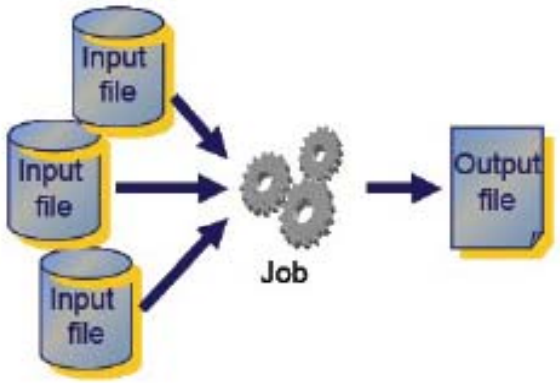
Laboratory: Hands using EGEE Grid (LCG, gLite)

Athanasia Asiki
aassiki@cslab.ece.ntua.gr
Computing Systems Laboratory,
National Technical University of Athens

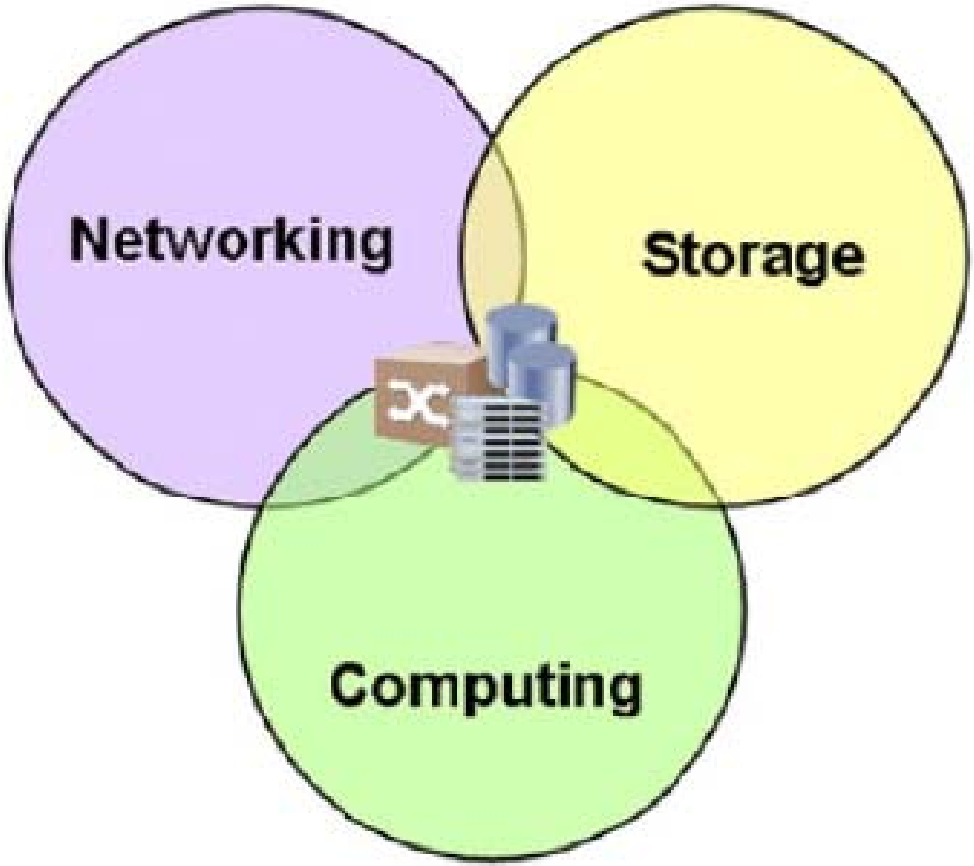
Let's enter the Grid…

Application's structure



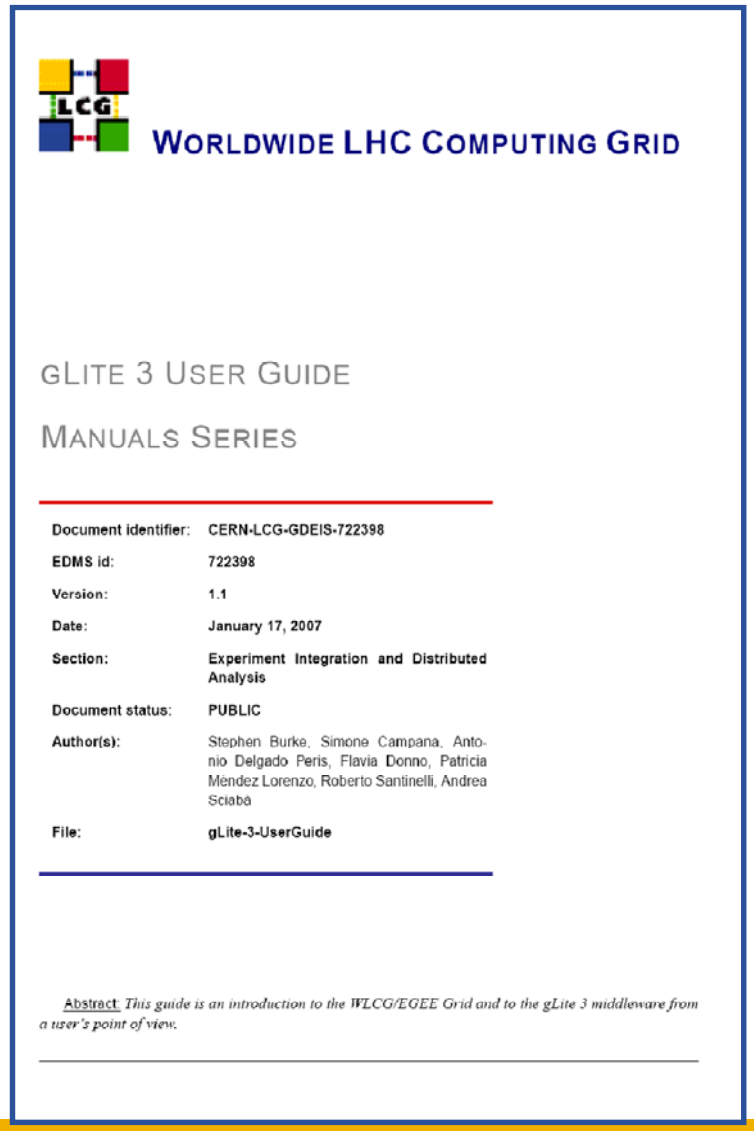
- **The execution of a typical Grid application follows this scenario:**
  - The user submits its application's job to the "Grid"
  - The job is being executed
  - The job's execution may include the processing of one or more **Input Files** stored in a Storage node
  - The job may produce one or more **Output Files**
  - The **Output Files** can be stored somewhere in the Grid system (perhaps in the Storage Element or in the User Interface)
  - The User can access the **Output Files** using the corresponding Grid mechanisms

Grid Resources



Documentation

http://glite.web.cern.ch/glite/documentation/



WORLDWIDE LHC COMPUTING GRID

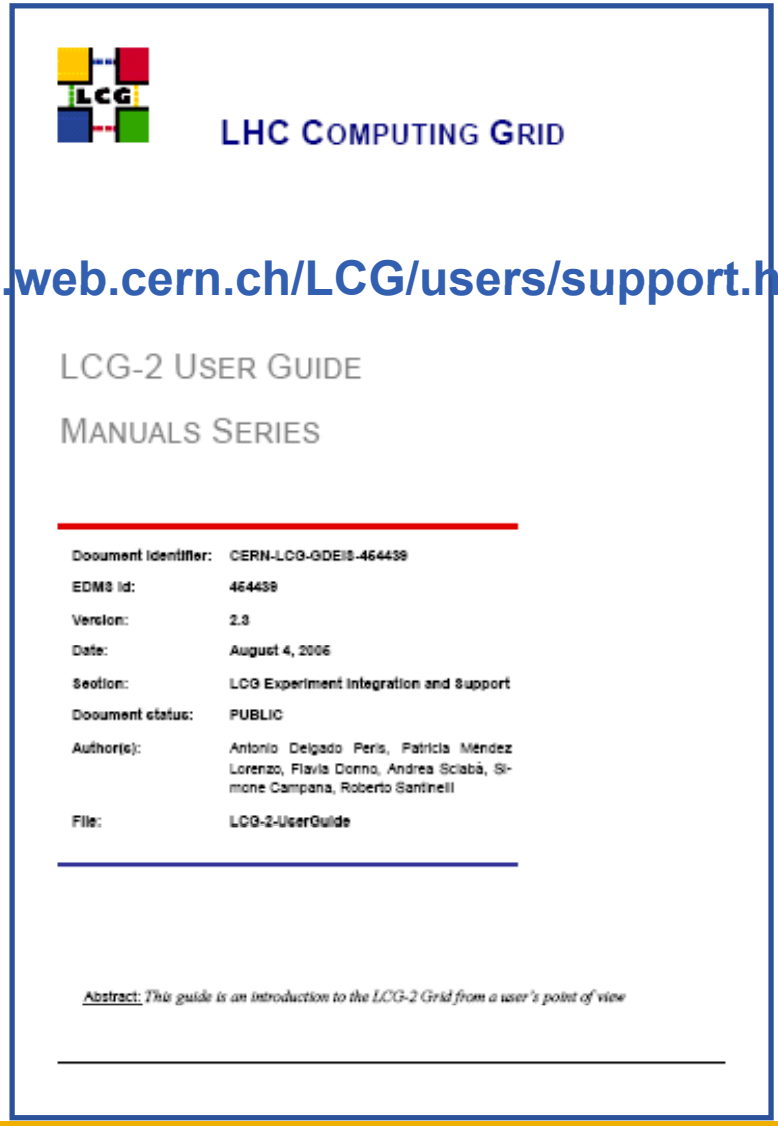
gLITE 3 USER GUIDE

MANUALS SERIES

Document identifier:	CERN-LCG-GDEIS-722398
EDMS id:	722398
Version:	1.1
Date:	January 17, 2007
Section:	Experiment Integration and Distributed Analysis
Document status:	PUBLIC
Author(s):	Stephen Burke, Simone Campana, Antonio Delgado Peris, Flavia Donno, Patricia Méndez Lorenzo, Roberto Santinelli, Andrea Sciabà
File:	gLite-3-UserGuide

Abstract: *This guide is an introduction to the WLCG/EGEE Grid and to the gLite 3 middleware from a user's point of view.*

http://lcg.web.cern.ch/LCG/users/support.html



LHC COMPUTING GRID

LCG-2 USER GUIDE

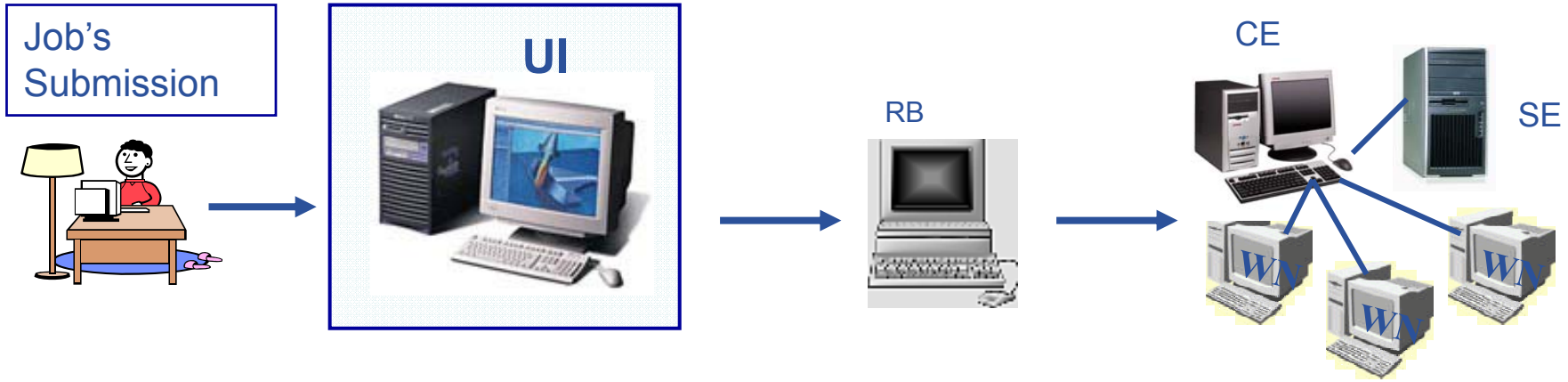
MANUALS SERIES

Document identifier:	CERN-LCG-GDEIS-454439
EDMS id:	454439
Version:	2.3
Date:	August 4, 2006
Section:	LCG Experiment Integration and Support
Document status:	PUBLIC
Author(s):	Antonio Delgado Peris, Patricia Méndez Lorenzo, Flavia Donno, Andrea Sciabá, Simone Campana, Roberto Santinelli
File:	LCG-2-UserGuide

