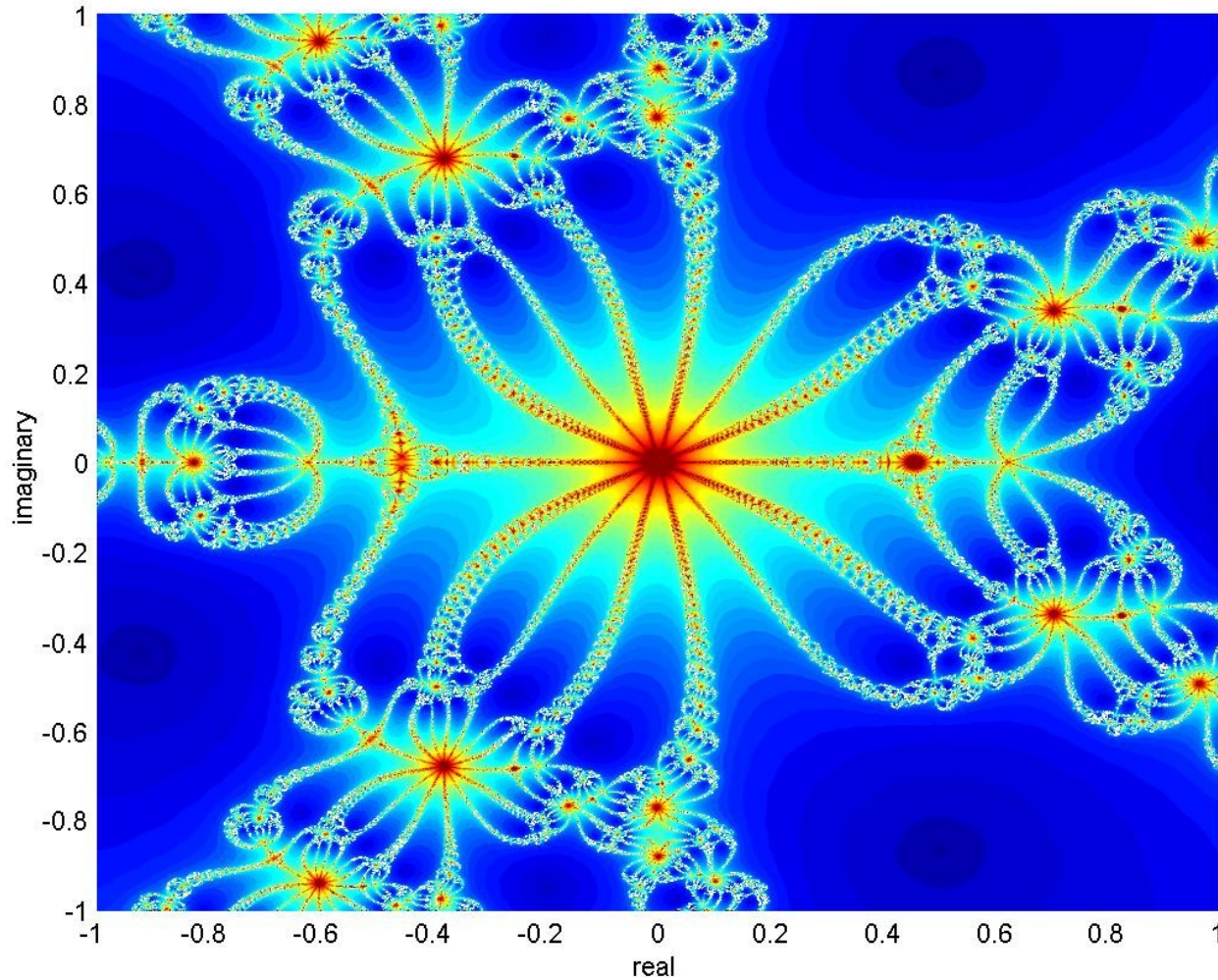
