gLite Training @ htw.

Yves Kemp, Hermann Hessling

Berlin, Nov.&Dec. 2010

Best use Linux desktop!

*Find the slides in **grid-lab00[2,3].desy.de:/home/common/**
http://www.desy.de/~kemp/htw_course.pdf*

Today's Task: The Physics Analysis

> ## Background
  - Velocity of particles in an ideal gas obeys a certain law:
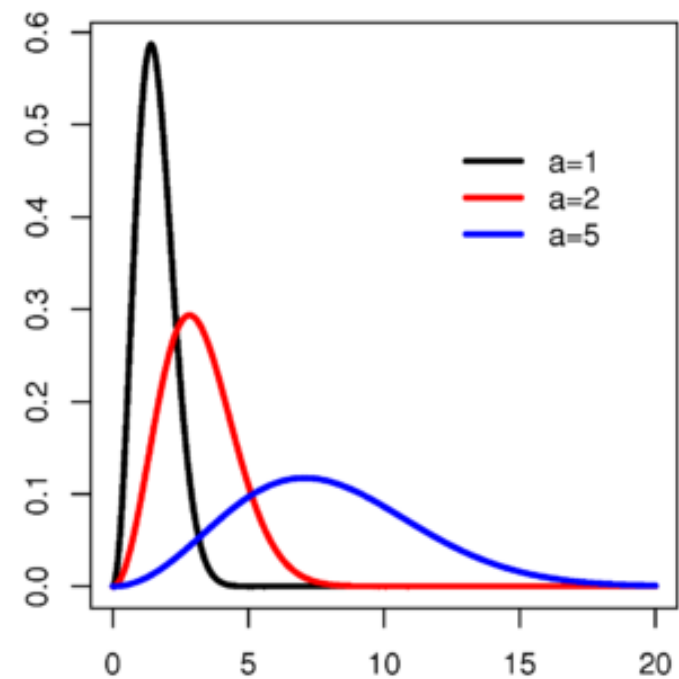 - Maxwell-Boltzmann distribution

> ## Question
  - If particle N hat velocity larger than mean, is it more probable for particle (N+1) to have a velocity below mean or above mean?

> ## Use a Monte-Carlo generator
  - Instead of performing experiment
  - Use Grid jobs for computation
  - Download files and do analysis
  - On your PC, with system tools

$$f(v) = 4\pi \left(\frac{m}{2\pi kT}\right)^{3/2} v^2 \exp\left[\frac{-mv^2}{2kT}\right]$$

Workflow

> **Monte-Carlo executable already exists**
  - You find it in $HOME/job/MaxBol
  - Has a bug: Sometimes goes into infinite loop: Needs watchdog
> **One Grid job**
  - Handles autonomously all (foreseen…) problems
  - $HOME/job/script.sh
  - Is / should be somewhat robust against failures
> **Send many jobs to the Grid**
  - Has been done for you over the weekend
  - We can send one more job, your task will be to retrieve the output
> **Data is stored on the Grid**
  - You have to retrieve the data files to your / course computer
  - And do the analysis
> **We will try to do the analysis with normal utilites: grep, sed, awk, …**
  - … and in the meantime learn something about the Grid, Physics and Statistics:-)

Topics

> **Preliminaries**
  - Networking

> **Security**
  - Virtual Organizations
  - Authentication
  - Certificates
  - Grid Proxy

> **Data Management**
  - Overview and LFC
  - LFN and PFN
  - Copy, Register, Replicate, Delete

- **gLite**
  – Job Workflow
  – Workload Management System
  – Job Description Language

- **Job Handling**
  - Simple Jobs
  - A "real life" ILC example

- **General recommendations**

Topics

> **Preliminaries**
>   - Networking
>
> **Security**
>   - Virtual Organizations
>   - Authentication
>   - Certificates
>   - Grid Proxy
>
> **Data Management**
>   - Overview and LFC
>   - LFN and PFN
>   - Copy, Register, Replicate, Delete

- **gLite**
  - Job Workflow
  - Workload Management System
  - Job Description Language
- **Job Handling**
  - Simple Jobs
  - A "real life" ILC example
- **General recommendations**

Preliminaries

> **Login Information:**
- User Interface (UI) available:
  - grid-lab00[2,3].desy.de (accessible only via htw network)
  - Please use your own account information to log in via
- SSH:
  - `ssh <username>@grid-lab00[2,3].desy.de`
  - Check with your neighbour and alternate machines.

