Grid Computing:
Running your Jobs around the World

LABORATÓRIO DE INSTRUMENTAÇÃO EM FÍSICA EXPERIMENTAL DE PARTÍCULAS

L I P

What is the GRID?

- GRID computing is a recent concept which takes distributing computing a step forward

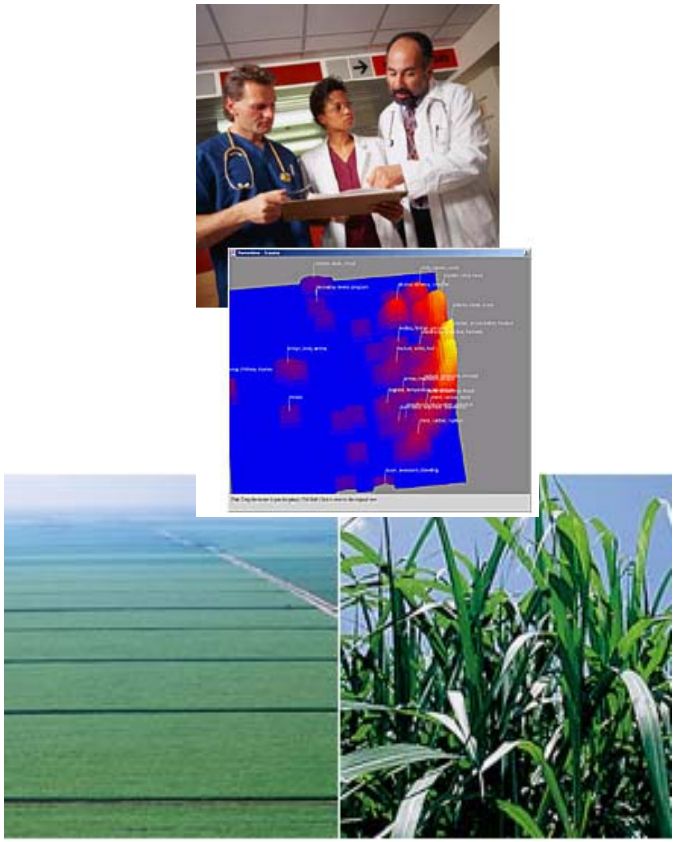
- The name GRID is chosen by analogy with the electric power grid:
  - Transparent: plug-in to obtain computing power without worrying where it comes from
  - Permanent and available everywhere