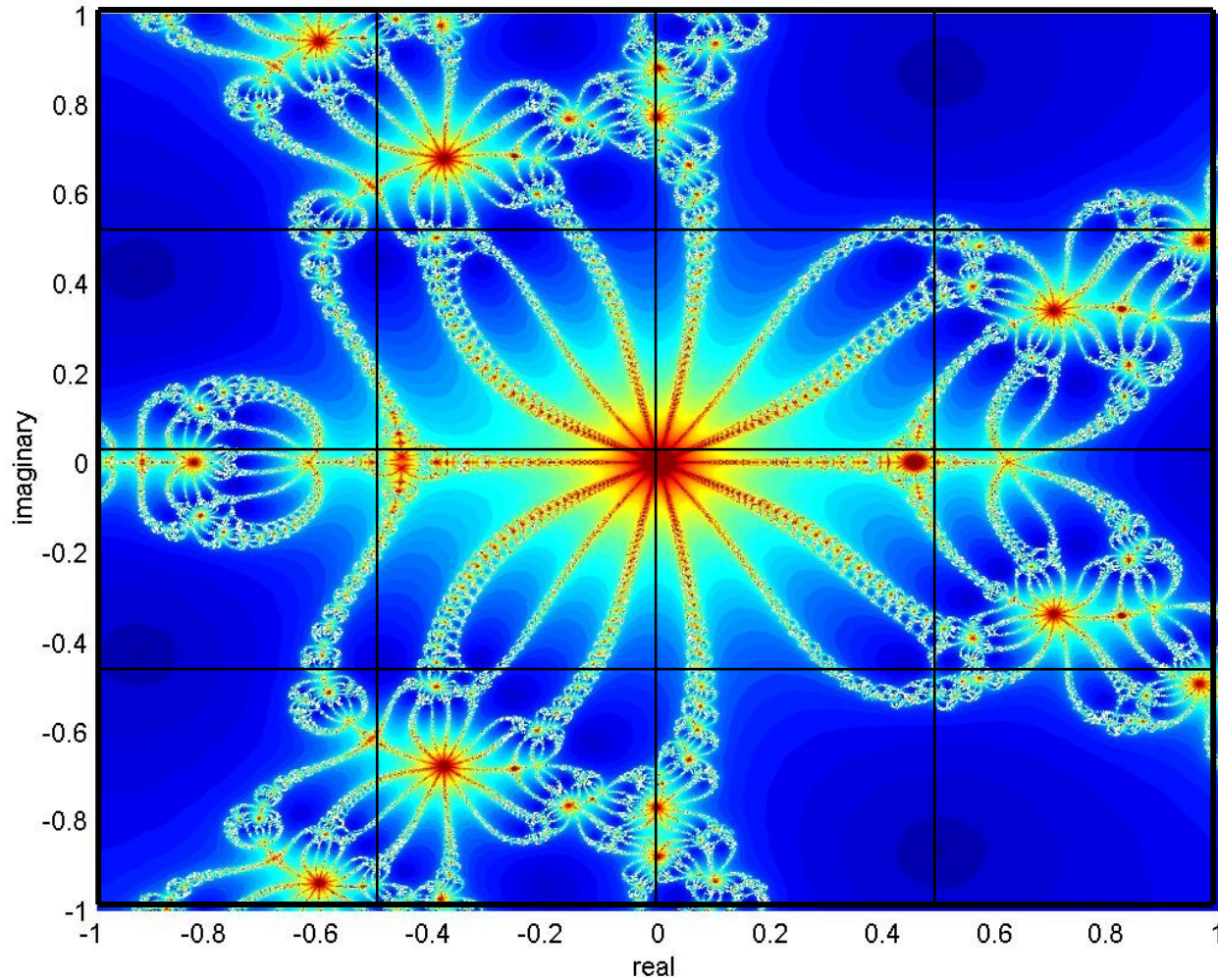