> **Only two computers, please do not waste resources!**

> **Why no local computer?**
- Installation experience at DESY
- Easiest with special OS:
- ScientificLinux (Recompilation of RedHat Enterprise Server)
  - Not available at htw
- In theory, you can use any computer with any OS (well, Linux…), even your laptop

- > **Preliminaries**
  - Network
- > **Security**
  - Virtual Organisations
  - Authentication
  - Certificates
  - Grid Proxy
- > **Data Management**
  - Overview and LFC
  - LFN and PFN
  - Copy, Register, Replicate, Delete

- **gLite**
  - Job Workflow
  - Workload Management System
  - Job Description Language
- **Job Handling**
  - Simple Jobs
  - A "real life" ILC example
- **General recommendations**

Who am I? What can I co?

> **For access to resources, two questions are of importance:**

> **"Who am I?" (authentication)**
  - The certificate answers this
  - The Certification Authority confirms this

> **"What can I do?" (authorization)**
  - You can do what your Virtual Organization allows you to do.

> **Comparision with "real life"**

Authentication:
Passport

Authorization:
Visa

  - Only the combination of Passport & Visa enables you to travel

Security: Authentication

> *Authentication and Authorization while using resources:*
>> ■ **VO does authorization ("What am I allowed to do?")**
>> ■ **Authentication through certificate and proxy ("Who am I?")**
> *Technical implementation: Public key, digitally signed by the Certification Authority*
>> ■ **The CA (e.g. GridKA @ FZK) makes correspondence certificate-identity by requiring the ID card of the user**
> *A certificate (or proxy) is required for all actions on the grid: job submission, data transfer…*
> *Your certificate might seem anonymous, but:*
>> ■ **Correspondence certificate-user known (one of the requirements from DECH management)**
>> ■ **Please keep your login data private!**

Security: Authorization

> ## "What am I allowed to do?"
  - Manged by VOMS: Virtual Organization Membership Service

> ## What is a VO? Virtual Organization?
  - General: A collection of people working on the same subject.
  - In HEP: An experiment or collaboration

> ## VOMS: Enables a very granular authorization scheme
  - Subgroups and different roles (e.g. /atlas/de, /zeus/Role=lcgadmin)
  - Steering at time of proxy creation

> ## Comparison
  - Real life: Passport -> Certificate
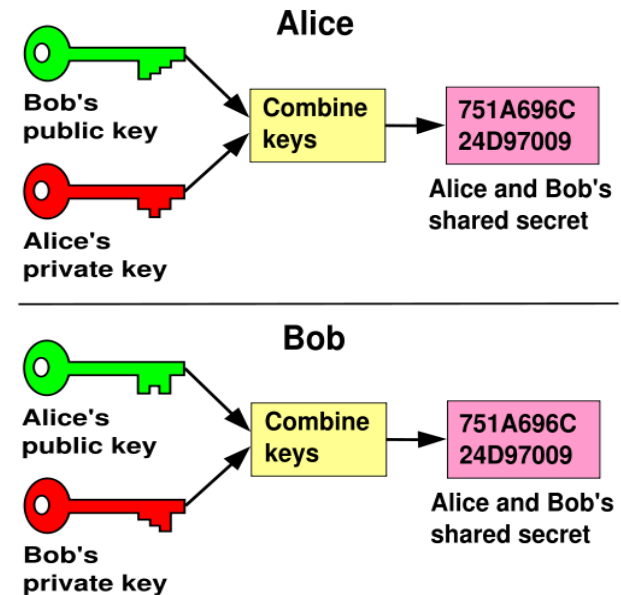 - Real life: Country -> VO
  - Real life: Duties/Rights -> VOMS

DESY

> **Certificates located on UI: `$HOME/.globus/`**
>   - `ls -la ~/.globus`
> **Each user has two files:**
>   - `usercert.pem`
>   - Public key, everyone can read it.
>   - File permissions: --rw--r--r or 644
>   - `userkey.pem`
>   - Private key, only owner should read it.
>   - File permissions: --rw------ or 600