- 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

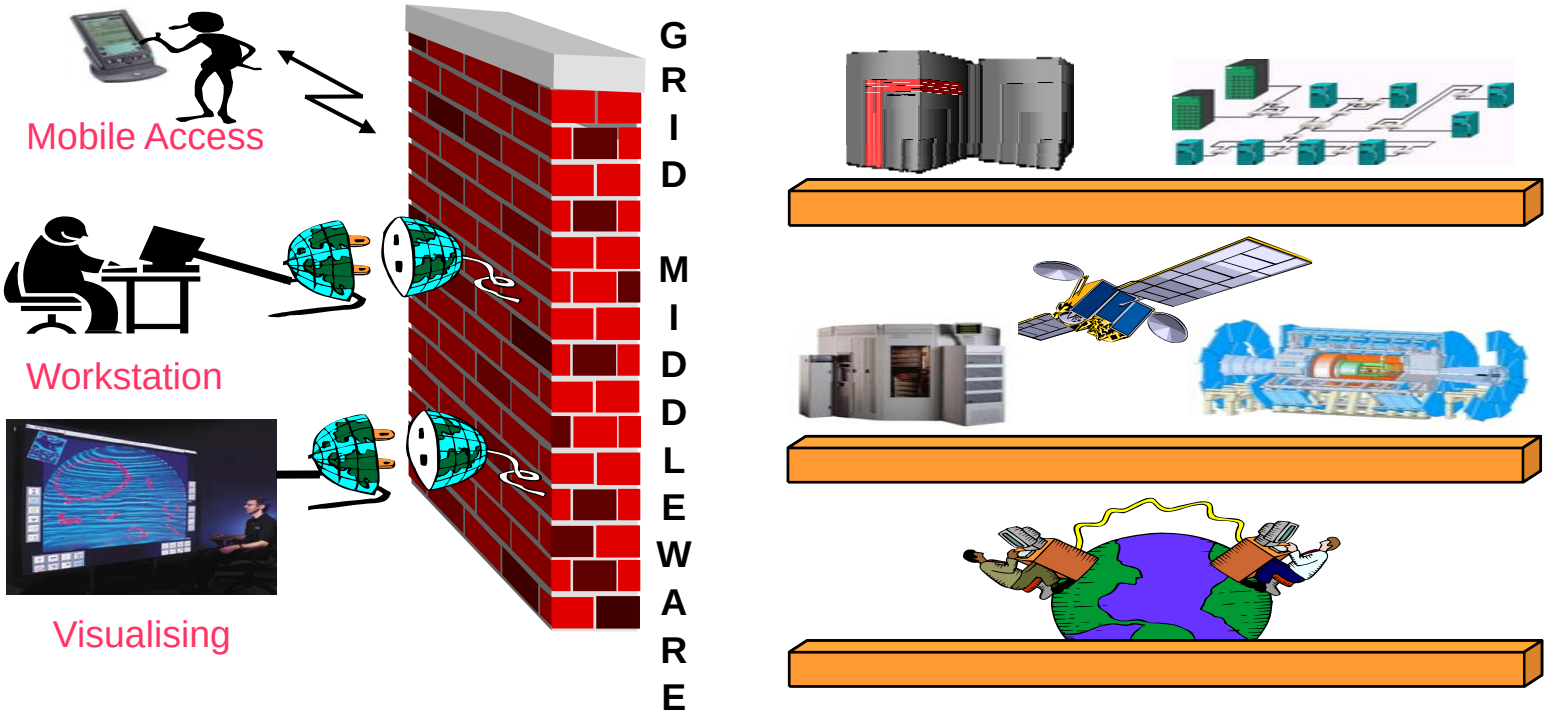
Motivation

- **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



The GRID Metaphor

- **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 vs Distributed Computing

- **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**

The LCG/EGEE Projects

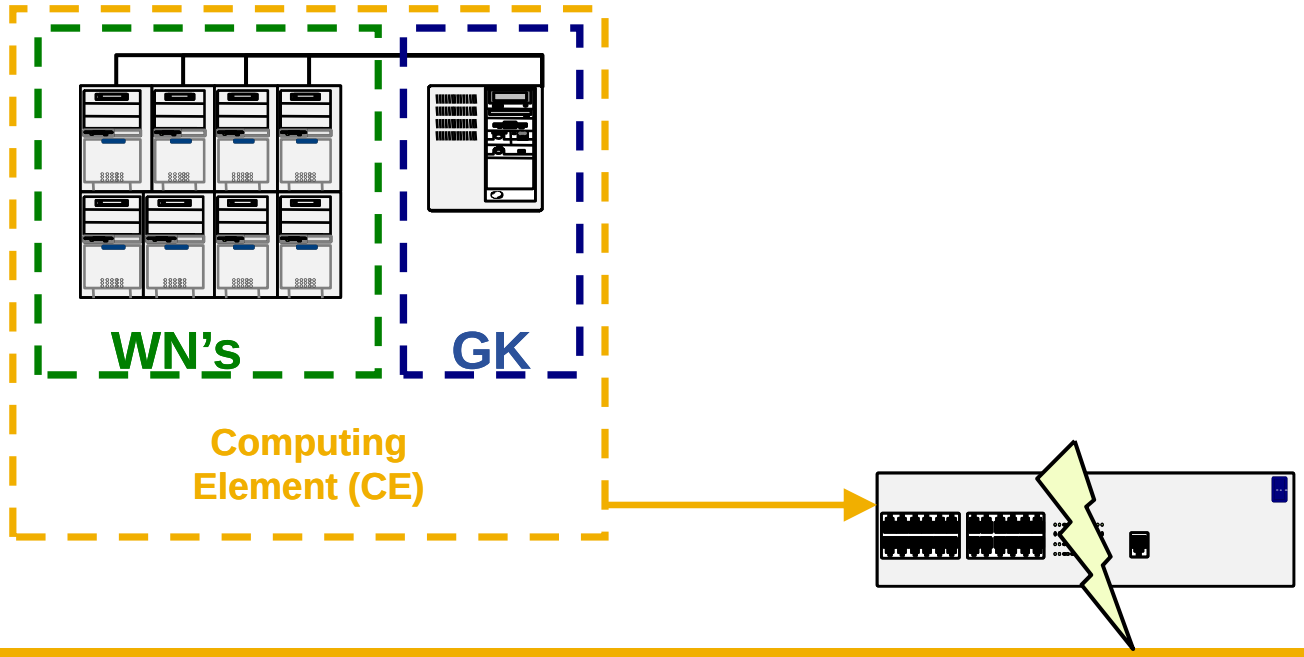
- **Among the several projects in place, LIP is involved in:**
  - **LHC Computing GRID (LCG)**
    - The biggest worldwide GRID infrastructure
    - Will be used in the data analysis produced by the LHC accelerator built at CERN, the european organization for nuclear research
  - **Enabling GRIDS for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

Grid Resources

- **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).