Abstract: *This guide is an introduction to the LCG-2 Grid from a user's point of view*

User Interface (1)

- Allows users to access Grid functionalities
- A machine where users have a personal account and where the user certificate is installed
- Gateway to Grid Services



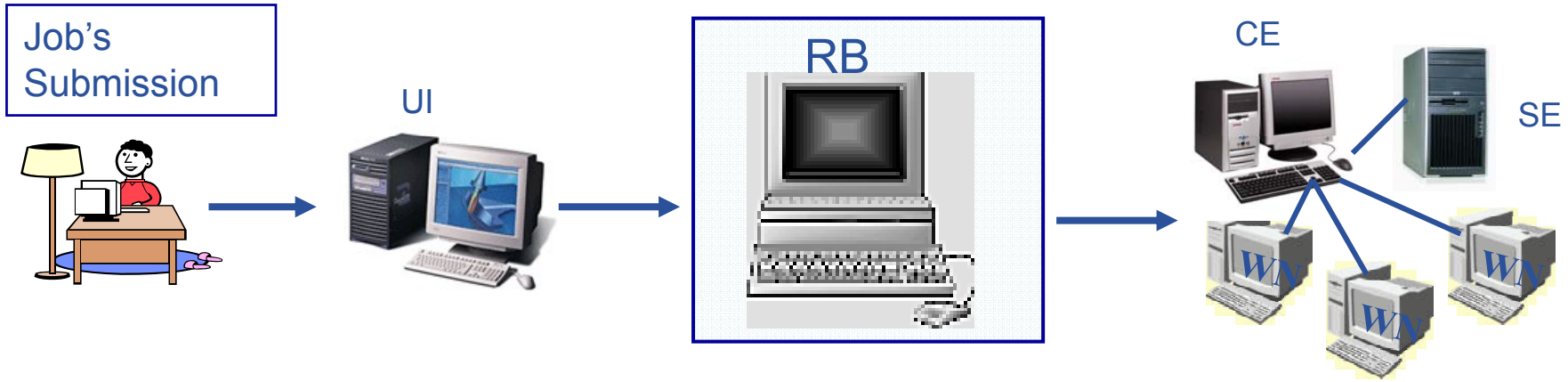
User Interface (2)

- *It provides a Command Line Interface to perform some basic Grid operations such as:*

- List all the resources suitable to execute a given job
- Submit jobs for execution
- Show the status of submitted jobs
- Cancel one or more jobs
- Retrieve the logging and bookkeeping information of jobs
- Retrieve the output of finished jobs
- Copy, replicate and delete files from Grid

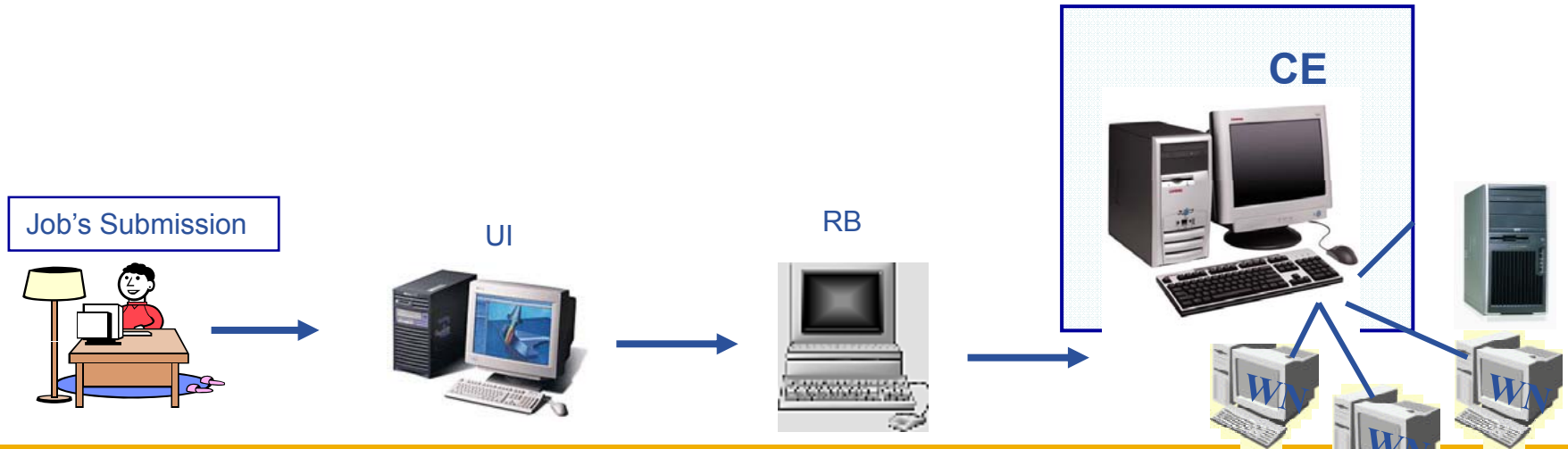
Workload Management System

- The resource broker is responsible for the acceptance of submitted jobs and for sending those jobs to the appropriate Computing Element
- Retrieves information from Information Catalogues so as to find the proper available resources depending on the job requirements



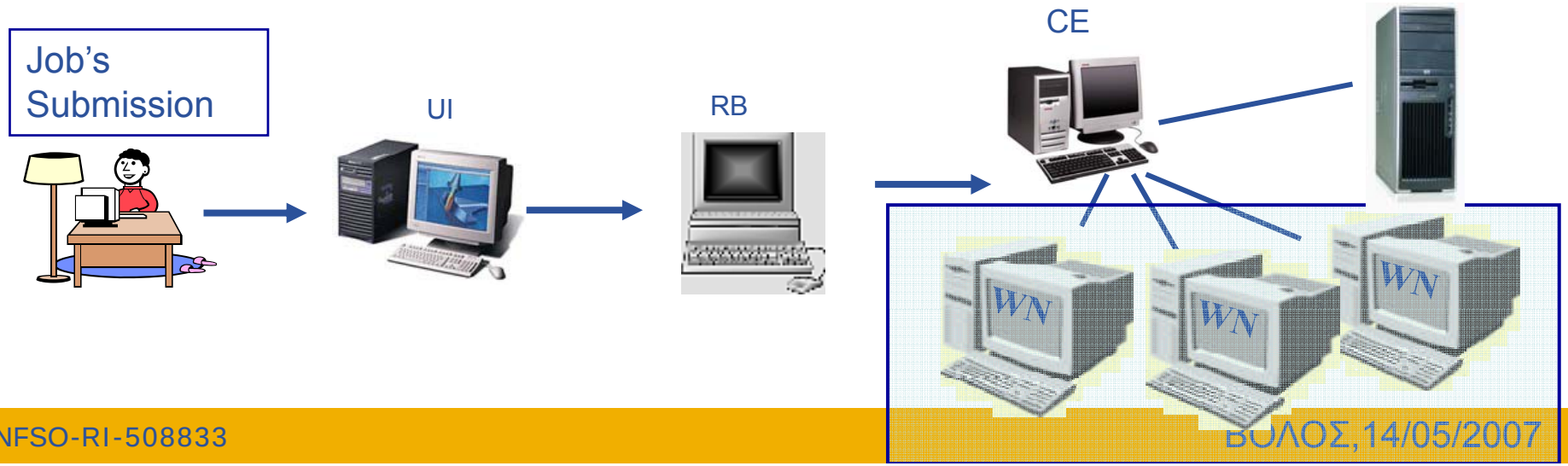
Computing Element

- **"Grid interface"**
- **It is built on a farm of a computing nodes called Worker Nodes (WNs)**
- **Executes the basic queues functions**
- **In the Computing Element, a process is being executed that accepts jobs and dispatch them for execution to the Worker nodes (WNs)**
- **The state of an executing job is being watched by the Computing Element**



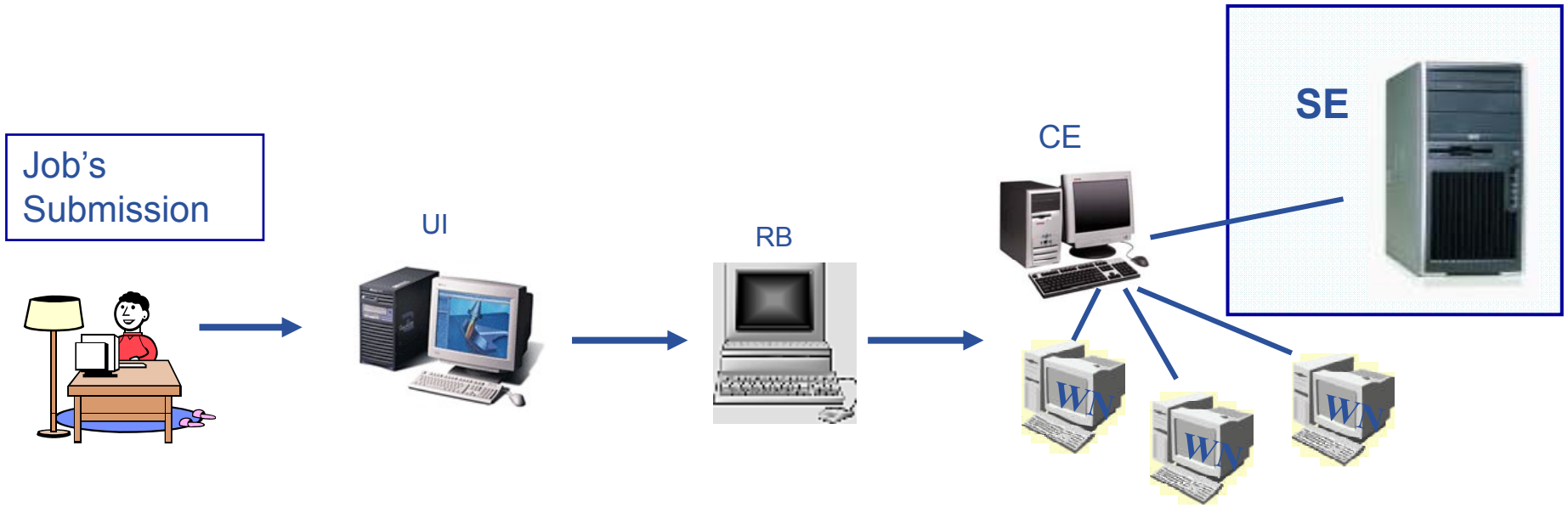
Worker Node

- The submitted jobs are being executed in the Worker nodes
- Need only outbound connectivity
- Only basic services of middleware are required to be provided by the Worker nodes such as
  - Application libraries
  - Application Programming Interfaces (API)
  - Commands for performing actions on Grid resources and Grid data



