



Enabling Grids for E-scienceE

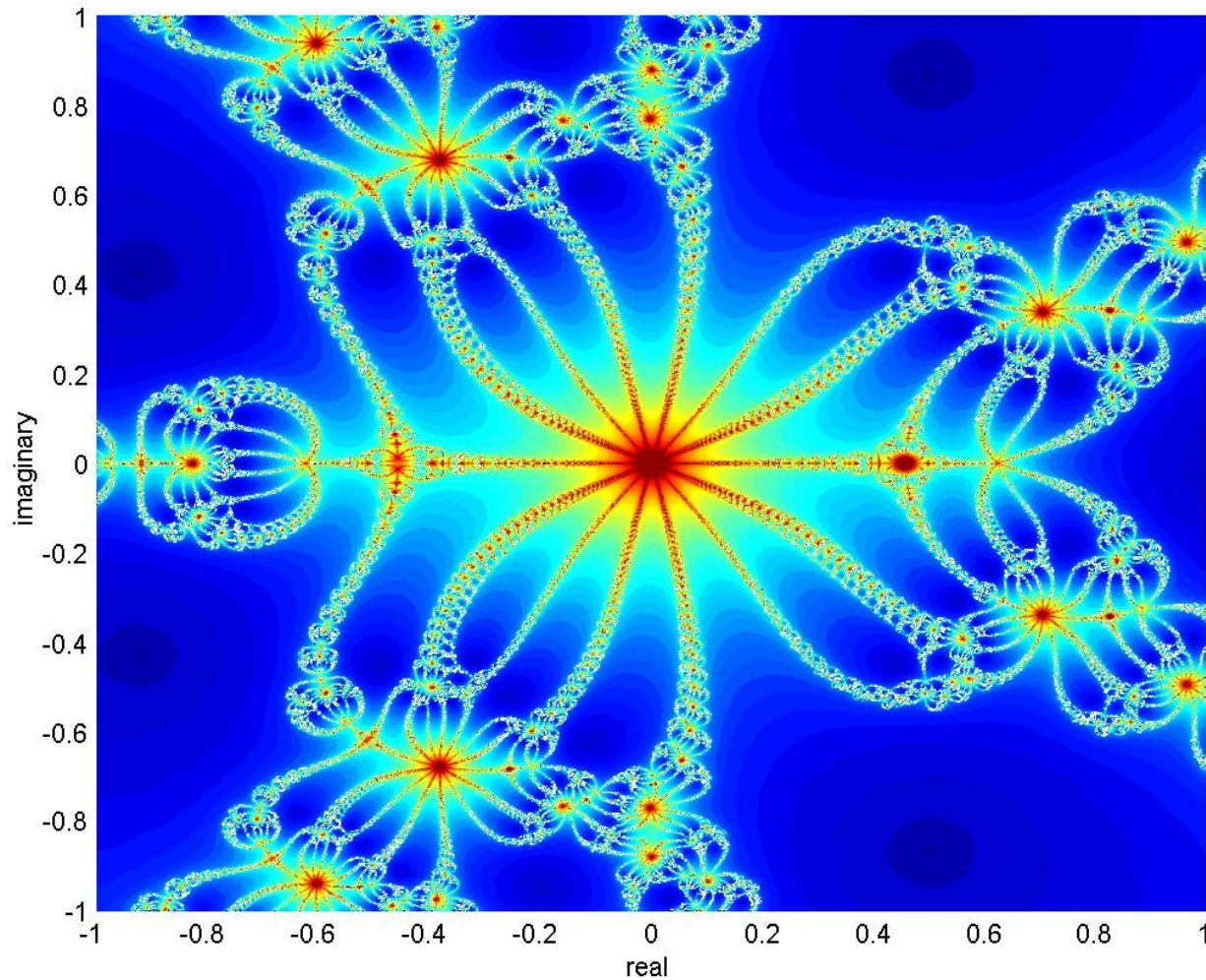
Hands-On LCG & EGEE Grids

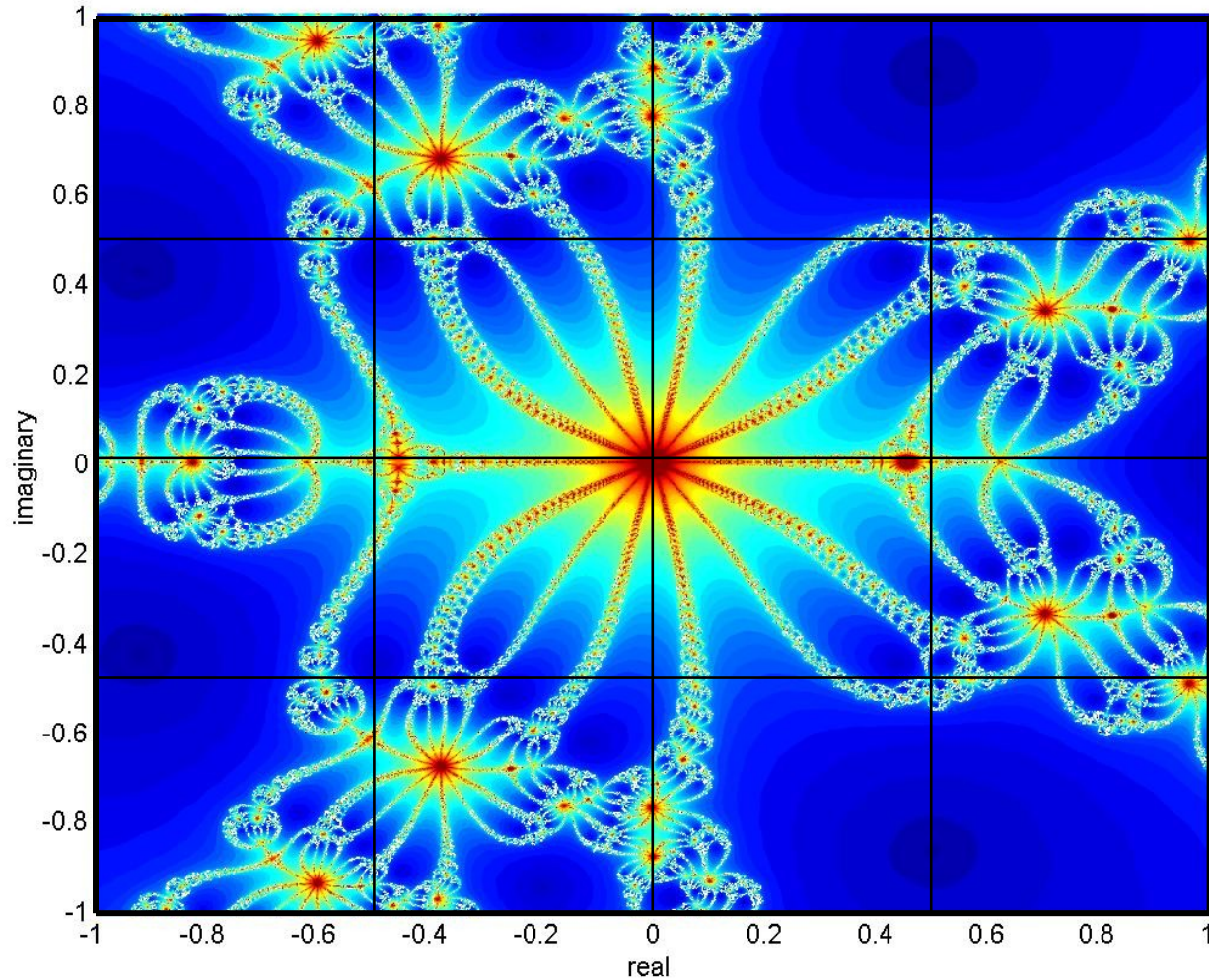
Fotis Georgatos <gef@grnet.gr>
Grid Technologies Trainer, GRNET

