt/data/itut/tut-14-11-07/dm-file-user01  
-rw          9621413      dm-file-user01
```

Exercise 3: SRM

Get metadata for files.

```
[ui03] > srm-get-metadata
srm://dcache01.lip.pt:8443/pnfs/lip.pt/data/itut/tut-14-11-07/dm-file-user01

FileMetaData(srm://dcache01.lip.pt:8443/pnfs/lip.pt/data/itut/tut-14-11-07/dm-file-user01)=
RequestFileStatus      SURL
:srm://dcache01.lip.pt:8443/pnfs/lip.pt/data/itut/tut-14-11-07/dm-file-user01
    size :9621413
    owner :12001
    group :12000
    permMode :420
    checksumType :adler32
    checksumValue :e7b38326
    isPinned :false
    isPermanent :true
    isCached :true
    state :
    fileId :0
    TURL :
    estSecondsToStart :0
    sourceFilename :
    destFilename :
    queueOrder :0
```

Exercise 4: SRM

Removing files.

```
[ui03] > srm-advisory-delete srm://dcache01.lip.pt:8443/pnfs/lip.pt/data/itut/  
tut-14-11-07/dm-file-user01
```

Check again with:

```
edg-gridftp-ls --verbose gsiftp://se03.lip.pt/pnfs/lip.pt/data/itut/  
tut-14-11-07/dm-file-user01
```

LFC client tools: Introduction

- Check that:

```
[ui03] > echo $LFC_HOST  
lfc01.lip.pt
```

- The directory structure of the LFC namespace has the form:

- `/grid/<V0>/<subpath>`

- Users should use the next commands carefully, keeping in mind that the operations they are performing affect the catalogue, but not the physical files that the entries represent.

LFC client tools: commands

<code>lfc-chmod</code>	Change access mode of a LFC file/directory
<code>lfc-chown</code>	Change owner and group of a LFC file/directory
<code>lfc-delcomment</code>	Delete the comment associated with a file/directory
<code>lfc-getacl</code>	Get file/directory access control lists
<code>lfc-ln</code>	Make a symbolic link to a file/directory
<code>lfc-ls</code>	List file/directory entries in a directory
<code>lfc-mkdir</code>	Create directory
<code>lfc-rename</code>	Rename a file/directory
<code>lfc-rm</code>	Remove a file/directory
<code>lfc-setacl</code>	Set file/directory access control lists
<code>lfc-setcomment</code>	Add/replace a comment
<code>lfc-entergrpmap</code>	Defines a new group entry in the Virtual ID table
<code>lfc-enterusrmap</code>	Defines a new user entry in Virtual ID table
<code>lfc-modifygrpmap</code>	Modifies a group entry corresponding to a given virtual gid
<code>lfc-modifyusrmap</code>	Modifies a user entry corresponding to a given virtual uid
<code>lfc-rmgrpmap</code>	Suppresses group entry corresponding to a given virtual gid or group name
<code>lfc-rmusrmap</code>	Suppresses user entry corresponding to a given virtual uid or user name.

Man pages are available for all the commands. Most of them work in a very similar way to their Unix equivalents, but operating on directories and files of the catalog namespace.

Creating directories, checking ACL's, ls.

```
[ui03] > lfc-mkdir /grid/itut/tut-15-05-08/user01
```

```
[ui03] > lfc-getacl /grid/itut/tut-15-05-08
```

```
# file: /grid/itut/tut-15-05-08
# owner: /C=PT/O=LIP/CN=user01
# group: itut
user::rwx
group::rwx                #effective:rwx
other::r-x
default:user::rwx
default:group::rwx
default:other::r-x
```

```
[ui03] > lfc-ls -l /grid/itut/tut-15-05-08/
```

```
drwxrwxr-x   1 227      108          0 Nov 08 17:48 mpi_and_interactivity
drwxrwxr-x   1 224      108          0 Nov 08 15:06 user01
```

Exercise 6: LFC

Adding metadata, changing permissions and removing directories.

```
[ui03] > lfc-setcomment /grid/itut/tut-15-05-08/user01 "My first comment"
```

```
[ui03] > lfc-ls -l --comment /grid/itut/tut-15-05-08/
drwxrwxr-x   1 227      108    0 Nov 08 17:48 mpi_and_interactivity
drwxrwxr-x   0 224      108    0 Nov 09 08:35 test
drwxrwxr-x   1 224      108    0 Nov 08 15:06 user01 My first comment
```

```
[ui03] > lfc-chmod 700 /grid/itut/tut-15-05-08/user01
[ui03] > lfc-ls -l /grid/itut/tut-15-05-08/
drwxrwxr-x   1 227      108    0 Nov 08 17:48 mpi_and_interactivity
drwx-----   1 224      108    0 Nov 08 15:06 user01
```