Storage Element

- **It provides uniform access to storage resources**
  (it may control simple disk servers, large disk arrays or Mass Storage Systems (MSS)
- **Each site may provide one or more SEs**



Information System (IS)

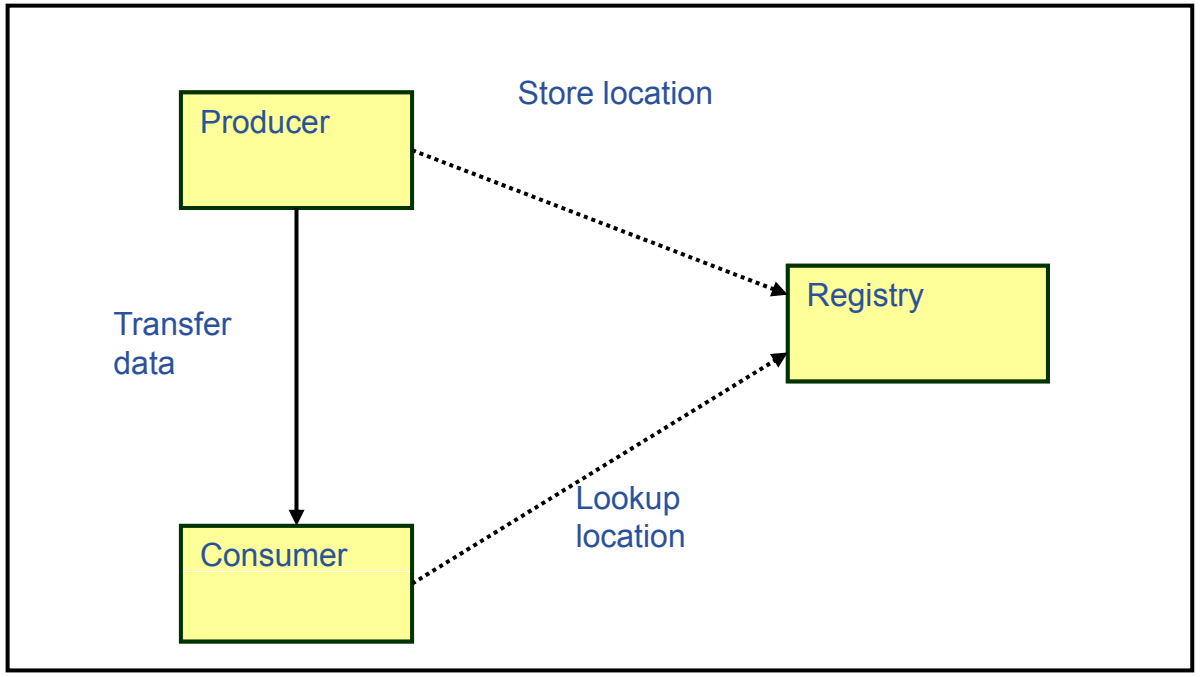
- **It provides information about the Grid resources and their status**
  ⇨ This information is essential for the operation of the whole Grid
- **Location of availiable Computing Elements to run jobs**
- **Finding of SEs that holding replicas of Grid files and the catalogs keeping the information on these files**
- **The information is stored in databases**
- **The published information is used for**
  - ✓ monitoring purposes ⇨ for analyzing usage and performance of the Grid, detecting fault situations and any other interesting events
  - ✓ accounting purposes ⇨ for creating statistics of the applications run by the users in the resources

MDS

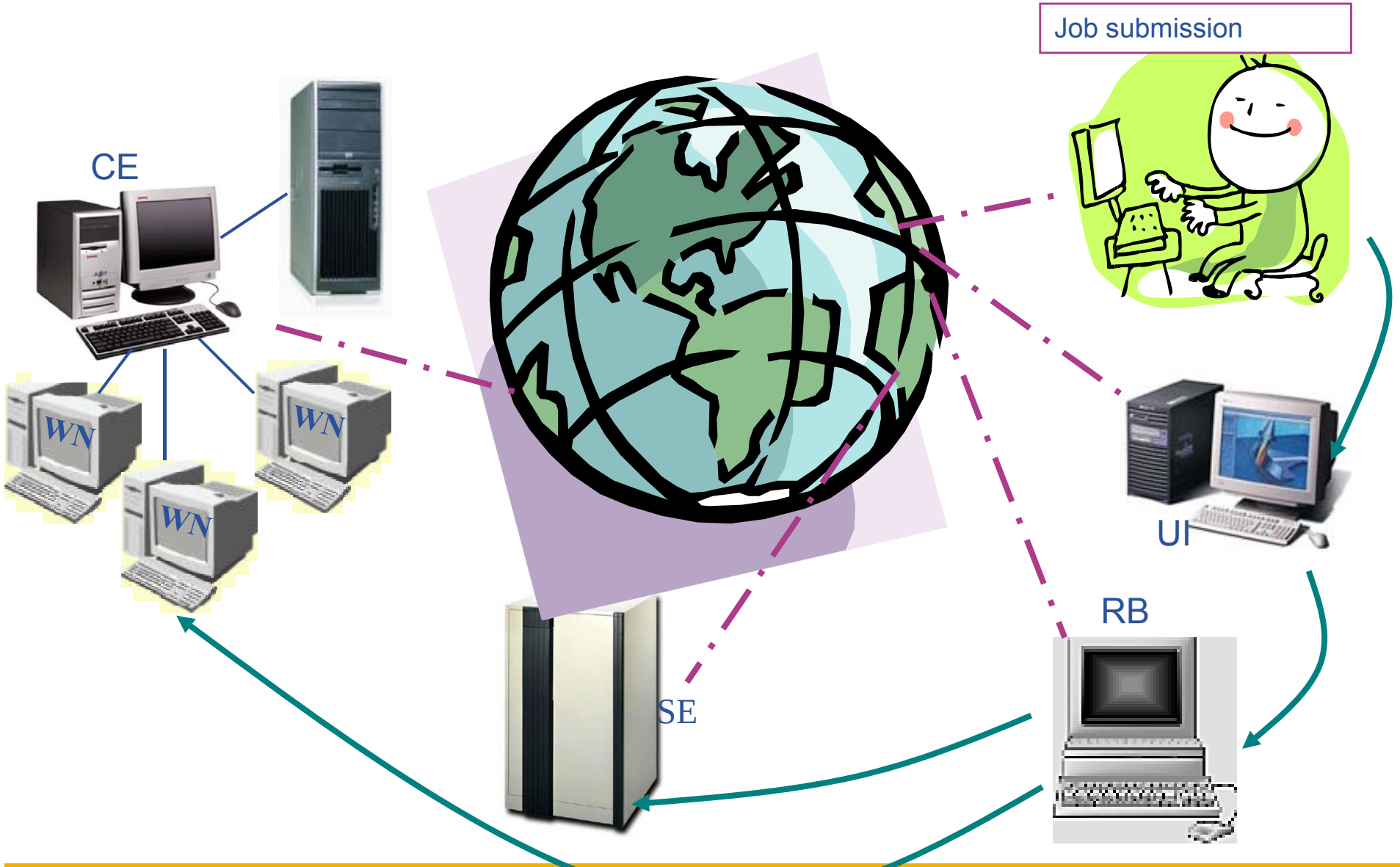
- **Globus Moinitoring and Discovery service**
- **Resource Discovery and publishing the resource status**
- **OpenLDAP** which is an open source implementation of the ***Lightweight Directory Access Protocol (LDAP)***, a specialised database optimised for reading, browsing and searching information
- **Hierarchical architecture:**
  - In every resource runs a ***Grid Resource Information Server (GRIS)*** providing relevant information about the resource
  - At each site runs a ***Site Grid Information Server (GIIS)*** that collects information from the local GRISes and republishes it. The GIIS uses a ***Berkeley Database Information Index (BDII)*** to store data
  - A BDII is used to read from a group of sites, depicting a view of the overall Grid resources (on top of the hierarchy)

R-GMA

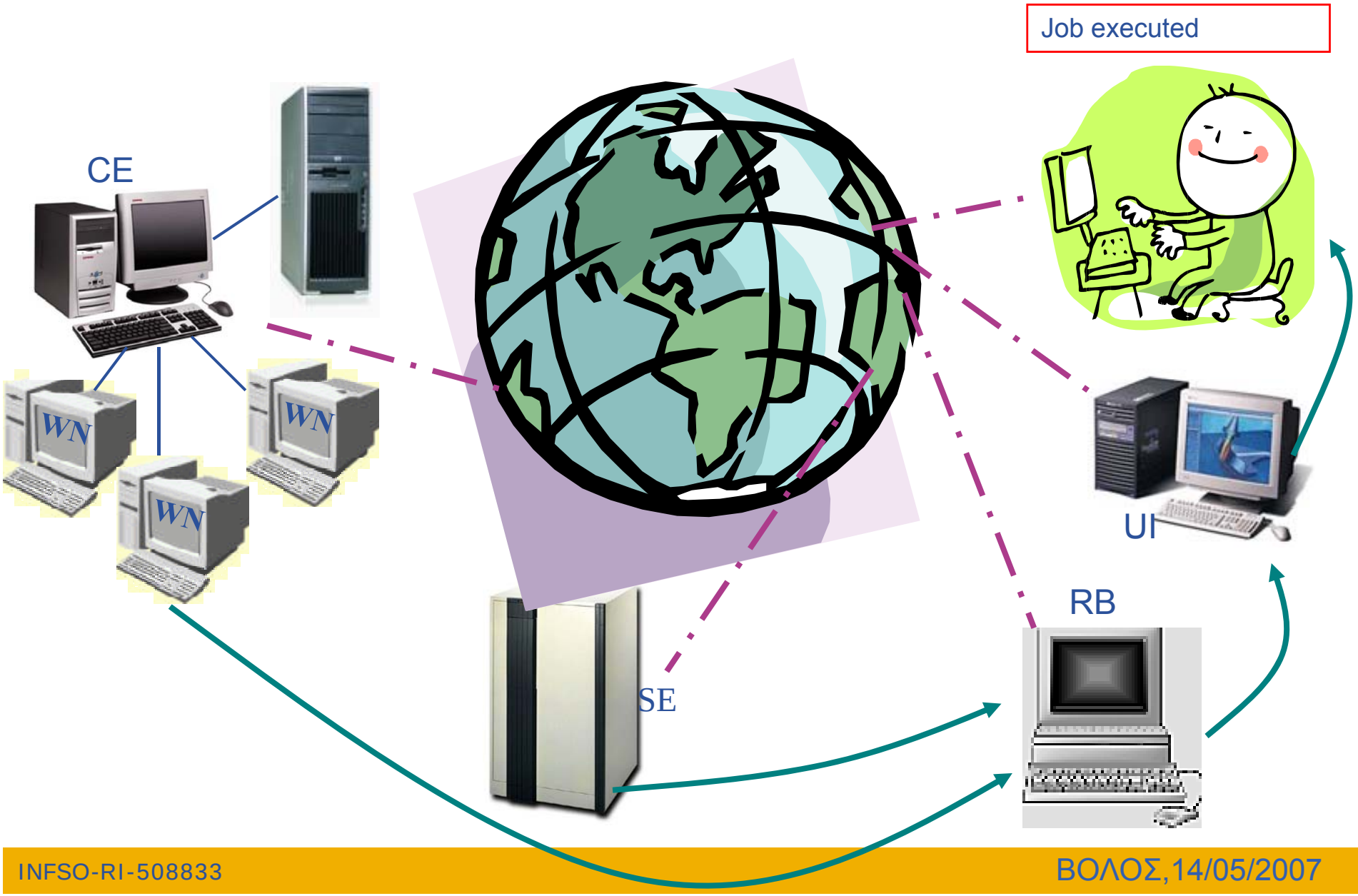
- Service used for accounting, monitoring and publication of user level information
- Presents information as a single virtual database containing a set of virtual tables
- A Consumer runs an SQL query on a table and the Registry selects the best Producers to answer the query through a process called *mediation*