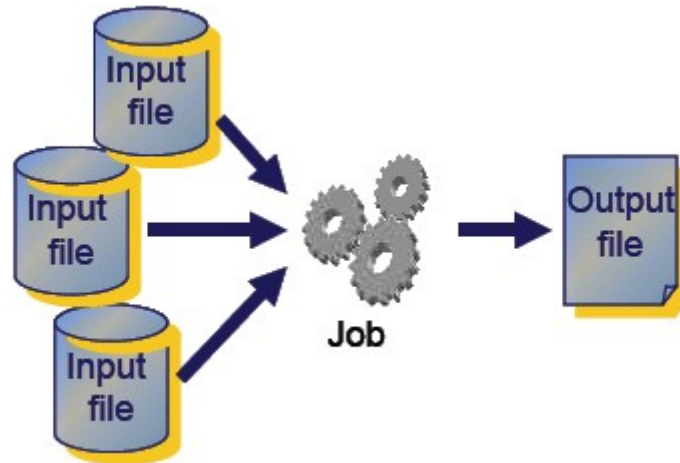
Chania, September 18th-19th, 2006

www.eu-egee.org



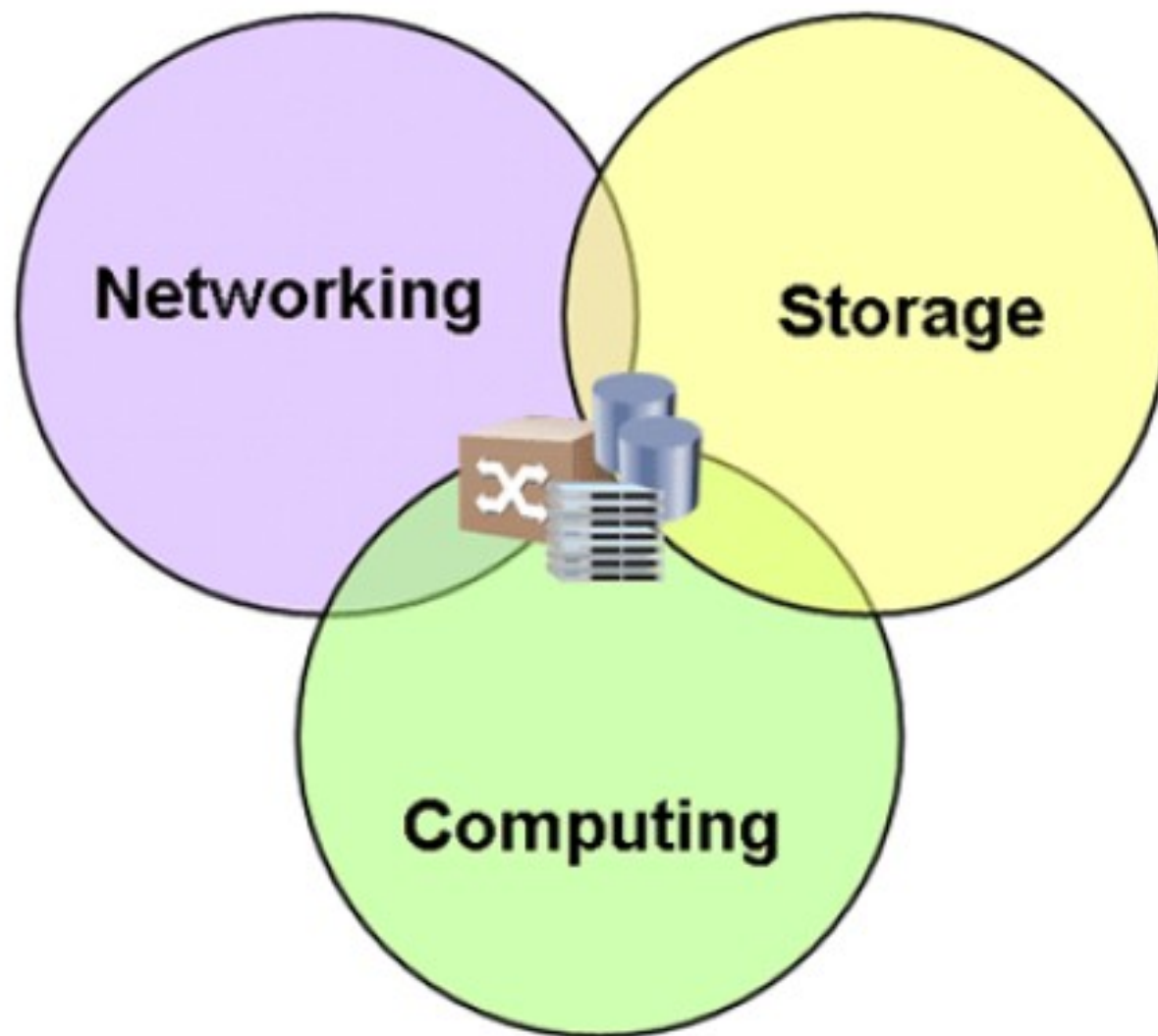


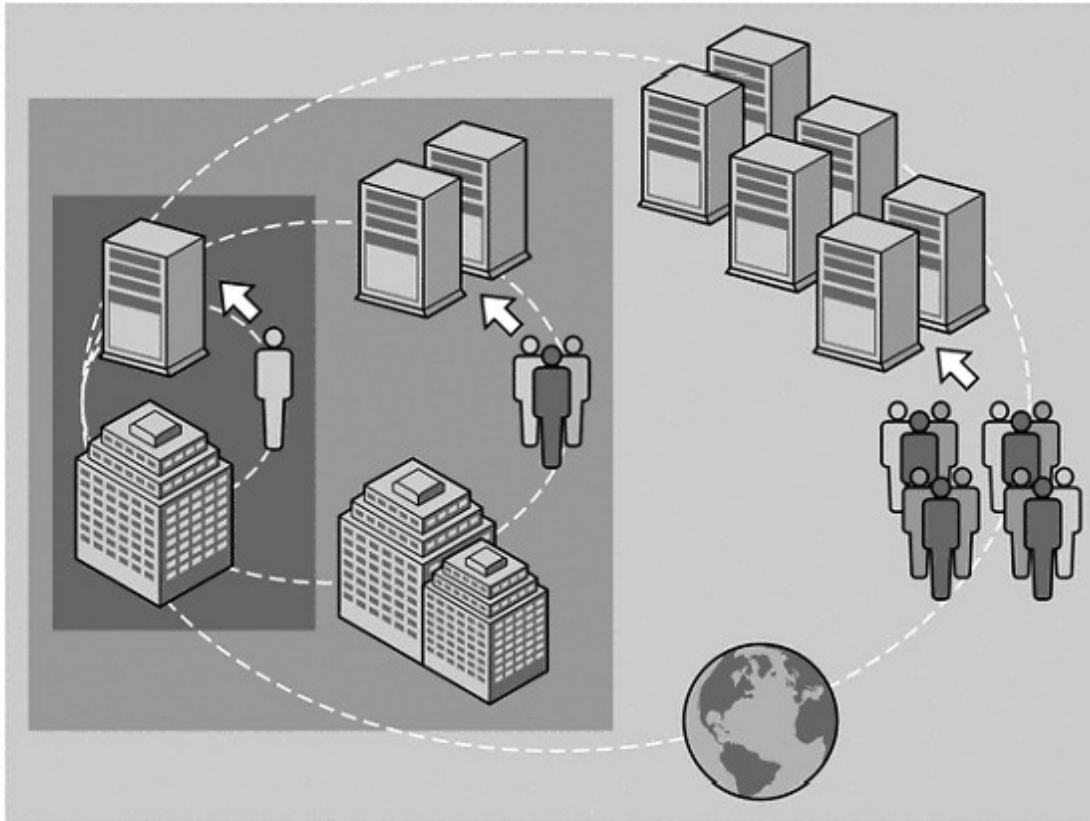




Ο χρήστης ζητάει:

- την εκτέλεση μιας εφαρμογής σε ένα μακρινό σύστημα,
- η οποία αναλύει κάποια δεδομένα εισόδου (τα οποία ενδεχομένως προϋπάρχουν στο Πλέγμα)
- παράγει κάποια δεδομένα εξόδου, τα οποία αποθηκεύονται επίσης στο Πλέγμα, πιθανώς σε κάποιο Storage Element ή στο UI.