# Using a certificate and simple job submission

How to get started

Fotis Georgatos <fotis@mail.cern.ch>  
(and a cast of thousands)





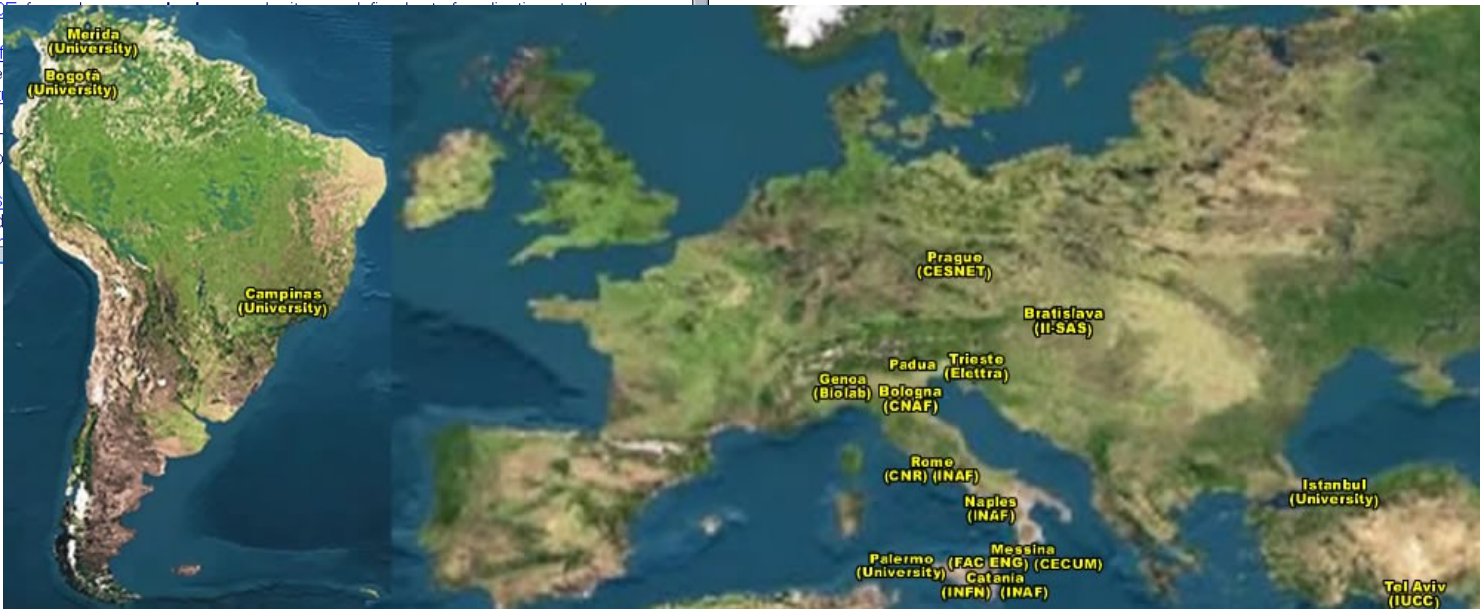
- **We are using the GILDA testbed today**
  - The production EGEE grid looks like this!
  - Current EGEE production middleware
  
- **The practical exercises are to illustrate “how”**
  - How grid services are used
  - How jobs are submitted
  - How output is retrieved
  - ...
  
- **We will use the Command-Line Interfaces on a “User Interface” (UI) machine**
  - “UI” is your interface to the GILDA Grid
    - Where your digital credentials are held
    - Middleware client tools are already installed

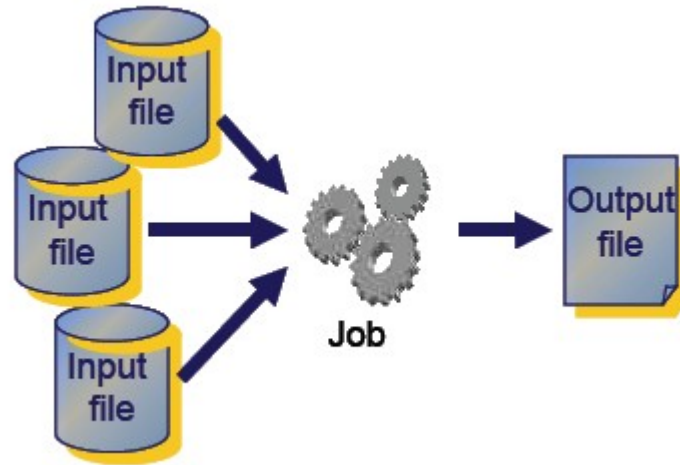
- **Why t-infrastructure?**
  - Training is necessary!
  - e-Infrastructure for production
  - t-Infrastructure for training
- **Need guaranteed response for tutorials; and limit the exposure of production systems**
  - use training grid
  - have training CA
  - able to change middleware to prepare participants for future releases on production system
  - need safe resources for installation training
  - easy entry point for new communities