Job submission to Grid



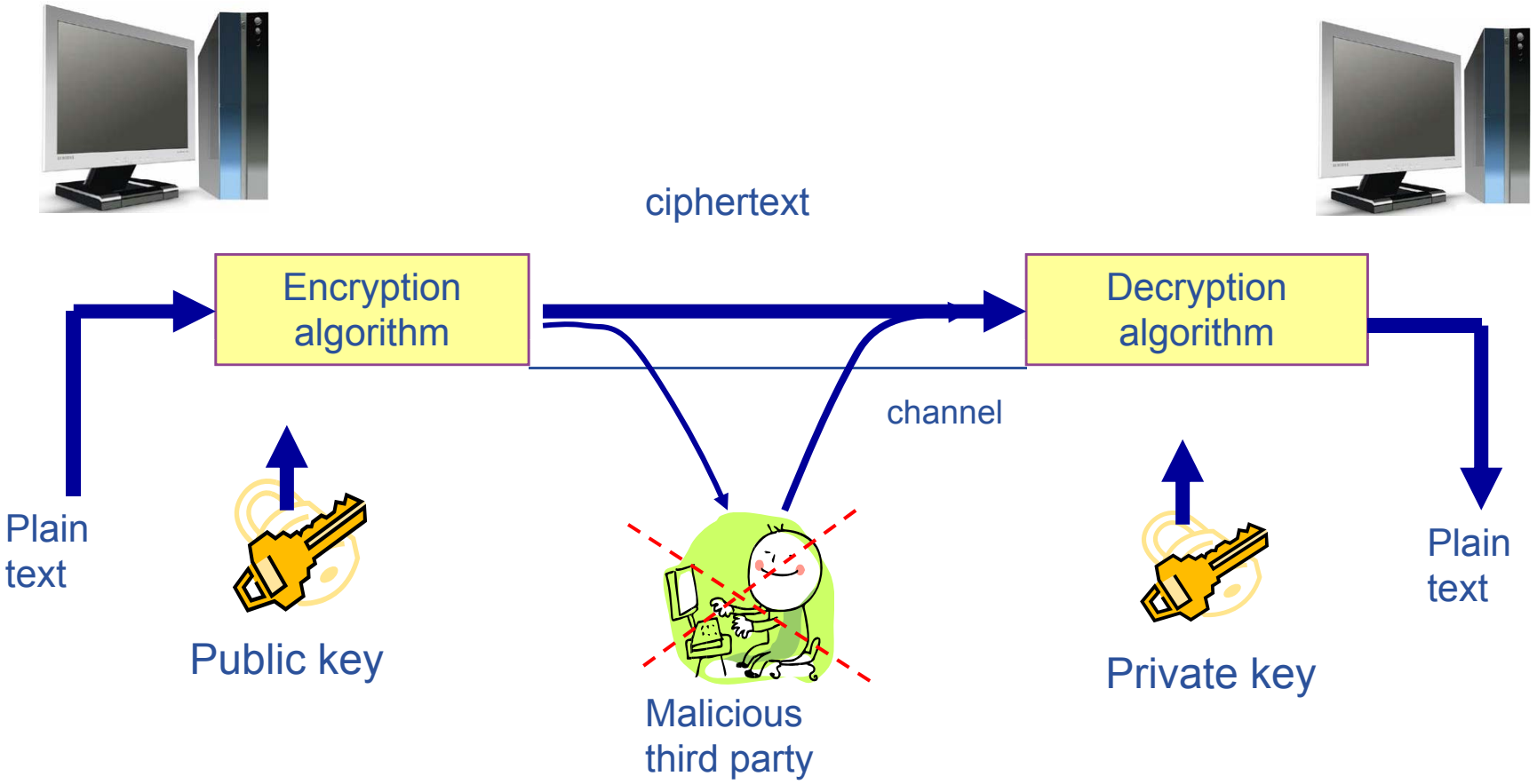
Job submission to Grid



Getting started

- **Obtaining a certificate**
- **Registering with LCG / EGEE**
- **Choosing a VO**
- **Accounts for the training events:**
  - ssh ui01.isabella.grnet.gr (Putty)
    - login as: **egee02 – egee50**

Cryptography components

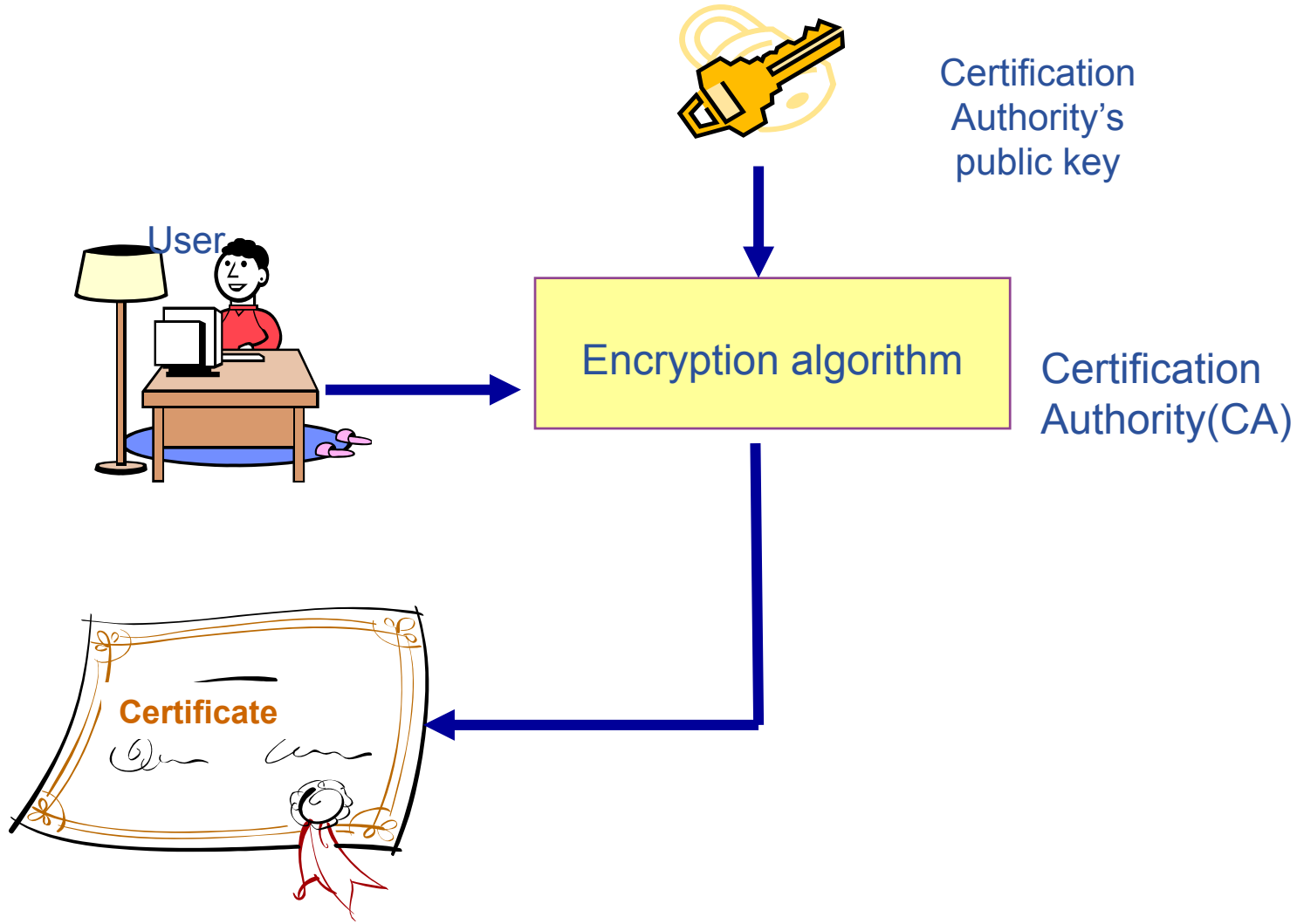


Digital certificate X.509

- **Each entity (user, resource) must obtain a certificate**
- **The certificate includes information, such as the expiration date, the Certification Authority that signed it, the owner's public key and a DN**
- **The DN defines uniquely the owner and has the following fields:**

C = Owner's country
O = Owner's organization
OU = Owner's group
CN = Owner's name

Certificate authority



Installing the certificate

✓ **[egee01@ui01 egee01]$ mkdir .globus**
- Create directory .globus under the user home directory

✓ **[egee01@ui01 egee01]$ ls -la**

drwxrwsr-x 2 egee01 training 4096 Apr 18 14:10 .globus
- List all files of $HOME in long format and do not hide entries starting with **.**

✓ **[egee01@ui01 egee01]$ mv *.pem ~/.globus/usercert.pem**
- Place the user certificate in directory .globus with name usercert.pem

✓ **[egee01@ui01 egee01]$ chmod 444 ~/.globus/usercert.pem**
- The user certificate can be public readable

✓ **[egee01@ui01 egee01]$ mv *.key ~/.globus/userkey.pem**
- Place the user key in directory .globus with name userkey.pem

✓ **[egee01@ui01 egee01]$ chmod 400 ~/.globus/userkey.pem**
- The key must be readable only by the user

✓ **[egee01@ui01 egee01]$ ls -l ~/.globus**

total 12
-r--r--r-- 1 egee01 training 5548 Apr 18 14:06 usercert.pem
-r-------- 1 egee01 training 963 Apr 18 14:06 userkey.pem

Certificate Info

- **Retrieving information of the user certificate**

✓ [egee01@ui01 egee01]$ grid-cert-info

```
    Certificate:
  Data:
    Version: 3 (0x2)
    Serial Number: 1788 (0x6fc)
    Signature Algorithm: sha1WithRSAEncryption
    Issuer: C=GR, O=HellasGrid Demos, OU=Certification Authoritites, CN=HellasGrid Demo CA 2006
    Validity
      Not Before: Apr 16 10:15:02 2007 GMT
      Not After : May  1 10:15:02 2007 GMT
    Subject: C=GR, O=HellasGrid Demos, OU=People, L=VOLOS TRAINING, CN=User 1788
    Subject Public Key Info:
      Public Key Algorithm: rsaEncryption
      RSA Public Key: (1024 bit)
        Modulus (1024 bit):
          00:c6:2e:31:bb:14:12:27:c3:a7:74:1a:48:3a:59:
```

- **Verifying a user certificate**

✓ [egee01@ui01 egee01]$ openssl verify -CApath /etc/grid-security/certificates ./.globus/usercert.pem

./.globus/usercert.pem: OK

Management of the certificates

- **Changing the pass phrase protecting the private key**
  - ✓ [egee01@ui01 egee01]$ grid-change-pass-phrase -file ~/.globus/userkey.pem
- **Checking the certificate subject**
  - ✓ [egee01@ui01 egee01]$ grid-cert-info -subject

  /C=GR/O=HellasGrid Demos/OU=People/L=VOLOS TRAINING/CN=User 1788
- **Checking the certificate expiration date**
  - ✓ [egee01@ui01 egee01]$ grid-cert-info -enddate