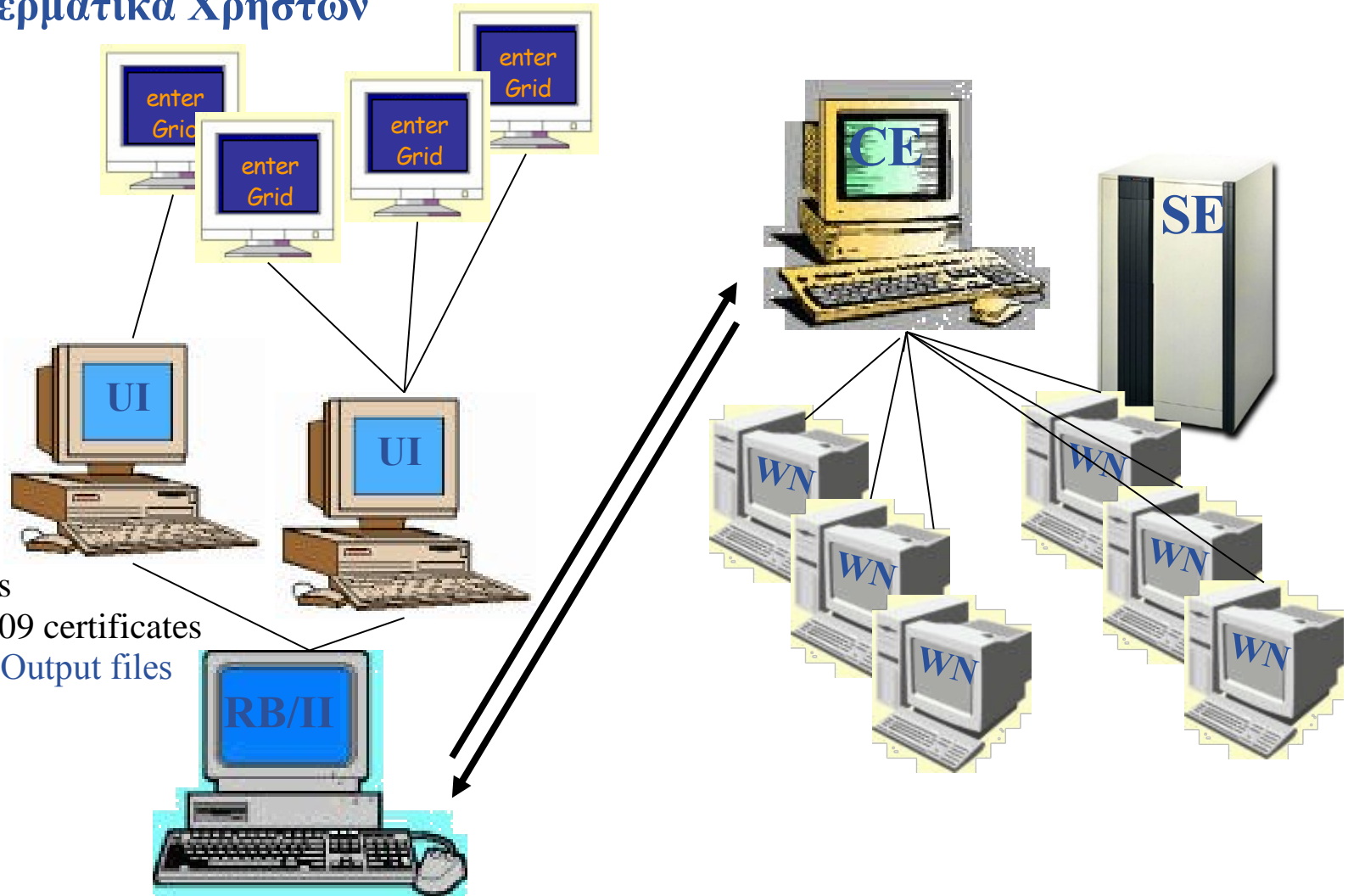




■ Cluster Grid ■ Campus Grid ■ Global Grid

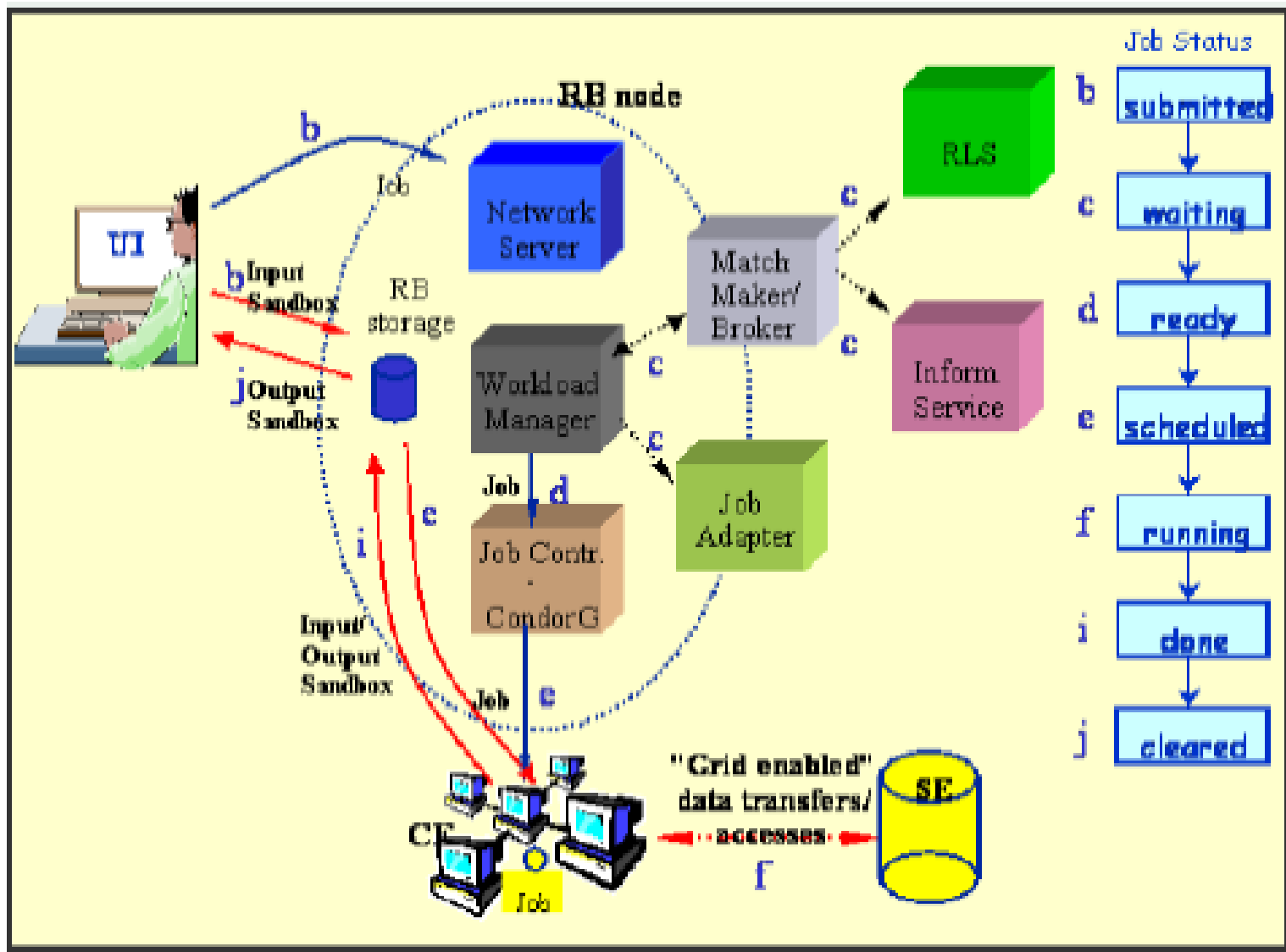
- Το Πλέγμα (LCG Grid) προσπαθεί να συντονίσει πόρους οι οποίοι ξεφεύγουν κατά πολύ από την τάξη μεγέθους ενός τοπικού ή campus-wide cluster.
- Οι αυξημένες δυνατότητες παρέχονται μεν, αλλά προφανώς με κάπως αυξημένη πολυπλοκότητα στο ενδιάμεσο λογισμικό.
- Απαιτείται, για να γίνεται η διαχείριση των εργασιών, οι εργασίες να εκτελούνται διά μέσω ειδικών LCG εντολών.