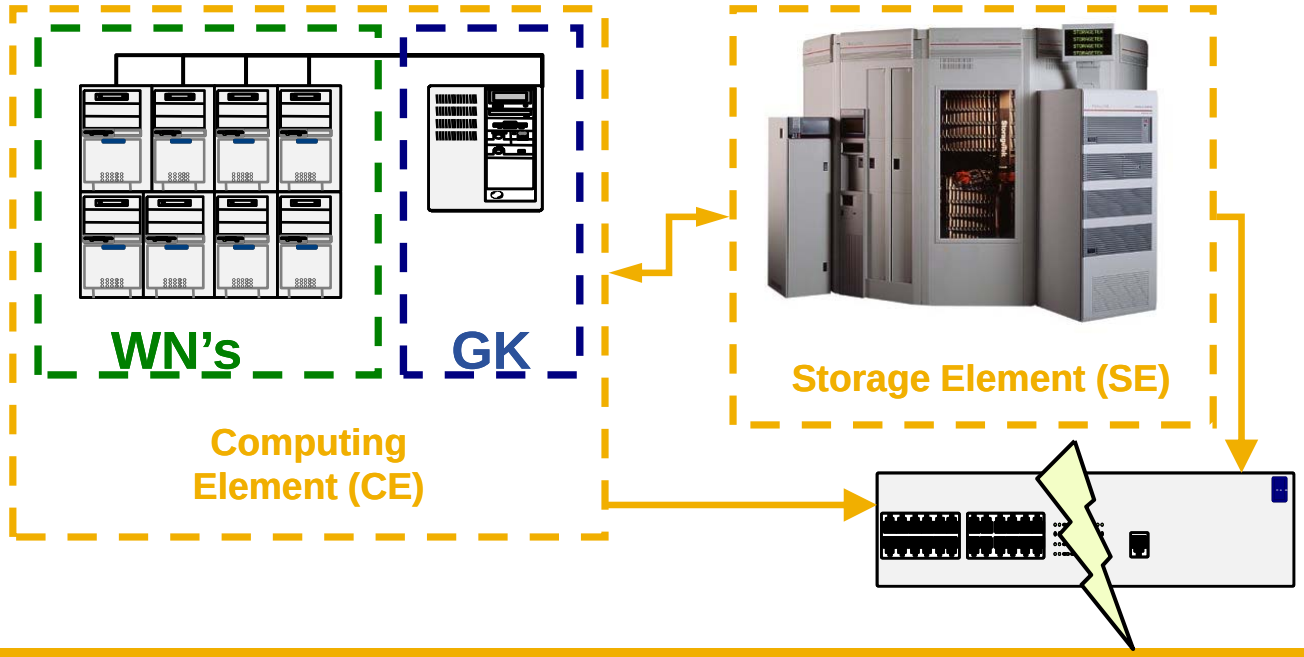
A GRID Site: The Computing Element

- **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.