> **Certificate are digital identity:**
>   - Working on the Grid
>   - Signing Emails
>   - Authentication through browser
>   - Different formats might be needed, all equivalent

Hands on

> **Certificate**
> ▪ Valid for (usually) one year, valuable, must be protected w. password
> **The authentication in the Grid is done by a short-living proxy created with your certificate:**
> ▪ Make a new grid-proxy:
>   ▪ `voms-proxy-init [-voms dech] [-valid hh:mm] [-vomslife hh:mm]`
>   ▪ Enter your GRID passphrase.
>   ▪ You can specify the VO during proxy creation
>   ▪ For long jobs you may need to change the proxy lifetime using e.g. the '-valid 24:00' option for a lifetime of one day.
>   ▪ The lifetime of the VOMS extension can be adapted separately
> ▪ Check the status of your proxy:
>   ▪ `voms-proxy-info [-all]`
>   ▪ `(try different proxies with/wo VOMS)`
> ▪ Delete your proxy:
>   ▪ `voms-proxy-destroy`
> ▪ What happens when you do?
>   ▪ `voms-proxy-init -voms dech:/dech/Role=lcgadmin`
>   ▪ `voms-proxy-init -voms atlas`

Hands on

Security - Authentication

> **Not every user is allowed to do everything on the grid.**
> **Therefore a grid-proxy can be created with different roles.**
> **An important role is the LCG administrator. He/she is allowed to install software on the CEs**
>   - voms-proxy-init -voms dech:/dech/Role=lcgadmin
> **The different privileges are connected with a different account mapping on the local resource**
>   - Different account mappings enable different access to local resources at site level (file permissions, batch queue share,…)
> **The user mapping on a CE/WN can be explored by the command** `whoami`
>   - Without special role: dechusrXXX
>   - As LCG administrator: dechsgmXXX
>   - (examples from DESY CE, other sites might be different)
>   - Try e.g. globus-job-run grid-ce5.desy.de /usr/bin/id

Topics

- **Preliminaries**
  - Networking
- **Security**
  - Virtual Organizations
  - Authentification
  - Certificates
  - Grid Proxy
- **Data Management**
  - Overview and LFC
  - LFN and PFN
  - Copy, Register, Replicate, Delete

- **gLite**
  - Job Workflow
  - Workload Management System
  - Job Description Language
- **Job Handling**
  - Simple Jobs
  - A "real life" ILC example
- **General recommendations**

Data Management - Overview

> **A file on the grid:**

- can be written, copied or deleted but not changed
  - Tribute to underlying Storage Elements that might have the file on tape
- has a global unique identifier (GUID)
  - e.g. guid:62bb0006-aaa2-4792-ae73-e282aef3bed1
- can have several replicas at different sites, each having a different physical file name (PFN/SFN)
  - e.g. sfn://se-fzk.gridka.de/storage/dech/generated/2006-09-11/filedeab4811-0e1f-4439-bffa-4efc46f40b82
- can be given several logical file names (LFN) by the user
  - e.g. lfn:/grid/dech/tarballs/dag.tar

- This information is stored in the LCG file catalogue (LFC).

Data Management - LFC

> **Where is the LCG File Catalog located?**
- `lcg-infosites --vo dech lfc`
  - In our case: `rb.scai.fraunhofer.de`
- `export LFC_HOST=rb.scai.fraunhofer.de`
  - `$LFC_HOST` may already be set on UIs (not on WNs)
  - Usually set this yourself
- (Set variable anyhow, even if lcg-infosites does not return result…)

> Data management: creating directories and show contents
- List the existing files and directories, e.g.:
  - `lfc-ls -l /grid/dech`
  - `-l` gives more information like the file size in the catalogue
- Create your own directory in the LFC (if not yet there):
  - `lfc-mkdir /grid/dech/htw/$USER`
  - (Please use the subdirectory …htw…)
- Already directory and files in there, do not delete them, we will need them later

> Hints:
- **More on lcg-infosites commands late**
- **All catalogue operations must have LFC_HOST set**
- **Spelling of fraunhofer:-)**

Hands on

Data Management - Storage

> **"How to store a file on the grid?" rises two questions:**
> - Where should the file be stored physically?
> - Which name should be assigned to the file?