  May 1 10:15:02 2007 GMT
- **Checking the CA issued the certificate**
  - ✓ [egee01@ui01 egee01]$ grid-cert-info -issuer

  /C=GR/O=HellasGrid Demos/OU=Certification Authoritites/CN=HellasGrid Demo CA 2006

Proxy certificates

- A new temporal certificate created taking into account the issued certificate by the corresponding CA
  ⇨ a new key pair is created to be used during the period that the proxy is valid
- The new private key is not secured by a password
- The use of a proxy is recommended because:
  - ✓ the *proxy* has a short lifetime
  - ✓ uses a different private key from the issued certificate

Proxy certificate

- **Creating a proxy**

✓ [egee01@ui01 egee01]$ **grid-proxy-init**

```
Your identity: /C=GR/O=HellasGrid Demos/OU=People/L=VOLOS TRAINING/CN=User 1788
 Enter GRID pass phrase for this identity:
Creating proxy ..................................... Done
Your proxy is valid until: Fri Apr 20 01:48:51 2007
```

- **Printing information on a proxy**

✓ [egee01@ui01 egee01]$ **grid-proxy-info**

```
subject  : /C=GR/O=HellasGrid Demos/OU=People/L=VOLOS TRAINING/CN=User 1788/CN=proxy
issuer   : /C=GR/O=HellasGrid Demos/OU=People/L=VOLOS TRAINING/CN=User 1788
identity : /C=GR/O=HellasGrid Demos/OU=People/L=VOLOS TRAINING/CN=User 1788
type     : full legacy globus proxy
strength : 512 bits
path     : /tmp/x509up_u1048
timeleft : 11:55:07
```

- **Destroying a proxy**

✓ [egee01@ui01 egee01]$ **grid-proxy-destroy**

VOMS

- **Virtual Organisation Membership Service (VOMS)**
  - A system which allows a proxy to have extensions containing information:
    - About the VO
    - The groups the user belongs to in the VO
    - Any roles the user is entitled to have

- ***Group*: subset of the VO containing members who share some responsibilities or privileges in the project**
  - Hierarchically organised
  - A user can be a member of any number of groups
  - VOMS proxy contains the list of all groups the user belongs to
  - Group ⇨ privileges the user **ALWAYS** has

- ***Role*: Attribute which typically allows a user to acquire special privileges to perform specific tasks**
  - Role ⇨ privileges the user needs to have only from time to time

VOMS proxies

- **FQAN (Fully Qualified Attribute Name)**

  FQAN = <group name>[/Role=<role name>]

  - **Mapping group and roles roles to specific privileges**
  - /cms/HeavyIons/Role=production.

- **Creating a VOMS proxy**

  ✓ **[egee01@ui01 egee01]$** voms-proxy-init -voms <alias>:<group name>[Role=<role name>]

  where alias: the server to be contacted

- **Retrieving information on a VOMS proxy**

  ✓ **[egee01@ui01 egee01]$** voms-proxy-info -all

A Command-line tool for querying MDS

- **lcg-infosites ⇨ obtain VO-specific information on existing Grid resources**

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

where:

--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: the number of CPUs, running jobs, waiting jobs and CE names (global, no VO-specific information)

se: the names of the SEs supporting the VO, the type of storage system and the used and available space;

-v 1: only the CE / SE names

-v 2: the cluster names, the amount of RAM, the operating system name and version and the processor model

all: the information given by ce and se

closeSE: the names of the CEs supporting the VO and their close SEs

tag: the software tags published by each CE supporting the VO

lfc: the hostname of the LFC catalogues available to the VO

lfcLocal: the hostname of the local LFC catalogues available to the VO

rb: the hostname and port of the RBs available to the VO

dli: the Data Location Index servers available to the VO

dliLocal: the local Data Location Index servers available to the VO

sitenames: the names of all WLCG/EGEE sites;

Examples (1)

- **Obtaining information about computing resources**
  - ✓ **[egee01@ui01 egee01]$ lcg-infosites --vo hgdemo ce**

```
valor del bdii: bdii.isabella.grnet.gr:2170
#CPU    Free    Total Jobs      Running Waiting ComputingElement
----------------------------------------------------------
 11      7       0               0       0       node001.grid.auth.gr:2119/jobmanager-pbs-hgdemo
 20      19      0               0       0       ce02.marie.hellasgrid.gr:2119/jobmanager-pbs-hgdemo
 64      12      0               0       0       ce01.isabella.grnet.gr:2119/jobmanager-pbs-hgdemo
 118     53      0               0       0       ce01.marie.hellasgrid.gr:2119/jobmanager-pbs-hgdemo
 118     53      0               0       0       glite-ce01.marie.hellasgrid.gr:2119/blah-pbs-hgdemo
 122     110     0               0       0       ce01.afroditi.hellasgrid.gr:2119/jobmanager-pbs-hgdemo
[…]
```

- **Obtaining information about storage resources**
  - ✓ **[egee01@ui01 egee01]$ lcg-infosites --vo see se**

```
Avail Space(Kb) Used Space(Kb)  Type    SEs
----------------------------------------------------------
45766072        100254364       n.a     se.phy.bg.ac.yu
28651116        81503928        n.a     se001.grid.bas.bg
170660000       16280000        n.a     se02.marie.hellasgrid.gr
1780407808      3362175488      n.a     se01.isabella.grnet.gr
2730000000      200000000       n.a     se01.marie.hellasgrid.gr
2360000000      550000000       n.a     se01.afroditi.hellasgrid.gr
2500000000      190000000       n.a     se01.kallisto.hellasgrid.gr
2640000000      350000000       n.a     se01.ariagni.hellasgrid.gr
[…]
```

Examples (2)

- **Listing the close Storage Elements**
  - ✓ **[egee01@ui01 egee01]$ lcg-infosites --vo hgdemo closeSE**

```
valor del bdii: bdii.isabella.grnet.gr:2170
Name of the CE: node001.grid.auth.gr:2119/jobmanager-pbs-hgdemo
      node004.grid.auth.gr

Name of the CE: ce02.marie.hellasgrid.gr:2119/jobmanager-pbs-hgdemo
      se02.marie.hellasgrid.gr
      se02.marie.hellasgrid.gr

Name of the CE: ce01.isabella.grnet.gr:2119/jobmanager-pbs-hgdemo
      se01.isabella.grnet.gr
[…]
```

- **Listing the hostname of the LFC catalogues**
  - ✓ **[egee01@ui01 egee01]$ lcg-infosites --vo hgdemo lfc**

    lfc.isabella.grnet.gr
  - ✓ **[egee01@ui01 egee01]$ lcg-infosites --vo atlas lfc**

    prod-lfc-atlas-central.cern.ch
  - ✓ **[egee01@ui01 egee01]$ lcg-infosites --vo see lfc**

    lfc.isabella.grnet.gr

Examples (3)

- **Listing computing and storage resources together**
  - ✓ **[egee01@ui01 egee01]$ lcg-infosites --vo hgdemo all**

```
valor del bdii: bdii.isabella.grnet.gr:2170
#CPU    Free    Total Jobs      Running Waiting ComputingElement
----------------------------------------------------------
 11      8       0               0       0    node001.grid.auth.gr:2119/jobmanager-pbs-hgdemo
 20      20      0               0       0    ce02.marie.hellasgrid.gr:2119/jobmanager-pbs-hgdem
[…]
Avail Space(Kb) Used Space(Kb)  Type    SEs
----------------------------------------------------------
151768428       76384460        n.a     node004.grid.auth.gr
170660000       16280000        n.a     se02.marie.hellasgrid.gr
[…]
```

- **Listing the software tags published by each CE supporting the VO**
  - ✓ **[egee01@ui01 egee01]$ lcg-infosites –vo see tag**

```
valor del bdii: bdii.isabella.grnet.gr:2170
Name of the CE: g02.phy.bg.ac.yu
        VO-seegrid-vive-0.4.2
        VO-seegrid-vive-0.4.3
Name of the CE: ce01.isabella.grnet.gr
        VO-see-R-2.4.0
        VO-see-octave-2.1.73
        VO-seegrid-vive-0.4.3
```

- **Listing all WLCG/EGEE sitenames**
  - ✓ **[egee01@ui01 egee01]$ lcg-infosites –vo hgdemo sitenames**

Job Description Language

- A high-level language based on the ***Classified Advertisement (ClassAd) language***
- JDL describes jobs and aggregates of jobs with arbitrary dependency relations
- JDL specifies the desired job characteristics and constraints, which are taken into account by the WMS to select the best resource to execute the job
- A JDL file consists of lines having the format:

  attribute = expression;

  - Expressions can span several lines, but only the last one must be terminated by a semicolon
  - Literals are enclosed in double quotes
  - " in strings must be escaped with a backslash ("\"Hallo")
  - The character " ' " cannot be used in the JDL
  - Comments of each line begin with # or //
  - Multi-line comments must be enclosed between "/*" and "*/"
  - **No blank characters or tabs should follow the semicolon at the end of a line**

Attributes (1)

Executable	✓ **The value of this attribute is the executable filename or the command to be run by the job** ✓ If the command is already present on the WN, it must be expressed as a absolute path
StdOutput	✓ The name of the files containing the standard output
StdError	✓ The name of the files containing the standard error
StdInput	✓ The names of the files used as Input files
OutputSandbox	✓ The files to be transferred back to the UI after the job is finished
Environment	✓ Modifies the shell environment of the job
Virtual Organisation	✓ Explicitly specify the VO of the user
Requirements	✓ Expresses constraints on the resources where the job should run ✓ Its value is a Boolean expression that must evaluate to true for a job to run on that specific CE **(example: Requirements = other.GlueCEInfoLRMSType == "PBS" && other.GlueCEInfoTotalCPUs > 1;)**

Attributes (2)

RetryCount MaxRetryCount	✓ Times that the WMS automatically resubmitts jobs which failed for some reason (deep resubmission ⇨ when the job failed after started running in a WN)
ShallowRetryCount MaxShalowRetryCount	✓ Times that the WMS automatically resubmitts jobs which failed for some reason (shallow resubmission – gLite)
MyProxyServer	✓ The Proxy server to be used for certificate renewal
Rank	✓ The CE with the highest rank is selected by the WMS to execute a job ✓ by default *Rank* = *other.GlueCEStateEstimatedResponseTime* (but other.GlueCEStateFreeCPUs other.GlueCEStateWaitingJobs)

A simple job

✓ [egee01@ui01 egee01]$ less testJob.sh

```
#!/bin/bash
echo "****** Running...  date ****** "
date
echo "****** Running... hostname ******"
hostname
echo "****** Running... pwd ****** "
pwd
echo "****** Running... ls ****** "
ls -l
echo "****** Running... uname ****** "
uname
echo "****** Running... uptime ****** "
uptime
echo "****** Running... iostat ****** "
iostat
echo "****** Running... free ****** "
free
echo "****** CPU_INFO ****** "
cat /proc/cpuinfo >cpu.out
echo "****** Downloading... A file to
    WN****** "
wget http://www. …
echo "****** Running... ls ****** "
ls -l
echo "****** Printing Input files ****** "
echo "First file:"
cat $1 > >merge.out
echo "Second file:"
cat $2 >> merge.out
```

A simple JDL file

- [egee01@ui01 egee01]$ **less testJob.jdl**

```
Executable = "testJob.sh";
Arguments = "fileA fileB";
StdOutput = "std.out";
StdError = "std.err";
InputSandbox = {“./testJob.sh", “./fileA", “./fileB"};
OutputSandbox = {"std.out", "std.err", "cpu.out", " merge.out" ,"gLite-3-
    UserGuide.pdf"};
```