```
[ui03] > lfc-rm -r /grid/itut/tut-15-05-08/user01
[ui03] > lfc-ls -l /grid/itut/tut-15-05-08
drwxrwxr-x   1 227      108    0 Nov 08 17:48 mpi_and_interactivity
drwxrwxr-x   0 224      108    0 Nov 09 08:35 test
```

A file can be removed only if there are no SURL's associated to it.
A directory can be removed (-r option) only if it is empty.

LCG DM client Tools: lcg_utils

The LCG Data Management tools (lcg_util) allow users to copy files between UI, CE, WN and a SE, to register entries in the file catalogue and replicate files between SEs.

Replica Management

lcg-cp	Copies a Grid file to a local destination (download)
lcg-cr	Copies a file to a SE and registers the file in the catalogue (upload)
lcg-del	Deletes one file (either one replica or all replicas)
lcg-rep	Copies a file from one SE to another SE and registers it in the catalogue (replicate)
lcg-gt	Gets the TURL for a given SURL and transfer protocol
lcg-sd	Sets file status to "Done" for a given SURL in an SRM's request

File Catalog interactions

lcg-aa	Adds an alias in the catalogue for a given GUID
lcg-ra	Removes an alias in the catalogue for a given GUID
lcg-rf	Registers in the catalogue a file residing on an SE
lcg-uf	Unregisters in the the catalogue a file residing on an SE
lcg-la	Lists the aliases for a given LFN, GUID or SURL
lcg-lg	Gets the GUID for a given LFN or SURL
lcg-lr	Lists the replicas for a given LFN, GUID or SURL

Check: **> echo \$LCG_GFAL_INFOSYS
i2g-ii01.lip.pt:2170**

And: **> export LCG_GFAL_V0=itut**

Exercise 7: lcg_utils

Copy and register a file in the LFC:

```
[ui03] > lcg-cr -v -l lfn:/grid/itut/tut-15-05-08/user01/dm-file-user01
file:/home/tutorial/user01/data-manag/dm-file-user01
Using grid catalog type: lfc
Using grid catalog : lfc01.lip.pt
Using LFN : /grid/itut/tut-15-05-08/user01/dm-file-user01
Using SURL : sfn://i2g-
se01.lip.pt/flatfiles/itut/generated/2007-11-09/file39c282cd-3071-4d45-a0f3-
a31369a80fc8
Source URL: file:/home/tutorial/user01/data-manag/dm-file-user01
File size: 9621413
V0 name: itut
Destination specified: i2g-se01.lip.pt
Destination URL for copy: gsiftp://i2g-
se01.lip.pt/flatfiles/itut/generated/2007-11-09/file39c282cd-3071-4d45-a0f3-
a31369a80fc8
# streams: 1
# set timeout to 0 seconds
Alias registered in Catalog: lfn:/grid/itut/tut-14-11-07/user01/dm-file-user01
      1048576 bytes      401.02 KB/sec avg      401.02 KB/sec inst
Transfer took 12510 ms
Destination URL registered in Catalog: sfn://i2g-
se01.lip.pt/flatfiles/itut/generated/2007-11-09/file39c282cd-3071-4d45-a0f3-
a31369a80fc8
guid:1376d0cc-d09a-48f7-9e24-131f4999ff7e
```

Exercise 8: lcg_utils

List alias, GUID, replicas

```
[ui03] > lcg-la sfn://i2g-  
se01.lip.pt/flatfiles/itut/generated/2007-11-09/file39c282cd-3071-4d45-a0f3-  
a31369a80fc8  
lfn:/grid/itut/tut-14-11-07/user01/dm-file-user01
```

```
[ui03] > lcg-lg lfn:/grid/itut/tut-15-05-08/user01/dm-file-user01  
guid:1376d0cc-d09a-48f7-9e24-131f4999ff7e
```

```
[ui03] > lcg-lr lfn:/grid/itut/tut-15-05-08/user01/dm-file-user01  
sfn://i2g-  
se01.lip.pt/flatfiles/itut/generated/2007-11-09/file39c282cd-3071-4d45-a0f3-  
a31369a80fc8
```

```
[ui03] > lfc-ls -l /grid/itut/tut-15-05-08/user01/dm-file-user01  
-rw-rw-r-- 1 224 108 9621413 Nov 09 10:14 dm-file-user01
```

Exercise 9: lcg_utils

Replicate a file