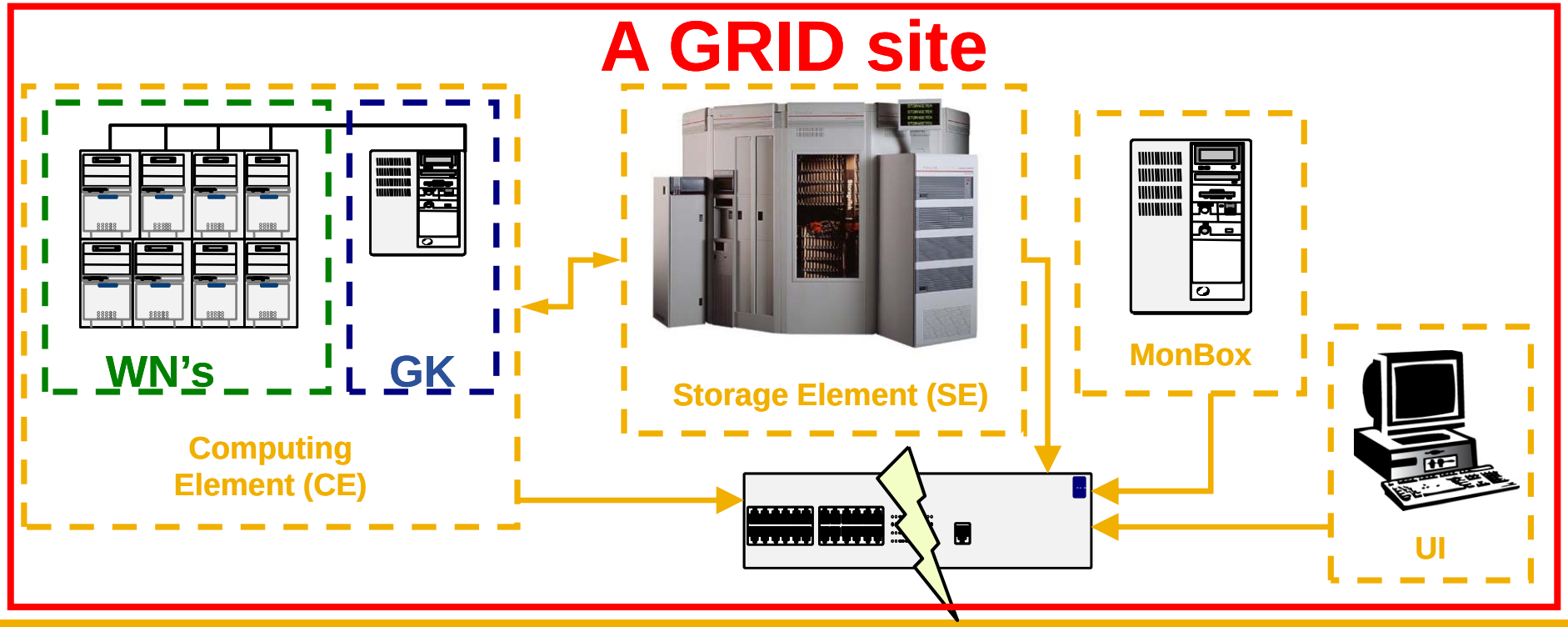
A GRID Site: The Storage Element

- **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.

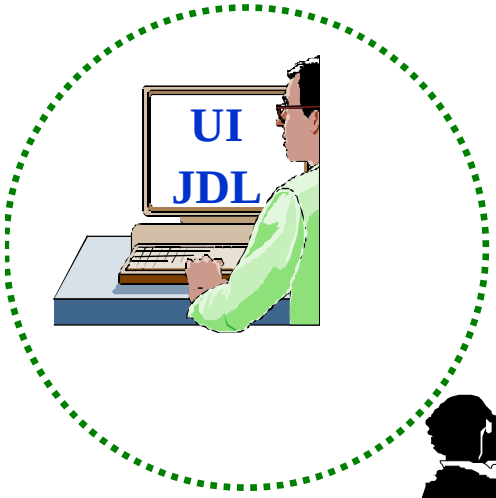


A GRID Site: The MONBOX & UI

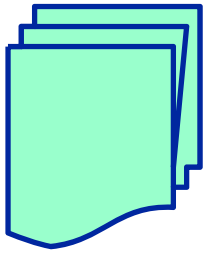
- The Monitoring Box (MonBox) service
  - Collects information given by sensors installed in the site machines.
- User Interfaces (UI)
  - Contain client middleware tools allowing the user to perform a large set of operations with grid resources (submission of jobs and storage and management of files).



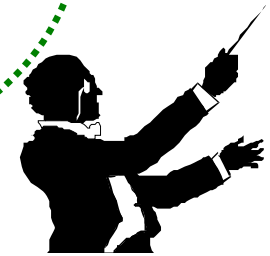
A job submission example



Replica Catalogues (RC)

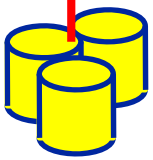


Information Service (IS)



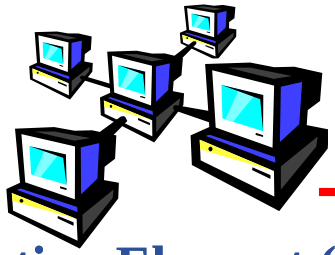
Resource Broker (RB)