Enabling Grids for E-science



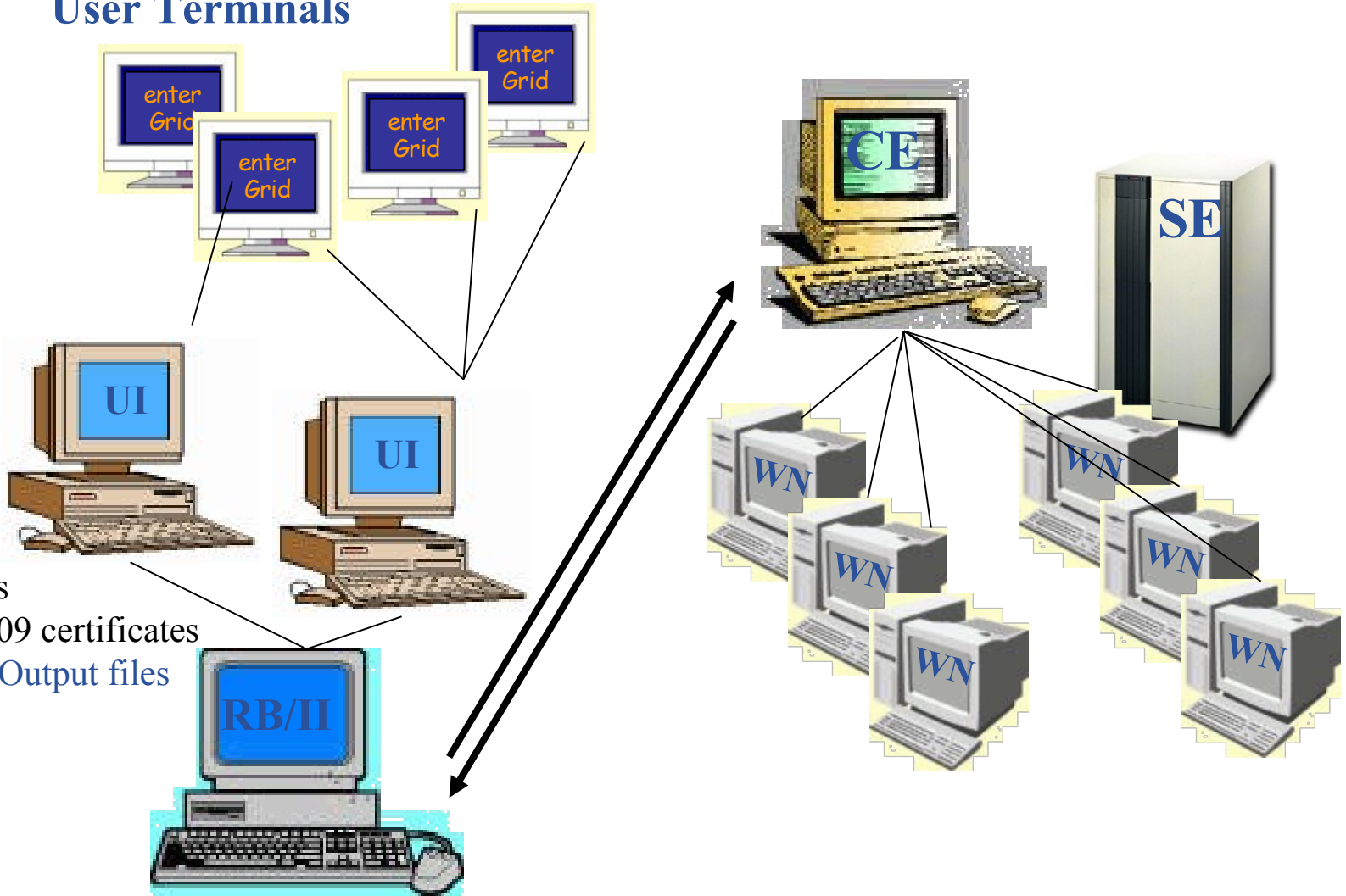
- 19 sites in 3 continents
- > 3000 certificates issued, >15% renewed at least once
- > 100 tutorials and demos performed in 23 months
- > 1,000,000 hits (> 50,000 unique visits) on (of) the web site from 10's of different countries
- > 0.6 TB of training material downloaded from the web site