> **Get a list of available storage elements:**
> - `lcg-infosites --vo dech se`

> **What is the meaning of the different columns?**
> **Some SEs might not work, for different reasons…**
> **Make your own experience:-)))**

Hands on

> **Most file handling commands begin with lcg-**
> **Copy a file to a SE:**
> - `(example of input file: /bin/bash :-))`
> - Choose an SE
> - `lcg-cr --vo dech -v -d dcache-se-desy.desy.de  \`
>   `    file:/bin/bash \`
>   `  -l lfn:/grid/dech/htw/$USER/testfile0`
> - → The `-v` option displays verbose output like used LFC, GUID, etc.

> **Where is this file now?**
> - `lcg-lr --vo dech  lfn:/grid/dech/htw/$USER/ \`
>   `testfile0`
>   - You get the storage file name of the file: `srm://<host>/<path`

> Hints:
> - "\" means long command (place this in one line)
> - Is LFC_HOST set?

Data Management - Replicates

> **Replicate the file to another SE:**
>   - `lcg-rep -v --vo dech -d globe-door.ifh.de \`
>   `lfn:/grid/dech/htw/$USER/testfile0`

> **List replicas and inspect the catalog:**
>   - Do all the files have the same size?

> **Copy the file back in your working directory:**
>   - `lcg-cp -v --vo dech \`
>   `lfn:/grid/dech/htw/$USER/testfile0 \`
>   `file://$HOME/junk`
> **Exercise: Download all the files that are registered in your /grid/…/ htwNNN directory**
>   - A small bash-for-loop might come in handy

Hands on

Constructing some commands

> **Look into the downloaded files: They contain the MC generated files from jobs (have been sent to the Grid earlier)**
> **Opening tar archive (compressed)**
> **Looking into tar archive**
> **Open all the tar archives you downloaded earlier**
> **Find all output files and put all "Randomly generated in one big file"**

Data Management - Deleting

> **Cleaning up:**

- Remove the file from one specific SE:
  - ```
    lcg-del --vo dech -s dcache-se-desy.desy.de \
        lfn:/grid/dech/htw/$USER/testfile0
    ```
  - Check the replicas and the catalog

- Remove the file from all SEs:
  - ```
    lcg-del --vo dech -a \   lfn:/grid/dech/htw/$USER/
    test.txt
    ```
  - Check the replicas and the catalog