Job Submission Service (JSS)



Storage Element (SE)

Logging & Book-keeping (LB)



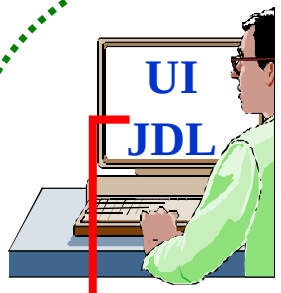
Computing Element (CE)

Job Status: Submitted

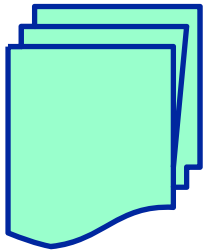
submitted

Replica Catalogues (RC)

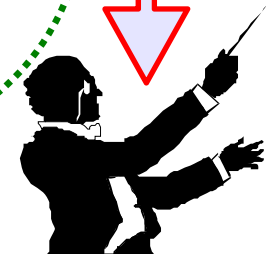
Information Service (IS)



Input Sandbox

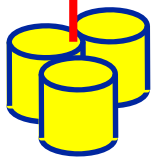


Job Submit Event



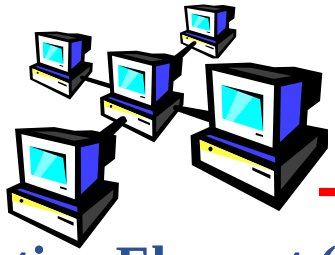
Resource Broker (RB)

Job Submission Service (JSS)



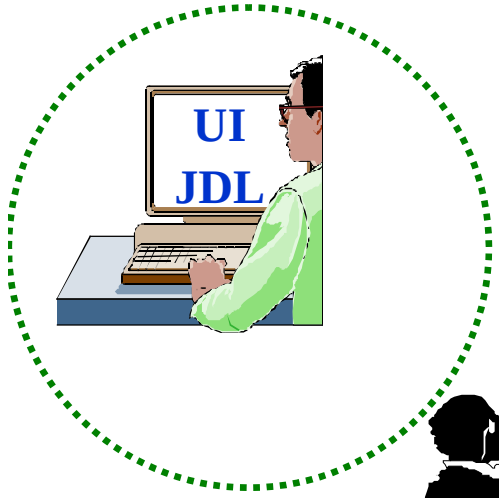
Storage Element (SE)

Logging & Book-keeping (LB)

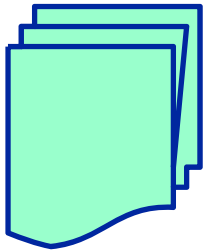


Computing Element (CE)

Job Status: Waiting



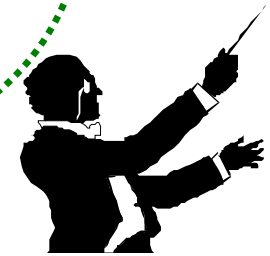
Replica Catalogues (RC)



Information Service (IS)

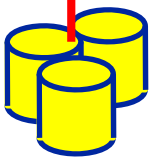
submitted

waiting



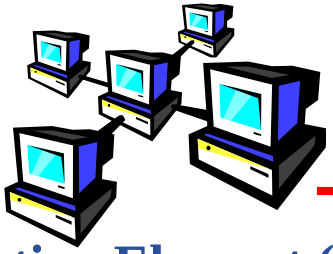
Resource Broker (RB)

Job Submission Service (JSS)



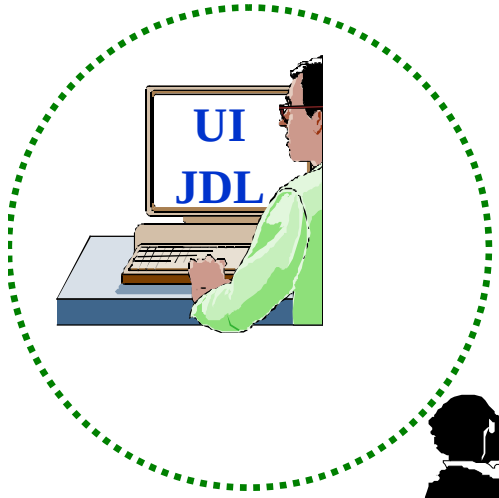
Storage Element (SE)

Logging & Book-keeping (LB)