Job matching to CEs

- **Listing computing elements that match a job description**
  - ✓ [egee01@ui01 egee01]$ **edg-job-list-match --vo hgdemo testJob.jdl**

```
Selected Virtual Organisation name (from proxy certificate extension): hgdemo
Connecting to host rb01.egee-see.org, port 7772

**********************************************************************
                    COMPUTING ELEMENT IDs LIST
 The following CE(s) matching your job requirements have been found:
                *CEId*
 ce01.ariagni.hellasgrid.gr:2119/jobmanager-lcgpbs-hgdemo
 ce01.athena.hellasgrid.gr:2119/jobmanager-pbs-hgdemo
[…]
node001.grid.auth.gr:2119/jobmanager-pbs-hgdemo
**********************************************************************
```

- **glite:**
  - ✓ [egee01@ui01 egee01]$ **glite-job-list-match testJob.jdl**

- **gLite via WMProxy**
  - ✓ [egee01@ui01 egee01]$ **glite-wms-job-submit –a --ranking testJob.jdl**

    --ranking: display the ranking value of each matching resource

Job submission

- **Single Job submission**
  - ✓ **[egee01@ui01 egee01]$ edg-job-submit --vo hgdemo –o test1 testJob.jdl**

```
Selected Virtual Organisation name (from proxy certificate extension): hgdemo
Connecting to host rb01.egee-see.org, port 7772
Logging to host rb01.egee-see.org, port 9002
============= edg-job-submit Success ==============================
 The job has been successfully submitted to the Network Server.
 Use edg-job-status command to check job current status. Your job identifier (edg_jobId) is:
 - https://rb01.egee-see.org:9000/-gL4JRE6UigY1fZ0scyTjw
 The edg_jobId has been saved in the following file:
 /home/training/egee01/test1
==========================================================================
```

  - ✓ **edg-job-submit –o test1 –r ce01.isabella.grnet.gr:2119/jobmanager-pbs-hgdemo testJob.jdl**
    - – -r : sends the job directly to the specified CE
- **glite:**
  - ✓ **[egee01@ui01 egee01]$ glite-job-submit –o test1 testJob.jdl**
- **gLite via WMProxy**
  - ✓ **[egee01@ui01 egee01]$ glite-wms-job-submit –a testJob.jdl**

Job status' retrieval

- **Retrieving the status of a job**
  - ✓ **[egee01@ui01 egee01]$ edg-job-status** *<edg_jobId>*

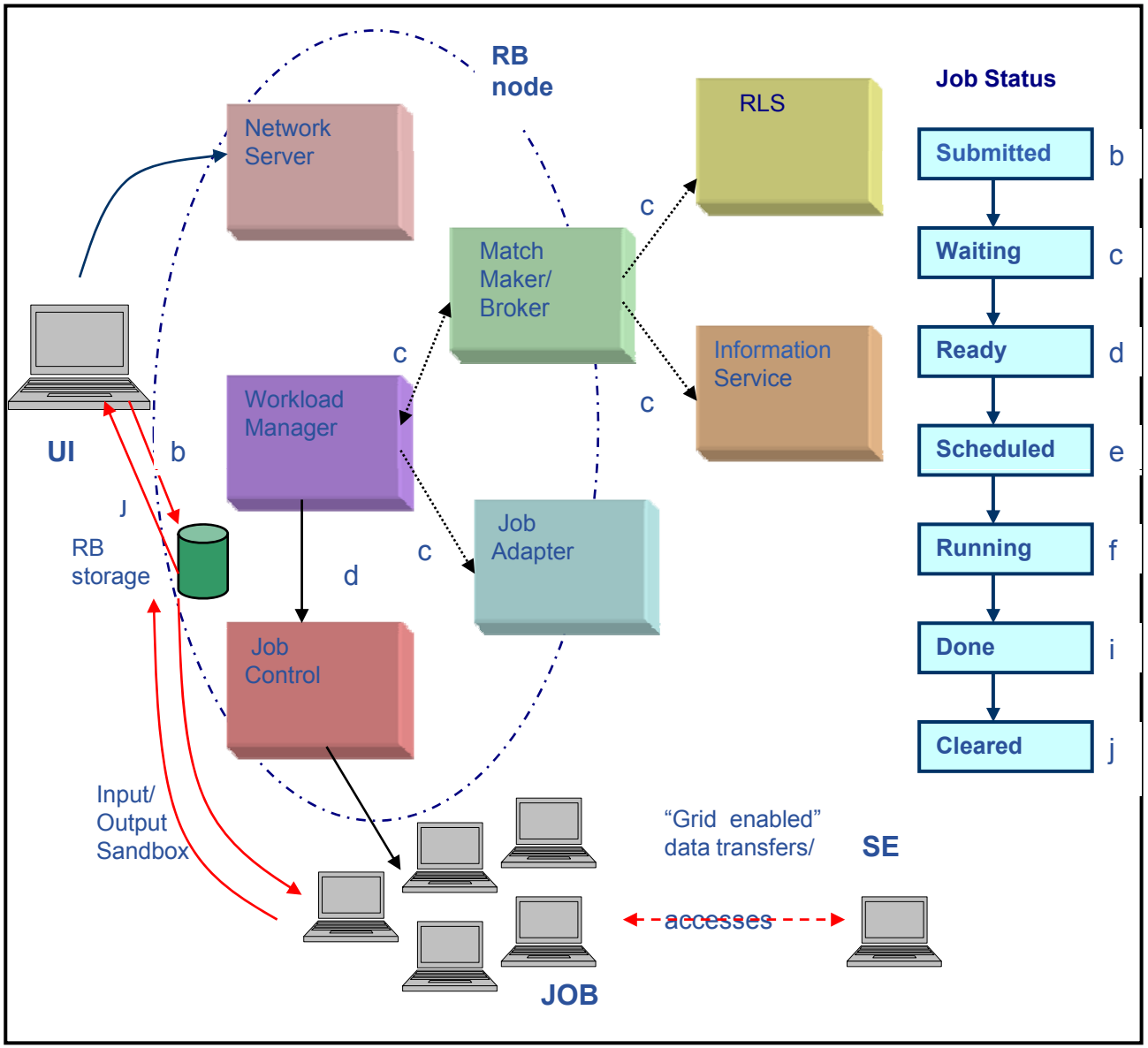
```
**********************************************************************
BOOKKEEPING INFORMATION:

Status info for the Job : https://rb01.egee-see.org:9000/wXt7Cn2MeijygXpD4qV4ng
Current Status:     Scheduled
Status Reason:      Job successfully submitted to Globus
Destination:        ce01.athena.hellasgrid.gr:2119/jobmanager-pbs-hgdemo
reached on:         Thu Apr 19 19:10:31 2007
**********************************************************************
```

  - ✓ **[egee01@ui01 egee01]$ edg-job-status –i test1**
    -i : the jobId read from the file that it was stored
  - ✓ **[egee01@ui01 egee01]$ edg-job-status –all –s <status>**
    Displays all the jobs owned by the user
  - ✓ **[egee01@ui01 egee01]$** watch **"edg-job-status –i test1"**

- **glite:**
  - ✓ **[egee01@ui01 egee01]$ glite-job-status –i test1**

- **gLite via WMProxy**
  - ✓ **[egee01@ui01 egee01]$ glite-wms-job-status –i test1**

Job flow (1)


RB node
Network Server
RLS
Match Maker/ Broker
Information Service
Workload Manager
Job Adapter
Job Control
UI
RB storage
b
c
c
c
c
d
j
Input/ Output Sandbox
"Grid enabled" data transfers/
SE
accesses
JOB
Job Status
Submitted
b
Waiting
c
Ready
d
Scheduled
e
Running
f
Done
i
Cleared
j

Job flow (2)

- **Job submission**
  - The user logs in the UI and submits the job to a Resource broker.
  - If one or more files need to be copied from the UI to the WN, this is specified in the job description and the files are initially copied to the RB. This set of files is called the **Input Sandbox**
  - **Job status ⇨ SUBMITTED**
- **Finding the proper CE**
  - The **WMS** interrogates the **Information Supermarket (ISM)** *(*an internal cache of information read from the BDII) , to determine the status of computational and storage resources.
  - The WMS interogates the File Catalogue to find the location of any required input files
  - **Job status ⇨ WAITING**
- **Job submission from the RB to the selected CE**
  - The RB prepares a wrapper script that will be passed together with other parameters, to the selected CE.
  - **Job status ⇨ READY**
- **Job arrival to the CE**
  - **The CE receives the request and sends the job for execution to the local LRMS.**
  - **Job status ⇨SCHEDULED**

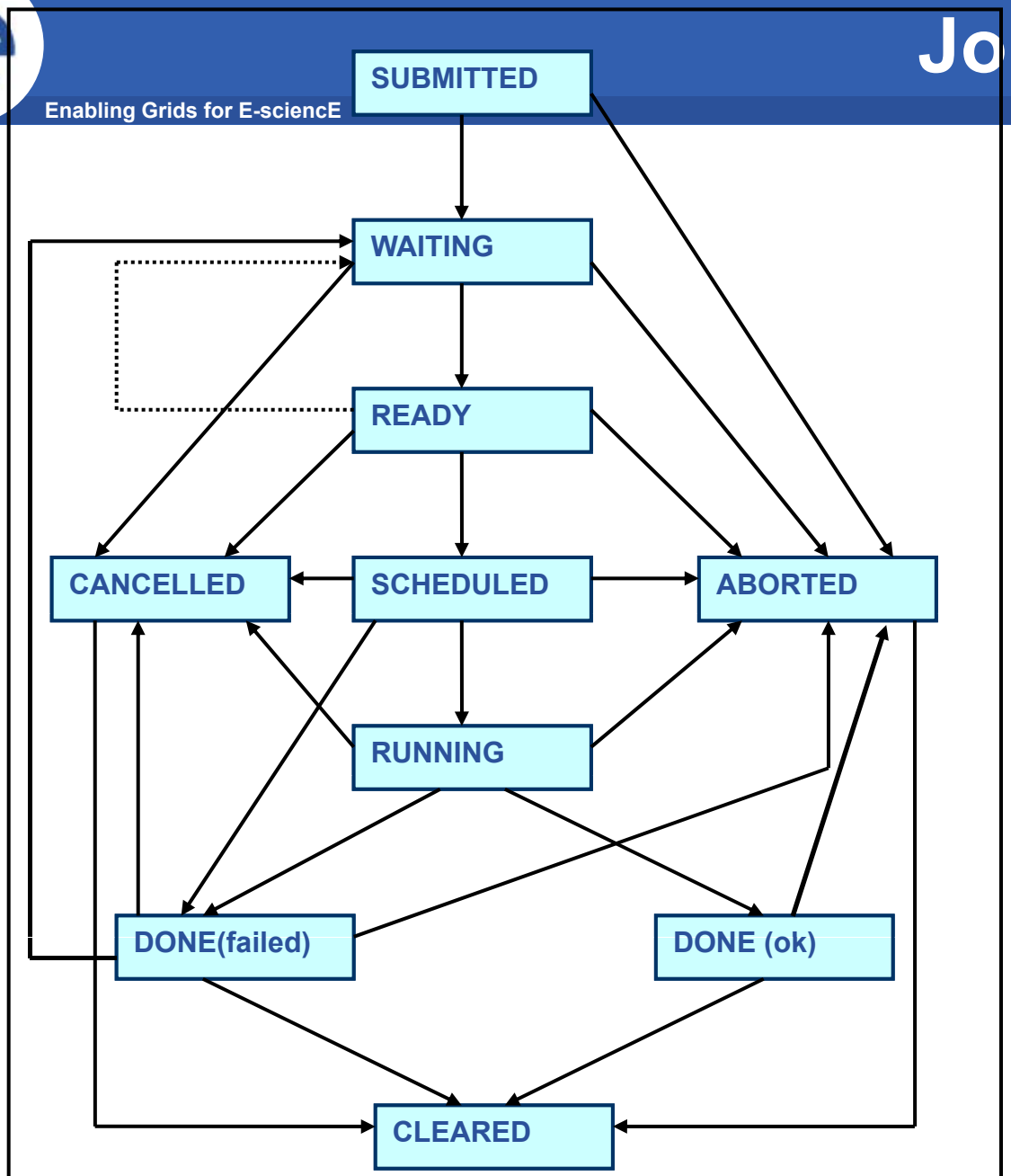
Job flow (3)