Τερματικά Χρηστών



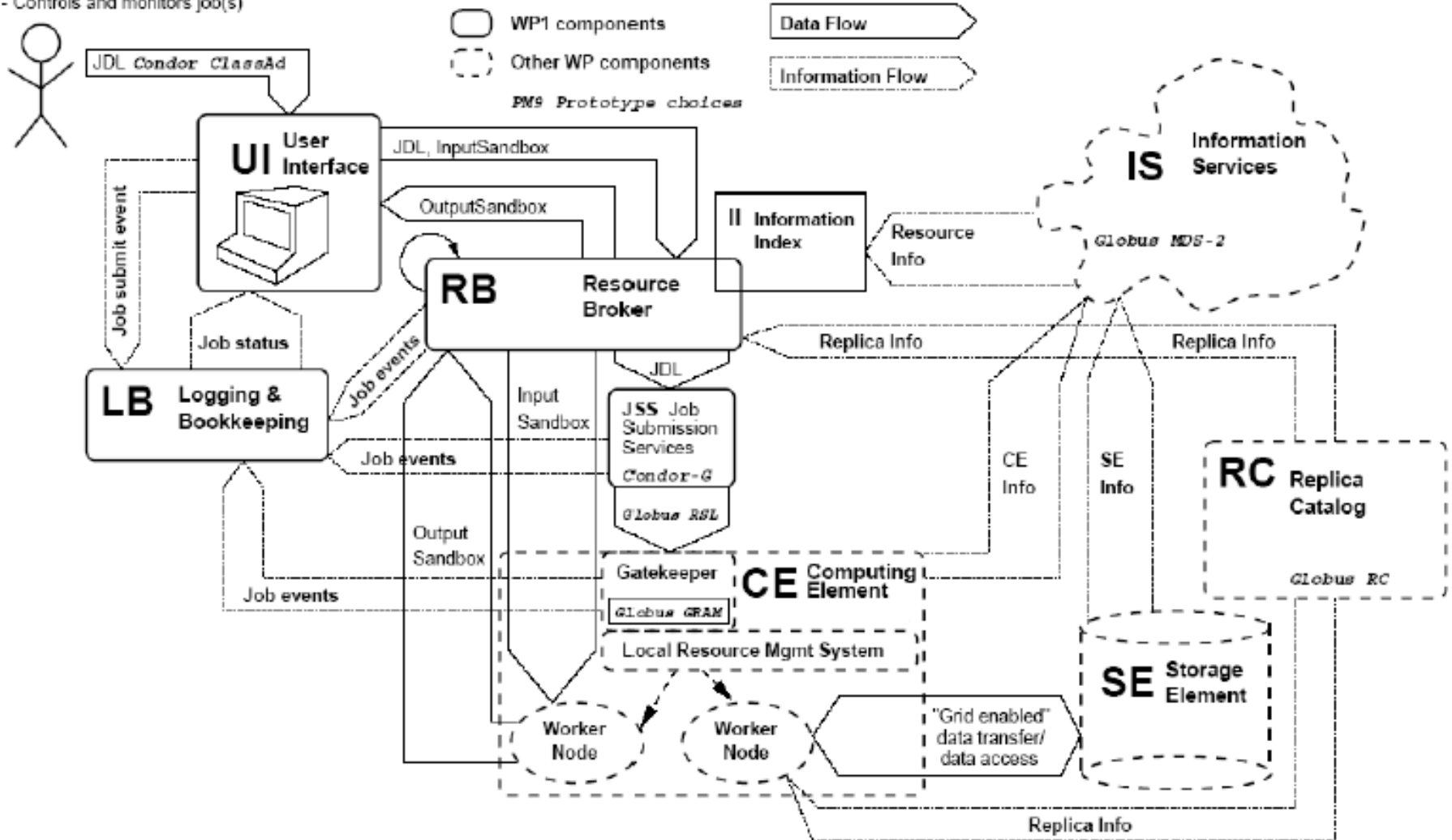
- JDL files
- PKI X.509 certificates
- Input & Output files