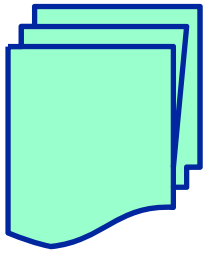


Computing Element (CE)

Job Status: Ready



Replica Catalogues (RC)

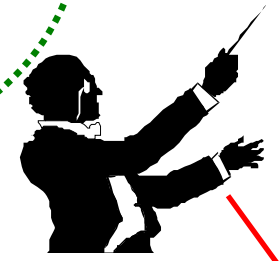


Information Service (IS)

submitted

waiting

ready

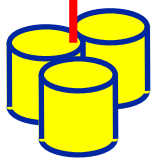


Resource Broker (RB)

Job Submission Service (JSS)

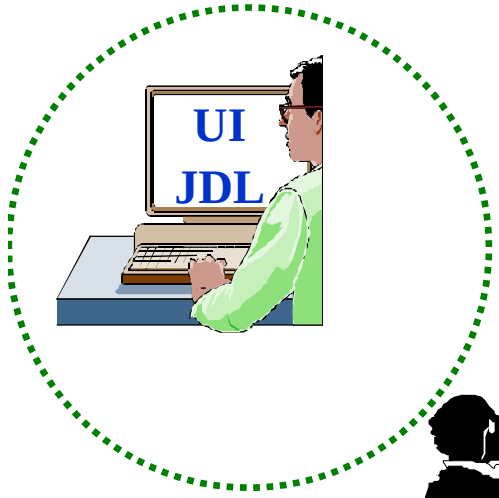
Logging & Book-keeping (LB)

Computing Element (CE)

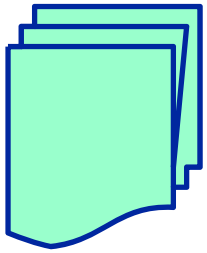


Storage Element (SE)

Job Status: Scheduled

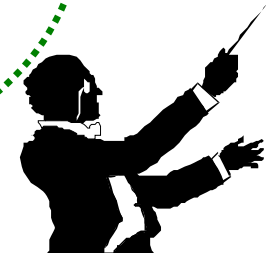


Replica Catalogues (RC)



Information Service (IS)

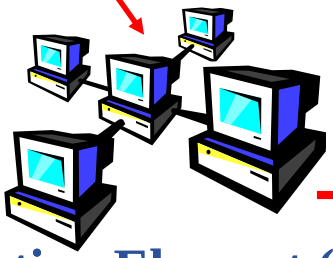
submitted
waiting
ready
scheduled



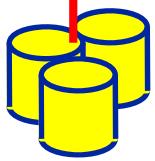
Resource Broker (RB)

Job Submission Service (JSS)

Logging & Book-keeping (LB)

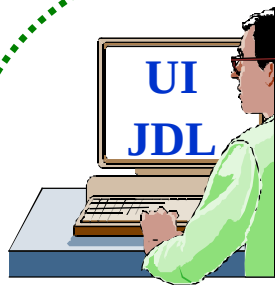


Computing Element (CE)

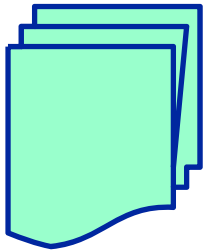


Storage Element (SE)

Job Status: Running



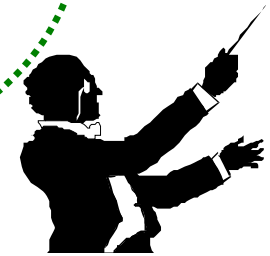
Replica Catalogues (RC)



Information Service (IS)

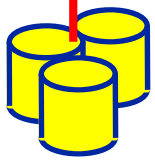
submitted
waiting
ready
scheduled
running

Input Sandbox



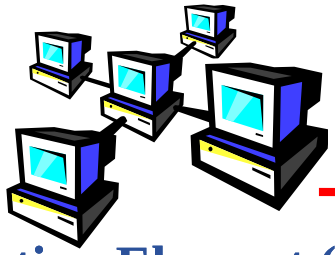
Resource Broker (RB)

Job Submission Service (JSS)