- **The user requests the execution of an application at a remote system, describing .jdl: input(s), job, output(s).**
- **The application handles input data (that might exist on the Grid) and produces some output data, which might also be stored on the Grid, either on a Storage Element or the User Interface itself.**

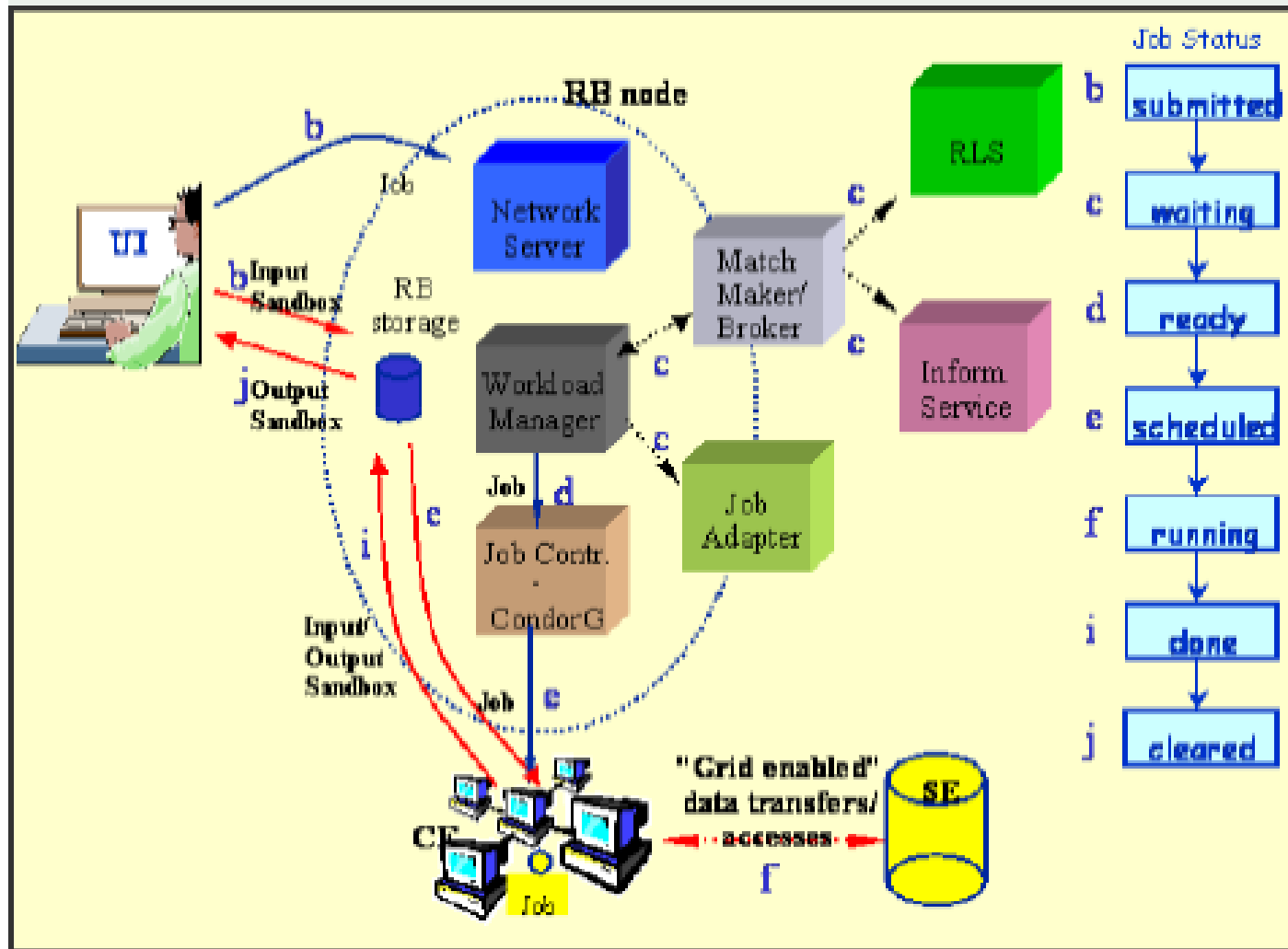
## User Terminals



- JDL files
- PKI X.509 certificates
- Input & Output files

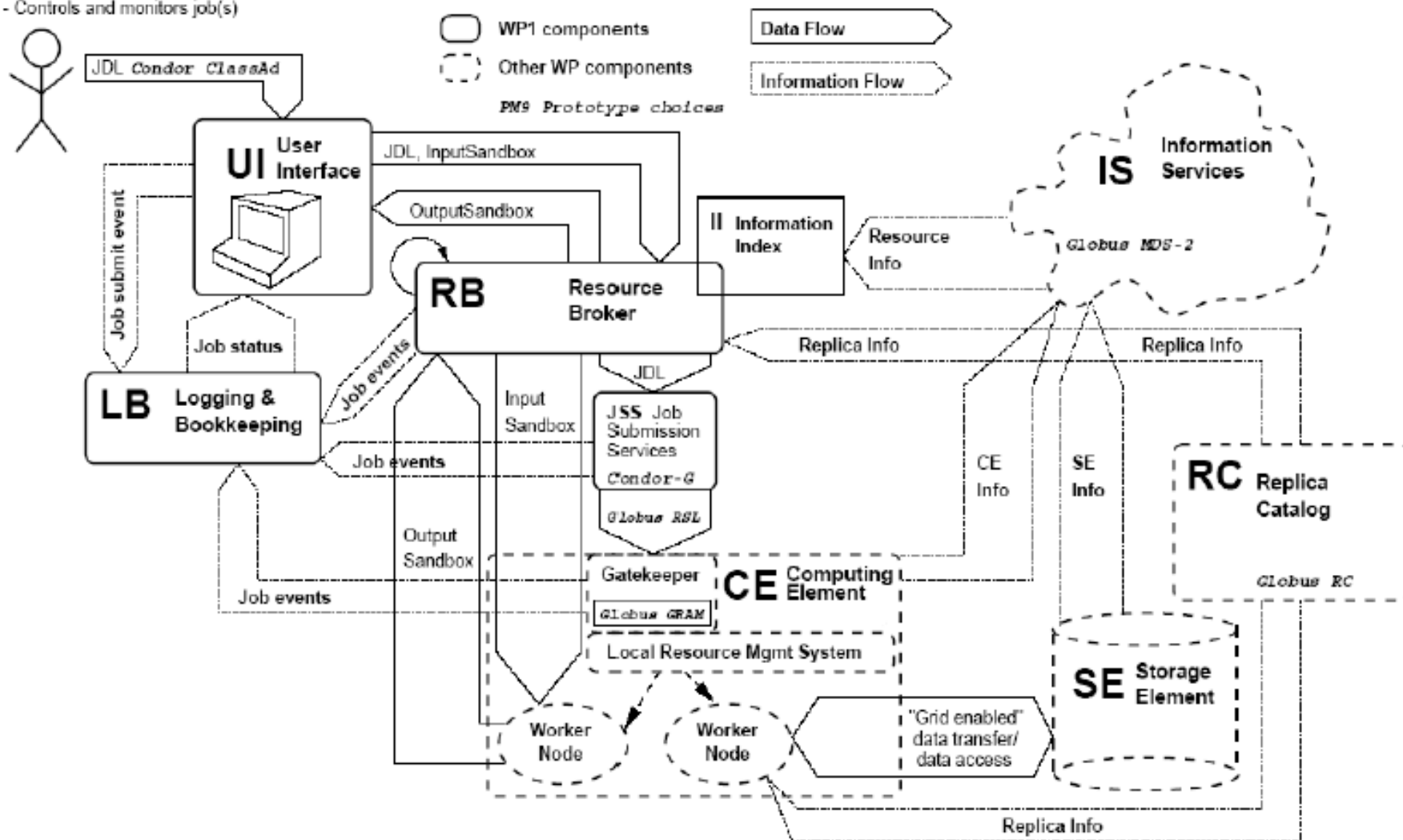


- **UI: User Interface**
  - The system through which Users submit their jobs
  - It might be any Linux system, including your own Laptop.
- **RB: Resource Broker (or WMS)**
  - The node where User's jobs arrive, before being sent to a CE
- **CE: Computing Element**
  - A front-end node that provides access to a set of Worker Nodes
- **WN: Worker Node**
  - The nodes that provide the computational services to the Grid
- **SE: Storage Element**
  - These nodes provide access to disk and tape subsystems
- **BDII: Berkeley Database Information Index**
  - It provides the information «Where/who/what is the Grid»