- **UI: User Interface - Διεπαφή Χρηστών**
 - Είναι το σύστημα από το οποίο οι χρήστες υποβάλουν διεργασίες
 - Μπορεί να είναι οποιοδήποτε Linux σύστημα, ακόμη και ένα laptop
- **RB: Resource Broker - Κατανεμητής Πόρων**
 - Οι διεργασίες υποβάλλονται εδώ από τον Χρήστη, πριν πάνε σε κάποιο CE
- **CE: Computing Element - Υπολογιστικό Στοιχείο**
 - Είναι ο κόμβος που παρέχει πρόσβαση σε μία ομάδα Worker Nodes
- **WN: Worker Node - Κόμβος Εξυπηρέτησης Διεργασιών**
 - Οι κόμβοι αυτοί παρέχουν υπολογιστικές υπηρεσίες στο Πλέγμα
- **SE: Storage Element - Αποθηκευτικό Στοιχείο**
 - Οι κόμβοι SE παρέχουν πρόσβαση σε συστήματα δίσκων και ταινιών
- **BDII: (...) - Βάση Δεδομένων Υπηρεσιών Καταλόγου**
 - Είναι ο κατανεμημένος μηχανισμός που παρέχει την πληροφόρηση «Που/Ποιος/Τι είναι το Πλέγμα» (Δουλεύει περίπου όπως το DNS)



End User

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



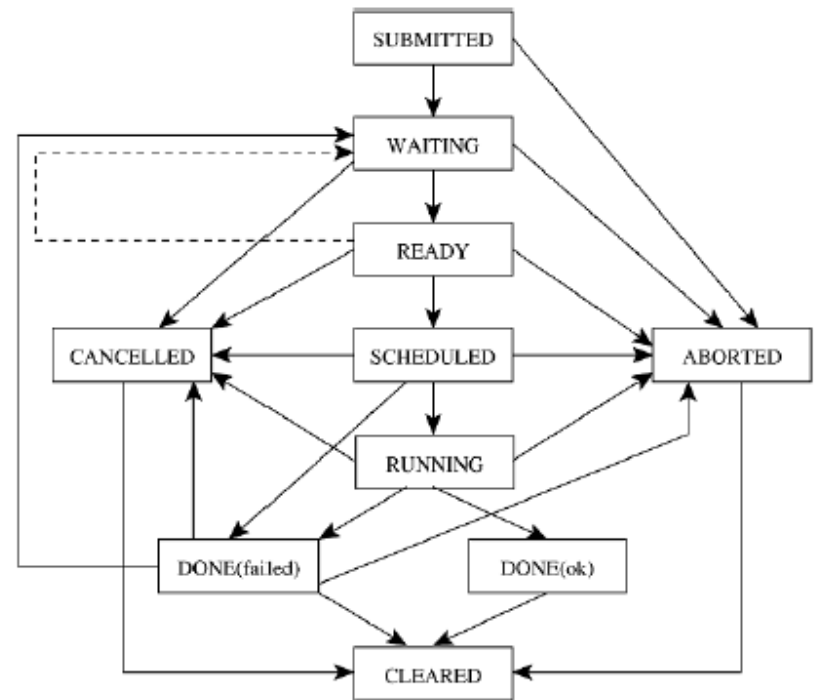
- **Χρήση ενός ssh client, πχ putty**
 - Το βρίσκετε εύκολα, πχ με ένα search engine όπως το google, ή κατευθείαν στην διεύθυνση <http://www.putty.nl/download.html>
- **Πρόσβαση σε ένα σύστημα τύπου User Interface:**
 - `ssh <myaccount>@ui01.isabella.grnet.gr`
 - `password: <mypassword>`
 - `ls -als .globus` (`mkdir .globus` εάν δεν υπάρχει ήδη)
- **Το ιδιωτικό κλειδί σας θα είναι στο αρχείο `userkey.pem`**
- **Το πιστοποιητικό σας θα είναι στο αρχείο `usercert.pem`**
- **Ενεργοποίηση πιστοποιητικού (συνήθως για 12 ώρες)**
 - `grid-proxy-init` και απαντάτε με `keyforcert`
 - `grid-proxy-info`, για να επιβεβαιώσετε ότι είναι πλέον ενεργό το πιστοποιητικό και πόσες ώρες απομένουν για την χρήση του

- **Job Description Language**

- Κατά το πρότυπο Condor classified advertisements
- Πληροφορία που αξιοποιείται για την βέλτιστη εκτέλεση της εργασίας
- Παράδειγμα ενός **hostname.jdl**:

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

- `edg-job-list-match --vo hgdemo testJob.jdl`
- `lcg-infosites --vo hgdemo ce`
- `lcg-infosites --vo hgdemo se`
- `lcg-infosites --vo see ce`
- `lcg-infosites --vo see se`
- `lcg-infosites --vo biomed ce`
- `lcg-infosites --vo biomed se`
- `lcg-infosites --vo atlas ce`
- `lcg-infosites --vo atlas se`
- `lcg-infosites --vo dteam ce`
- `lcg-infosites --vo dteam se`