```
[ui03] > lcg-rep -v -d dcache01.lip.pt lfn:/grid/itut/tut-15-05-08/user01/dm-
file-user01
Using grid catalog type: lfc
Using grid catalog : lfc01.lip.pt
Source URL: lfn:/grid/itut/tut-14-11-07/user01/dm-file-user01
File size: 9621413
VO name: itut
Destination specified: dcache01.lip.pt
Source URL for copy: gsiftp://i2g-
se01.lip.pt/flatfiles/itut/generated/2007-11-09/file39c282cd-3071-4d45-a0f3-
a31369a80fc8
Destination URL for copy:
gsiftp://se03.lip.pt:2811//pnfs/lip.pt/data/itut/generated/2007-11-09/file3c6474f
c-361f-4f70-8112-ea974c5e2138
# streams: 1
# set timeout to 0
0 bytes      0.00 KB/sec avg      0.00 KB/sec inst
Transfer took 5050 ms
Destination URL registered in LRC:
srm://dcache01.lip.pt/pnfs/lip.pt/data/itut/generated/2007-11-09/file3c6474fc-361
f-4f70-8112-ea974c5e2138
```

Exercise 10: lcg_utils

List the replicas, and removing a file

```
[ui03] > lcg-lr lfn:/grid/itut/tut-15-05-08/user01/dm-file-user01
sfn://i2g-se01.lip.pt/flatfiles/itut/generated/2007-11-09/file39c282cd-3071-4d45-
a0f3-a31369a80fc8
srm://dcache01.lip.pt/pnfs/lip.pt/data/itut/generated/2007-11-09/file3c6474fc-361
f-4f70-8112-ea974c5e2138
```

```
[ui03] > lcg-del -v
srm://dcache01.lip.pt/pnfs/lip.pt/data/itut/generated/2007-11-09/file3c6474fc-361
f-4f70-8112-ea974c5e2138
V0 name: (null)
Using GUID : 1376d0cc-d09a-48f7-9e24-131f4999ff7e
```

```
[ui03] > lcg-del -v -a lfn:/grid/itut/tut-15-05-08/user01/dm-file-user01
V0 name: (null)
Using GUID : 1376d0cc-d09a-48f7-9e24-131f4999ff7e
set timeout to 0 seconds
sfn://i2g-se01.lip.pt/flatfiles/itut/generated/2007-11-09/file39c282cd-3071-4d45-
a0f3-a31369a80fc8 is deleted
sfn://i2g-se01.lip.pt/flatfiles/itut/generated/2007-11-09/file39c282cd-3071-4d45-
a0f3-a31369a80fc8 is unregistered
```

Exercise 11: lcg_utils

Copying a file and register that file

```
[ui03] > lcg-cp -v file:/home/tutorial/user01/data-manag/dm-file-user01
srm://dcache01.lip.pt/pnfs/lip.pt/data/itut/tut-15-05-08/dm-file-user01
Source URL: file:/home/tutorial/user01/data-manag/dm-file-user01
File size: 9621413
Source URL for copy: file:/home/tutorial/user01/data-manag/dm-file-user01
Destination URL: srm://dcache01.lip.pt/pnfs/lip.pt/data/itut/tut-14-11-07/dm-
file-user01
# streams: 1
# set timeout to 0 (seconds)
      1048576 bytes   1423.13 KB/sec avg   1423.13 KB/sec inst
Transfer took 2040 ms
```

```
[ui03] > lcg-rf -v -l lfn:/grid/itut/tut-15-05-08/user01/dm-file-user01
srm://dcache01.lip.pt/pnfs/lip.pt/data/itut/tut-15-05-08/dm-file-user01
Using grid catalog type: lfc
Using grid catalog : lfc01.lip.pt
set timeout to 0 seconds
Site URL to be registered:
srm://dcache01.lip.pt/pnfs/lip.pt/data/itut/tut-15-05-08/dm-file-user01
File size: 9621413
lfc01.lip.pt: /grid/itut/tut-15-05-08/user01/dm-file-user01: No such file or
directory
Alias created in Catalog: lfn:/grid/itut/tut-15-05-08/user01/dm-file-user01
guid:7ae088c8-79b5-4849-9e4b-4a139fd418ec
```

Exercise 12: lcg_utils

Get the TURL, un-registering a file

```
[ui03] > lcg-gt srm://dcache01.lip.pt/pnfs/lip.pt/data/itut/tut-15-05-08/dm-file-user01 gsiftp  
gsiftp://stcalo04.lip.pt:2811//pnfs/lip.pt/data/itut/tut-15-05-08/dm-file-user01
```

```
[ui03] > lcg-uf guid:7ae088c8-79b5-4849-9e4b-4a139fd418ec  
srm://dcache01.lip.pt/pnfs/lip.pt/data/itut/tut-15-05-08/dm-file-user01
```

Confirm with: `[ui03] > lfc-ls -l /grid/itut/tut-15-05-08/user01/`