- **Job submission to the Worker node**
  - The LRMS handles the execution of jobs on the local Worker Nodes.
  - The Input Sandbox files are copied from the RB to an available WN where the job is executed.
  - While the job runs, Grid files can be directly accessed from an SE using either the RFIO or gsidcap protocol
  - Any new produced output files which can be uploaded to the Grid and made available for other Grid users to use. This can be achieved using the Data Management tools described later. Uploading a file to the Grid means copying it to a Storage Element and registering it in a file catalogue.
  - **Job status ⇨ RUNNING**

- **Job finished**
  - If the job ends without errors, the small output files specified by the user in the ***Output Sandbox*** are transferred back to the RB node.
  - **Job status ⇨ DONE**

- **Output retrieval**
  - The user can retrieve the output files to the UI
  - **Job status ⇨ Cleared**

Job status


SUBMITTED
WAITING
READY
CANCELLED
SCHEDULED
ABORTED
RUNNING
DONE(failed)
DONE (ok)
CLEARED

Job cancelling

- **Cancelling a job**
  - ✓ **[egee01@ui01 egee01]$ edg-job-cancel –i test1 testJob.jdl**

```
Are you sure you want to remove specified job(s)? [y/n]n :y

============================== edg-job-cancel Success ==============================
The cancellation request has been successfully submitted for the following job(s):

- https://rb01.egee-see.org:9000/FN3BMRlxUhH0UVcgu7ptvQ

====================================================================================
```

- **glite:**
  - ✓ **[egee01@ui01 egee01]$ glite-job-cancel –i test1**

- **gLite via WMProxy**
  - ✓ **[egee01@ui01 egee01]$ glite-wms-job-cancel –i test1**

Output retrieval

- If the job's status is DONE, then its output can be copied to the UI with the commands:
  - ✓ [egee01@ui01 egee01]$ edg-job-get-output –i test1 –dir .

```
Retrieving files from host: rb01.egee-see.org ( for https://rb01.egee-see.org:9000/-
    gL4JRE6UigY1fZ0scyTjw )
**********************************************************************************

                    JOB GET OUTPUT OUTCOME
Output sandbox files for the job:
- https://rb01.egee-see.org:9000/-gL4JRE6UigY1fZ0scyTjw
have been successfully retrieved and stored in the directory:
/home/training/egee01/egee01_-gL4JRE6UigY1fZ0scyTjw
**********************************************************************************
```

- glite:
  - ✓ [egee01@ui01 egee01]$ glite-get-output –i test1 –dir .
- gLite via WMProxy
  - ✓ [egee01@ui01 egee01]$ glite-wms-get-output –i test1 –dir .

Retrieval of job logging info

- **Retrieval of logging information about submitted jobs**
  - ✓ **[egee01@ui01 egee01]$ edg-job-get-logging-info –i test1**

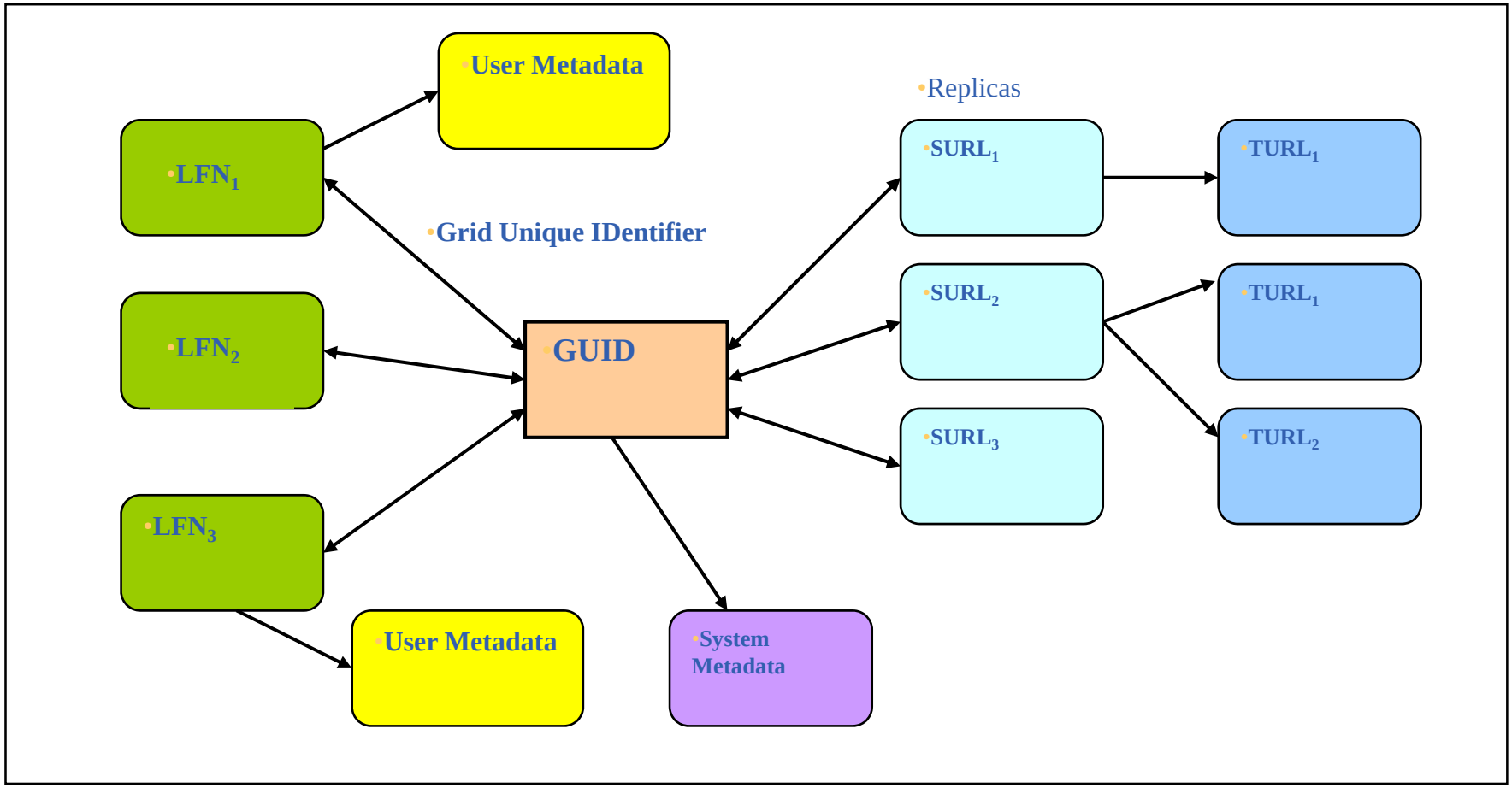
```
*************************************************************************
  LOGGING INFORMATION:
  Printing info for the Job : https://rb01.egee-see.org:9000/-gL4JRE6UigY1fZ0scyTjw
        ---
   Event: RegJob
  - source              =    UserInterface
  - timestamp             =    Thu Apr 19 18:44:57 2007
        ---
   Event: Transfer
  - destination          =    NetworkServer
  - result              =    START
  - source              =    UserInterface
  - timestamp             =    Thu Apr 19 18:44:58 2007
        ---
  […]
  Event: Clear
  - source              =    NetworkServer
  - timestamp             =    Thu Apr 19 19:31:37 2007

  *************************************************************************
```

- **glite:**
  - ✓ **[egee01@ui01 egee01]$ glite-job-logging-info –i test1**
- **gLite via WMProxy**
  - ✓ **[egee01@ui01 egee01]$ glite-wms-logging-info –i test1**

Filenames(1)

- **Grid Unique Identifier (GUID)**
  - Identifies a file uniquely
  - *Example*: guid:ab993b98-8bc9-4984-901e-91290276090c
- **Logical File Name (LFN) (User Alias)**
  - Refers to a file instead of a GUID
  - lfn:<any_string>
  - LFC catalogue: lfn:/grid/<MyVO>/<MyDirs>/<MyFile>
  - *Example*: lfn:/grid/hgdemo/test_egee01/test_file
- **Storage URL (SURL) (Physical File Name-PFN)**
  - Identifies a replica in a SE
  - <sfn|srm>://<SE_hostname>/<some_string>
  - *Example*: sfn://se01.isabella.grnet.gr/storage/hgdemo/generated/2007-04-20/filec4087974-dbaa-4890-91e2-3c105fa0a3df
- **Transport URL (TURL)**
  - A valid URI with the necessary information to access a file in a SE
  - <protocol>://<some_string>
  - *Example*:gsiftp://se01.isabella.grnet.gr/storage/hgdemo/generated/2007-04-20/file1a08d327-d7dc-4d89-bb01-2c86f59eae37

Filenames(2)


User Metadata
LFN1
Grid Unique IDentifier
LFN2
GUID
LFN3
User Metadata
System Metadata
Replicas
SURL1
TURL1
SURL2
TURL1
SURL3
TURL2

File Catalogue

- **File Catalogue (gLite) or LCG File Catalogue**
  - Maintains mappings between LFNs, GUID, SURLs
  - Local File Catalogue, holding only replicas stored at a given site
  - Global File Catalogue, containing information about all files in the Grid
  - Consists of a unique catalogue, where the LFN is the main key
  - System metadata are supported

- **Grid file**
  - Physically present in a SE
  - Registered in the file catalogue

- **High Level Tools (lcg_util) ⇨ Consistency between files in the SEs and entries in the file catalogue**

- **Low level Data Management tools ⇨ Inconsistency between SEs physical files and catalogue entries**

LFC commands

- lfc-ls: List file / directory entries in a directory
- lfc-mkdir: Create directory
- lfc-chmod: Change access mode of a LFC file / directory
- lfc-chown: Change owner and group of a LFC file / directory
- lfc-ln: Make a symbolic link to a file / directory
- lfc-rename: Rename a file / directory
- lfc-rm: Remove a file / directory
- lfc-setcomment: Add / replace a comment
- lfc-delcomment: Delete the comment associated with a file / directory

More advanced LFC commands

- lfc-getacl: Get file / directory access control lists
- lfc-setacl: Set file / directory access control lists
- lfc-entergrpmap: Defines a new group entry in the Virtual ID table
- lfc-enterusrmap: Defines a new user entry in Virtual ID table
- lfc-modifygrpmap: Modifies a group entry corresponding to a given virtual gid
- lfc-modifyusrmap: Modifies a user entry corresponding to a given virtual uid
- lfc-rmgrpmap: Suppresses group entry corresponding to a given virtual gid or group name
- lfc-rmusrmap: Suppresses user entry corresponding to a given virtual uid or user name.