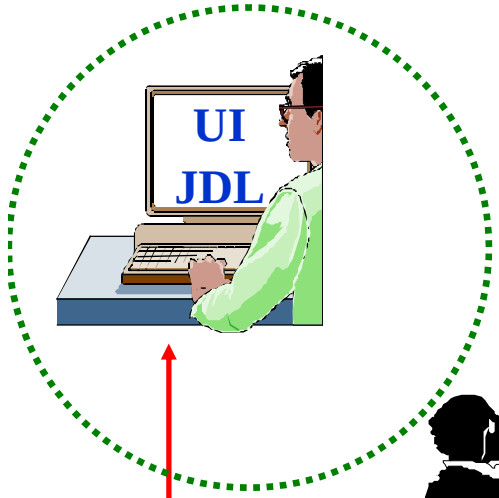
Storage Element (SE)

Logging & Book-keeping (LB)

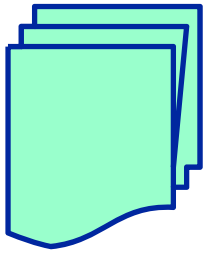


Computing Element (CE)

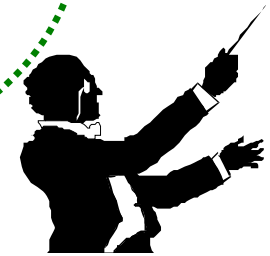
Job Status: Done



Replica Catalogues (RC)

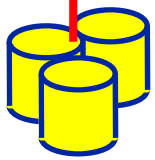


Information Service (IS)



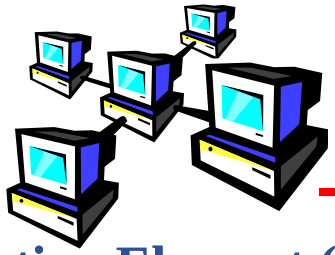
Resource Broker (RB)

Job Submission Service (JSS)



Storage Element (SE)

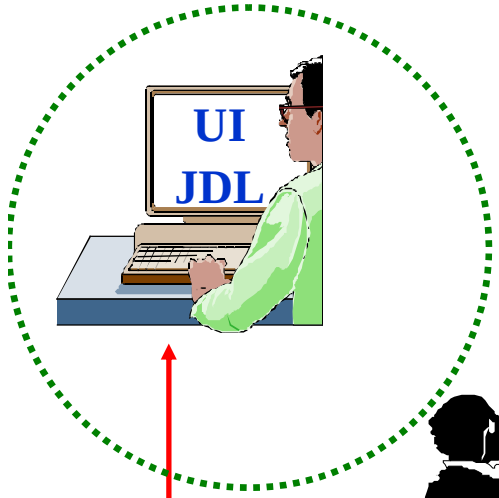
Logging & Book-keeping (LB)



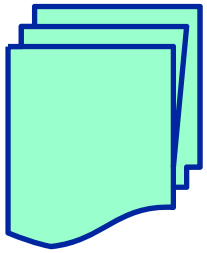
Computing Element (CE)

- submitted
- waiting
- ready
- scheduled
- running
- done

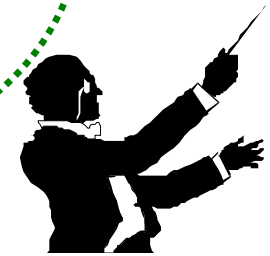
Job Status: OutputReady



Replica Catalogues (RC)

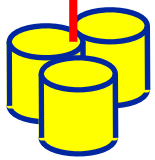


Information Service (IS)



Resource Broker (RB)

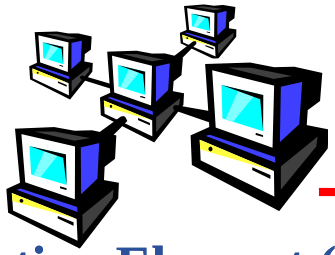
Job Submission Service (JSS)



Storage Element (SE)

Logging & Book-keeping (LB)

Output Sandbox



Computing Element (CE)

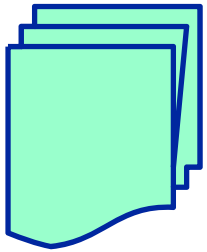
submitted
waiting
ready
scheduled
running
done
outputready

Job Status: Cleared

Example1; Example2

Output Sandbox

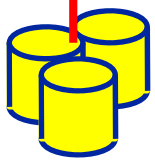
Replica Catalogues (RC)



Information Service (IS)

Resource Broker (RB)

Job Submission Service (JSS)



Storage Element (SE)

Computing Element (CE)

Logging & Book-keeping (LB)

- submitted
- waiting
- ready
- scheduled
- running
- done
- outputready
- cleared