- **edg-job-submit -o myjobids.txt --vo hgdemo testJob.jdl**
 - Στέλνει την διεργασία στον Resource Broker
 - Το RB εντοπίζει το πλέον κατάλληλο CE και στέλνει την διεργασία εκεί
 - Το CE αναθέτει σε ένα ελεύθερο WN την διεργασία
- **edg-job-status -i myjobids.txt**
 - submitted: Η διεργασία έχει υποβληθεί στο RB
 - waiting: Η διεργασία αναμένει επεξεργασία στο RB
 - ready: Η διεργασία έχει αντιστοιχηθεί ήδη σε κάποιο CE, αλλά δεν έχει ακόμα φτάσει στο LRMS του CE.
 - scheduled: Η διεργασία έχει μπει στην ουρά του LRMS στο CE
 - running: Η διεργασία τρέχει σε κάποιο Worker Node
 - done: Η διεργασία τερμάτισε και μένει η συλλογή από το RB
- **edg-job-cancel** εάν θέλουμε την ακύρωση μιας διεργασίας
- **edg-job-get-output -i myjobids.txt --dir .** παραλαβή εξόδου

- **Αναζήτηση Storage Element στα οποία έχουμε πρόσβαση:**
 - `lcg-infosites --vo hgdemo se`
- **Κατάθεση ενός αρχείου σε ένα Storage Element:**
 - `lcg-replica-manager cr -d <SE> -l lfn:<file> --vo=<myvo> file://`pwd`/<file>`
- **Ανάληψη ενός αρχείου από ένα Storage Element:**
 - `lcg-replica-manager cp -l lfn:<file> --vo=<myvo> file://`pwd`/<file>`
- **Δημιουργία αντίγραφου σε ένα άλλο SE:**
 - `lcg-replica-manager rep --vo hgdemo -d <SE> lfn:<file>`
- **Ερώτηση για τα αντίγραφα ενός αρχείου (με βάση το lfn):**
 - `lcg-replica-manager lr --vo <myvo> lfn:<file>`

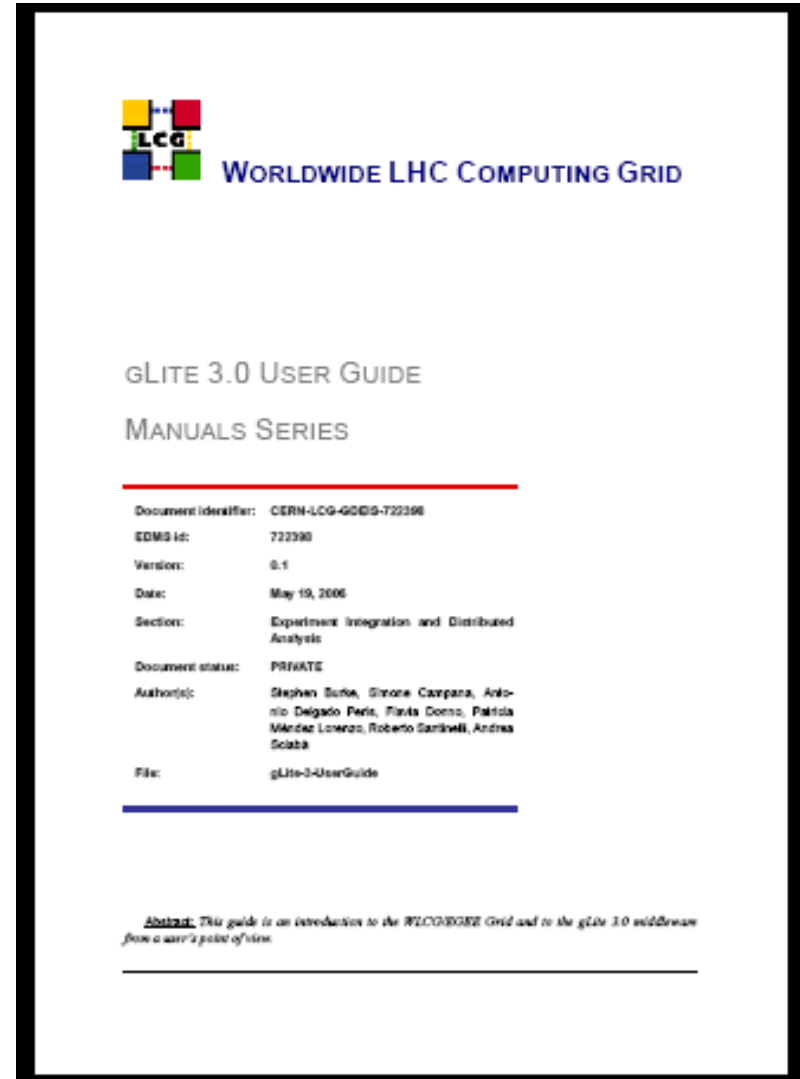
GLITE 3.0 USER GUIDE

- Οδηγίες χρήσης του gLite, με πολλές εντολές middleware, ένα .pdf με 166 σελίδες.
 - 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 (d2ol, for SARS, 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)

• 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)