LFN management examples

- **Creating a directory in the LFN namespace**
  - ✓ [egee01@ui01 egee01]$ **lfc-mkdir lfc:/grid/hgdemo/test_egee01**
- **Listing the entries of a LFC directory**
  - ✓ [egee01@ui01 egee01]$ **lfc-ls –l lfc:/grid/hgdemo/**
    drwxrwxr-x 0 26158 32000 0 Apr 20 12:45 test_egee01
- **Adding metadata information to LFC entries**
  - ✓ [egee01@ui01 egee01]$ **lfc-setcomment lfc:/grid/hgdemo/test_egee01/myfile "Created for theTraining"**
- **View metadata of a specific file**
  - ✓ [egee01@ui01 egee01]$ **lfc-ls --comment lfc:/grid/hgdemo/test_egee01/myfile**
    lfc:/grid/hgdemo/test_egee01/myfile Created for VOLOS Training
- **Removing metadata information to LFC entries**
  - ✓ [egee01@ui01 egee01]$ **lfc-delcomment lfc:/grid/hgdemo/test_egee01/myfile**
- **Removing LFNs from the LFC**
  - ✓ [egee01@ui01 egee01]$ **lfc-rm -r lfc:/grid/hgdemo/test_egee01**

LCG Data Management Client Tools

- ***lcg_util* tools**
  - Allow users to copy files between UI, CE, WN and a SE
  - Allow users 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 Data Management Client Tools

- ## File Catalogue Interaction

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

Upload a file

- **Uploading a file to the grid**
  - ✓ [egee01@ui01 egee01]$ lcg-cr --vo hgdemo -d se01.isabella.grnet.gr file:/home/training/egee01/fileA
    guid:80c7741d-f6bf-40a4-8fbb-936cdcde583c

- **Listing the LFNs associated to this file**
  - ✓ **[egee01@ui01 egee01]$ lcg-la --vo hgdemo guid:80c7741d-f6bf-40a4-8fbb-936cdcde5**
    lfn:/grid/hgdemo/generated/2007-04-20/file-94e97a14-469e-4540-a42e-330d61aae22c83c

- **Uploading a file to the grid with specific LFN**
  - ✓ **[egee01@ui01 egee01]$ lcg-cr --vo hgdemo -d se01.isabella.grnet.gr -l **
    lfn:/grid/hgdemo/test_egee01/myfile file:/home/training/egee01/fileB
    guid:4660d8a9-197e-4e8d-a58a-67f2fcb4e2d0

  - ✓ **[egee01@ui01 egee01]$ lcg-la --vo hgdemo guid:4660d8a9-197e-4e8d-a58a-67f2fcb4e2d0**
    lfn:/grid/hgdemo/test_egee01/myfile

  - ✓ **[egee01@ui01 egee01]$ lfc-ls lfc:/grid/hgdemo/test_egee01/**
    myfile

File management operations (1)

- **Replicating a file**

✓ [egee01@ui01 egee01]$ **lcg-rep -v --vo hgdemo -d se01.marie.hellasgrid.gr guid:80c7741d-f6bf-40a4-8fbb-936cdcde583c**

(guid of fileA)

```
Destination URL for copy: gsiftp://se01.marie.hellasgrid.gr/se01.marie.hellasgrid.gr:/data02/hgdemo/2007-04-20/file1429bc11-5bc8-4b23-af3a-09aca4dfed7e.101918.0
# streams: 1
# set timeout to 0
        0 bytes      0.00 KB/sec avg      0.00 KB/sec inst
Transfer took 2020 ms
Destination URL registered in LRC: srm://se01.marie.hellasgrid.gr/dpm/marie.hellasgrid.gr/home/hgdemo/generated/2007-04-20/file1429bc11-5bc8-4b23-af3a-09aca4dfed7e
```

✓ [egee01@ui01 egee01]$ **lcg-rep -v --vo hgdemo -d se01.kallisto.hellasgrid.gr lfn:/grid/hgdemo/test_egee01/myfile**

(fileB)

```
Using grid catalog type: lfc
Using grid catalog : lfc.isabella.grnet.gr
Source URL: lfn:/grid/hgdemo/test_egee01/myfile
File size: 54
VO name: hgdemo
Destination specified: se01.kallisto.hellasgrid.gr
Source URL for copy: gsiftp://se01.isabella.grnet.gr/storage/hgdemo/generated/2007-04-20/file1a08d327-d7dc-4d89-bb01-2c86f59eae37
Destination URL for copy: gsiftp://se01.kallisto.hellasgrid.gr/se01.kallisto.hellasgrid.gr:/data01/hgdemo/2007-04-20/file9264a98d-d406-4f9a-8e2c-9d58a8a56714.107843.0
# streams: 1
# set timeout to 0
        0 bytes      0.00 KB/sec avg      0.00 KB/sec inst
Transfer took 2020 ms
Destination URL registered in LRC: srm://se01.kallisto.hellasgrid.gr/dpm/kallisto.hellasgrid.gr/home/hgdemo/generated/2007-04-20/file9264a98d-d406-4f9a-8e2c-9d58a8a56714
```

Replicas listing

- **Listing replicas**
  - ✓ **[egee01@ui01 egee01]$ lcg-lr --vo hgdemo lfn:/grid/hgdemo/test_egee01/myfile**

    sfn://se01.isabella.grnet.gr/storage/hgdemo/generated/2007-04-20/file1a08d327-d7dc-4d89-bb01-2c86f59eae37

    srm://se01.kallisto.hellasgrid.gr/dpm/kallisto.hellasgrid.gr/home/hgdemo/generated/2007-04-20/file9264a98d-d406-4f9a-8e2c-9d58a8a56714

  - ✓ **[egee01@ui01 egee01]$ lcg-lr --vo hgdemo guid:4660d8a9-197e-4e8d-a58a-67f2fcb4e2d0**

    sfn://se01.isabella.grnet.gr/storage/hgdemo/generated/2007-04-20/file1a08d327-d7dc-4d89-bb01-2c86f59eae37

    srm://se01.kallisto.hellasgrid.gr/dpm/kallisto.hellasgrid.gr/home/hgdemo/generated/2007-04-20/file9264a98d-d406-4f9a-8e2c-9d58a8a56714

  - ✓ **[egee01@ui01 egee01]$ lcg-lr --vo hgdemo sfn://se01.isabella.grnet.gr/storage/hgdemo/generated/2007-04-20/file1a08d327-d7dc-4d89-bb01-2c86f59eae37**

    sfn://se01.isabella.grnet.gr/storage/hgdemo/generated/2007-04 20/file1a08d327-d7dc-4d89-bb01-2c86f59eae37

    srm://se01.kallisto.hellasgrid.gr/dpm/kallisto.hellasgrid.gr/home/hgdemo/generated/2007-04-20/file9264a98d-d406-4f9a-8e2c-9d58a8a56714

Map SURL / LFN to GUID

- **Listing guid**
  - Listing guids given the lfn

✓ **[egee01@ui01 egee01]$ lcg-lg --vo hgdemo lfn:/grid/hgdemo/test_egee01/myfile**

guid:4660d8a9-197e-4e8d-a58a-67f2fcb4e2d0

  - Listing guids given the sfn

✓ **[egee01@ui01 egee01]$ lcg-lg --vo hgdemo sfn://se01.isabella.grnet.gr/storage/hgdemo/generated/2007-04-20/file1a08d327-d7dc-4d89-bb01-2c86f59eae37**

guid:4660d8a9-197e-4e8d-a58a-67f2fcb4e2d0

Map GUID / SURL to LFN

- **Listing LFNs**

  ✓ **[egee01@ui01 egee01]$ lcg-la --vo hgdemo guid:4660d8a9-197e-4e8d-a58a-67f2fcb4e2d0**

  lfn:/grid/hgdemo/test_egee01/myfile

- ✓ **[egee01@ui01 egee01]$ lcg-la --vo hgdemo sfn://se01.isabella.grnet.gr/storage/hgdemo/generated/2007-04-20/file1a08d327-d7dc-4d89-bb01-2c86f59eae37**

  lfn:/grid/hgdemo/test_egee01/myfile

JDL file

```
Executable = "testJob.sh";
Arguments = "fileB fileA";
StdOutput = "std.out";
StdError = "std.err";
InputSandbox = {"./testJob.sh","./fileA","./fileB"};
OutputSandbox = {"std.out", "std.err", " cpu.out" ,"gLite-3-UserGuide.pdf"};
OutputData ={
            [
            OutputFile = "merge.out";
            LogicalFileName = "lfn:/grid/hgdemo/test_egee01/myfile2";
            StorageElement = "se01.isabella.grnet.gr";
            ]
      };
Rank = other.GlueCEStateFreeCPUs;
Requirements = other.GlueCEInfoTotalCPUs>1;
```

Job execution

- **Find CEs matching the job's requirements**
  edg-job-list-match --vo hgdemo --rank testJob2.jdl

- **Submit job**
  edg-job-submit --vo hgdemo –o test2 testJob2.jdl

- **Watch the job status**
  watch "edg-job-status –i test2 "

- **Retrieve the job output**
  edg-job-get-output –i test2 –dir .

LFC entries' management examples

- **Listing the entries of a LFC directory**
  - ✓ [egee01@ui01 egee01]$ lfc-ls –l lfc:/grid/hgdemo/
    drwxrwxr-x 0 26158 32000 0 Apr 20 12:45 test_egee01
- **Adding metadata information to LFC entries**
  - ✓ [egee01@ui01 egee01]$ lfc-setcomment lfc:/grid/hgdemo/test_egee01/myfile "Created for theTraining"
- **View metadata of a specific file**
  - ✓ [egee01@ui01 egee01]$ lfc-ls --comment lfc:/grid/hgdemo/test_egee01/myfile
    lfc:/grid/hgdemo/test_egee01/myfile Created for VOLOS Training
- **Copying a file out of the Grid**
  - ✓ [egee01@ui01 egee01]$ lcg-cp --vo hgdemo -t 100 –v lfn:/grid/hgdemo/test_egee01/myfile2 file:/home/training/egee01/merge
  - ✓ [egee01@ui01 egee01]$ ls
  - ✓ [egee01@ui01 egee01]$ less merge

Delete Grid files and LFC entries

- **Deleting all replicas with the specific lfn**
  - ✓ **[egee01@ui01 egee01]$ lcg-del --vo hgdemo -a lfn:/grid/hgdemo/test_egee01/myfile2**
  - ✓ **[egee01@ui01 egee01]$ lcg-lr --vo hgdemo lfn:/grid/hgdemo/test_egee01/mylife2**
    lcg_lr: No such file or directory
- **Deleting the file with given guid in the specific SE**
  - ✓ **[egee01@ui01 egee01]$ lcg-lg --vo hgdemo lfn:/grid/hgdemo/test_egee01/myfile**
    guid:a34193ac-9e7b-4b73-8c63-517ca59223e3
  - ✓ **[egee01@ui01 egee01]$ lcg-del --vo hgdemo -s se01. kallisto.hellasgrid.gr guid:e8fbab0d-9cff-40c0-a24e-9c665b144d1f**
- **Deleting all replicas with the specific guid**
  - ✓ **[egee01@ui01 egee01]$ lcg-lg --vo hgdemo lfn:/grid/hgdemo/test_egee01/myfile**
    guid:a34193ac-9e7b-4b73-8c63-517ca59223e3
  - ✓ **[egee01@ui01 egee01]$ lcg-del --vo hgdemo -a guid:e8fbab0d-9cff-40c0-a24e-9c665b144d1f**
- **Delete LFC directory**
  - ✓ **[egee01@ui01 egee01]$ lfc-ls lfc:/grid/hgdemo/test_egee01**
  - ✓ **[egee01@ui01 egee01]$ lfc-rm -r lfc:/grid/hgdemo/test_egee01**
  - ✓ **[egee01@ui01 egee01]$ lfc-ls lfc:/grid/hgdemo/**

Q&A

Thank you !