## End User

- Specifies job using JDL
- Submits job using UI
- Controls and monitors job(s)



- **Log on to desktop (please, work in pairs, collaborate!)**
- **You have been given a user number**
- **Get an ssh client, eg. putty**
  - You can easily find it, fi. with google, or going directly to the address:  
<http://www.putty.nl/download.html>
- **Then use putty to connect to glite-tutor.ct.infn.it**
  - usernames --> sofia01....25
  - passwords --> GridSOF01....25
  - Certificate passphrase: SOFIA
  - **ALWAYS USE THE SAME ACCOUNT DURING ALL OF THIS WEEK – REMEMBER YOUR NUMBER!!**

- **Once in the Linux system with User Interface middleware:**
  - `ls -als .globus`
- **Your certificate lives in the file `~/.globus/usercert.pem`**
- **Your private key lives in the file `~/.globus/userkey.pem`**
- **Activation of the certificate (by default for 12 hours)**
  - `voms-proxy-init --voms gilda`, then reply with “SOFIA”
  - `voms-proxy-info --all`
    - to confirm that the certificate is already active
    - to observe that you have two credentials (from CA and from VOMS)
- **`wget http://cern.ch/fotis/hello.jdl`**
- **`wget http://cern.ch/fotis/hello.sh`**

- **Job Description Language**

- According to the standard of Condor classified advertisements
- The information hereby included is used for optimal assignment of a job

- Example of a **hostname.jdl**:

- **Executable** = `"/bin/hostname"`;
- **Arguments** = `"-f"`;
- **StdOutput** = `"std.out"`;
- **StdError** = `"std.err"`;
- **OutputSandbox** = `{"std.out","std.err"}`;
- **VirtualOrganisation** = `"gilda"`;
- **Requirements** = `other.arch="i386"`;
- **Rank** = `other.FreeCPUs`;

- **edg-job-submit -o myjobids.txt hello.jdl**
  - It sends the job to the Resource Broker
  - RB finds the most suitable CE and assigns the job to it
  - CE identifies a free WN and assigns the job to it
- **edg-job-status -i myjobids.txt**
  - submitted: Job has been submitted to the RB
  - waiting: Job awaits processing at the RB
  - ready: Job has been assigned to CE, but is not yet into the LRMS
  - scheduled: Job is in the queue of the LRMS at the CE
  - running: Job is running within a Worker Node
  - done: Job is over at the RB and the collection by an UI is pending

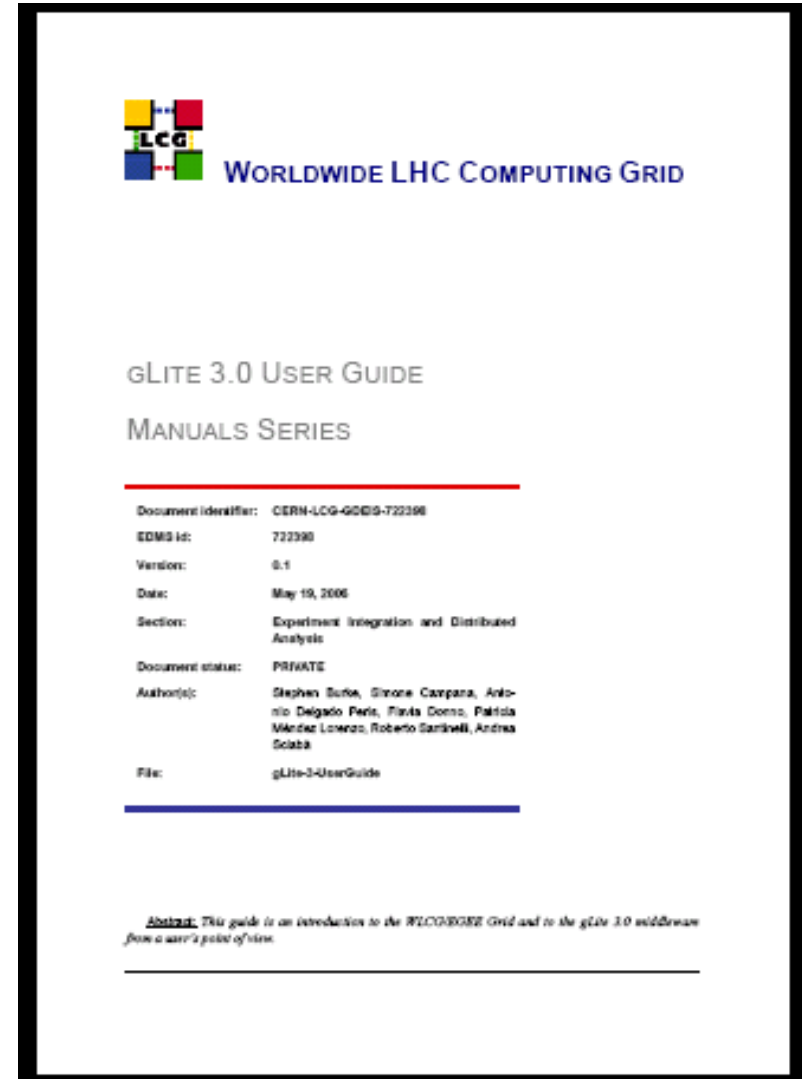
## GLITE 3.0 USER GUIDE

- Instructions about gLite, with many middleware commands, a .pdf of 166 pages.
  - Getting started
  - Information Service
  - Workload Management & .jdl
  - Data Management
  - User tools

• **RTFM!**

(=Read The Fine Manual!)

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





- **Medical/Healthcare**

Imaging  
 Diagnosis & Treatment  
 Drug design (malaria, d2ol, H5N1, anthrax, embola)

- **Bioinformatics**

Study of the human/other genomes (genome@home)  
 Protein folding (folding@home, predictor@home)

- **Geological & climate applications**

Weather Forecasting  
 Climate Simulation (climate@home)  
 Ocean current analysis  
 Oil and Gas Exploration  
 Seismic Signal Analysis

- **Pharmaceutical, Chemical, Biotechnology**

atmospheric chemistry  
 systems biology  
 molecular modeling  
 nanotechnology  
 materials science

- material interaction simulations
- catalysis investigations)

- **Digital Libraries**

- **Mathematics and Basic Research**

prime numbers (gimps/mprimes effort)  
 The (partial) verification of Riemann's hypothesis

- **Business decision support**

Financial analysis  
 Portfolio optimization  
 Risk management applications  
 Supply Chain and Demand Chain Optimization  
 Search and Retrieval (huge databases, data mining)  
 Route Optimization

- Transportation
- LAN and WAN Networking

- **Electrical, Mechanical and Civil Engineering**

Energy production/distribution strategy optimization  
 Engineering and digital design  
 CAD / CAM  
 Aerodynamic simulation (wind tunnel simulation)  
 Digital Rendering (raytracing, digital video synthesis)  
 Construction verification against earthquakes

- eg. finite elements method

- **Physics & Astrophysics**

High Energy Physics simulations and signal analysis  
 N-body problem simulation  
 space probe signal analysis (einstein@home)  
 radio telescope signal analysis (seti@home)

- **Computer Science**

Cryptography (distributed.net)  
 Search Engines (grud, a distributed Internet crawler)

Many! <http://distributedcomputing.info/distrib-2003/distrib-projects.html>

- **edg-job-submit -o myjobids.txt hello.jdl**
  - It sends the job to the Resource Broker
  - RB finds the most suitable CE and assigns the job to it
  - CE identifies a free WN and assigns the job to it
- **edg-job-status -i myjobids.txt**
  - submitted: Job has been submitted to the RB
  - waiting: Job awaits processing at the RB
  - ready: Job has been assigned to CE, but is not yet into the LRMS
  - scheduled: Job is in the queue of the LRMS at the CE
  - running: Job is running within a Worker Node
  - done: Job is over at the RB and the collection by an UI is pending
- **edg-job-cancel , this is only useful for cancelling a job**
- **edg-job-get-output -i myjobids.txt --dir .**

- **Go to the agenda page for this event as follows:**
  - Go to <http://www.egee.nesc.ac.uk/schedreg/>
  - Click on this event and bookmark it
- **Further Information**
  - Look for <more information> link on the agenda