- Remove your directory in the LFC
  - ```
    lfc-rm -r /grid/dech/htw/$USER
    ```

> **Try to remove a file in the lfc which was registered by your neighbour (with his/her consent).**

> **Do not delete the files that were originally there**

Hands on

Data Management - Commands

> **LFC commands:**
  - `lfc-ls`
  - `lfc-mkdir`
  - `lfc-rm`
  - `lfc-ln`
  - `lfc-rename`

File commands:
  `lcg-cr`
  `lcg-lr`
  `lcg-rep`
  `lcg-cp`
  `lcg-del`

> **LFC commands for meta data:**
  - `lfc-setcomment <file> "comment"`
  - `lfc-ls --comment <file>`
  - `lfc-delcomment <file>`

Topics

> **Preliminaries**
- Networking

> **Security**
- Virtual Organizations
- Authentification
- Certificates
- Grid Proxy

> **Data Management**
- Overview and LFC
- LFN and PFN
- Copy, Register, Replicate, Delete

- **gLite**
  - Job Workflow
  - Workload Management System
  - Job Description Language

- **Job Handling**
  - Simple Jobs
  - A "real life" ILC example

- **General recommendations**

LFC
Catalog

Information
Service

Resource
Broker (WMS)

Author.
&Authen.

Logging &
Book-keeping

Job Submission
Service

Storage
Element

Computing
Element

DESY

gLite - Job Workflow

Input "sandbox"

Output "sandbox"

voms-proxy-init

Author. &Authen.

Resource Broker

LFC Catalog

DataSets info

Information Service

SE & CE info

Expanded JDL

Input "sandbox" + Broker Info

Output "sandbox"

Publish

Job Submit Event

Job Query

Job Status

Globus RSL

Job Submission Service

Storage Element

Job Status

Logging & Book-keeping

Job Status

Computing Element

DESY
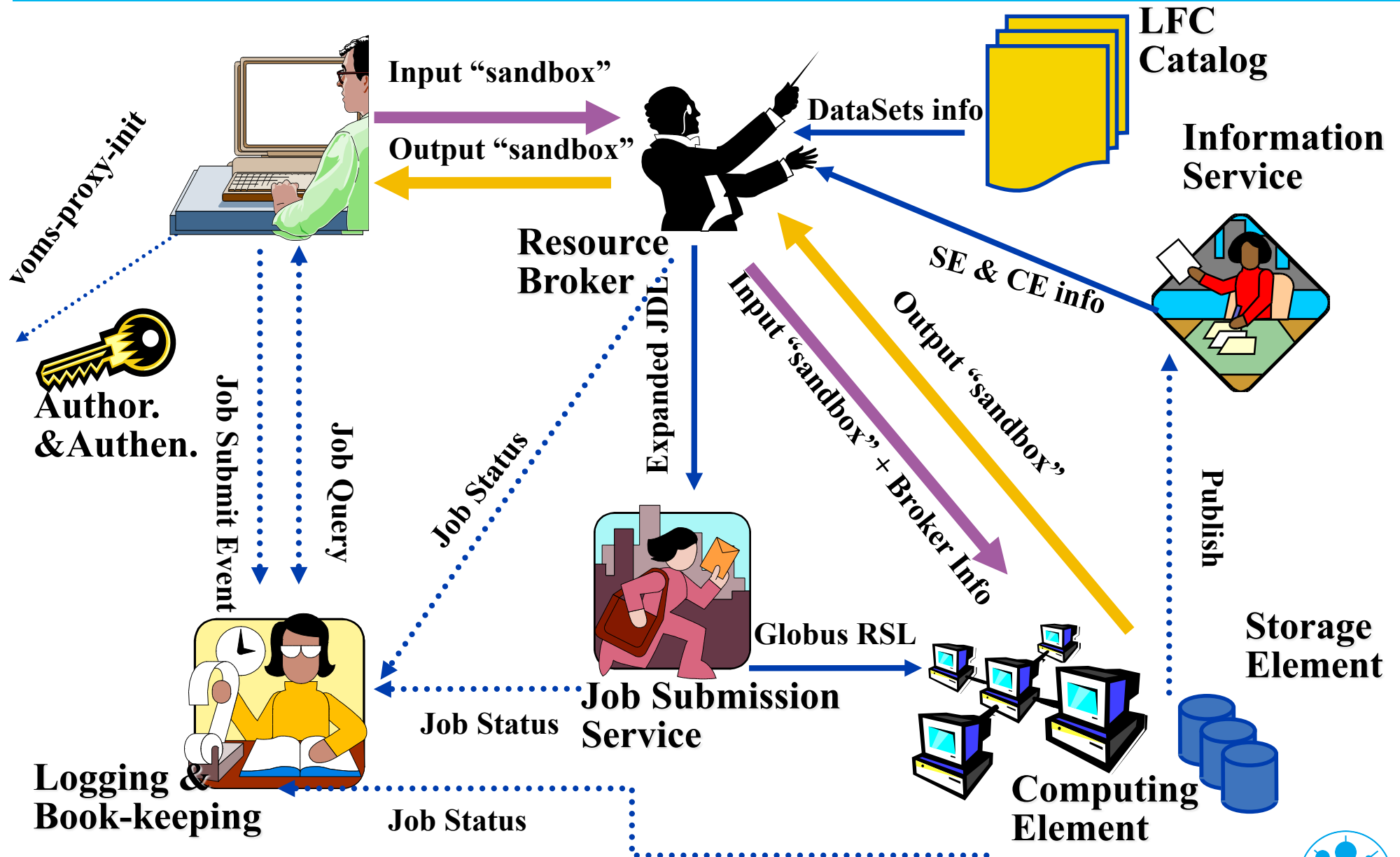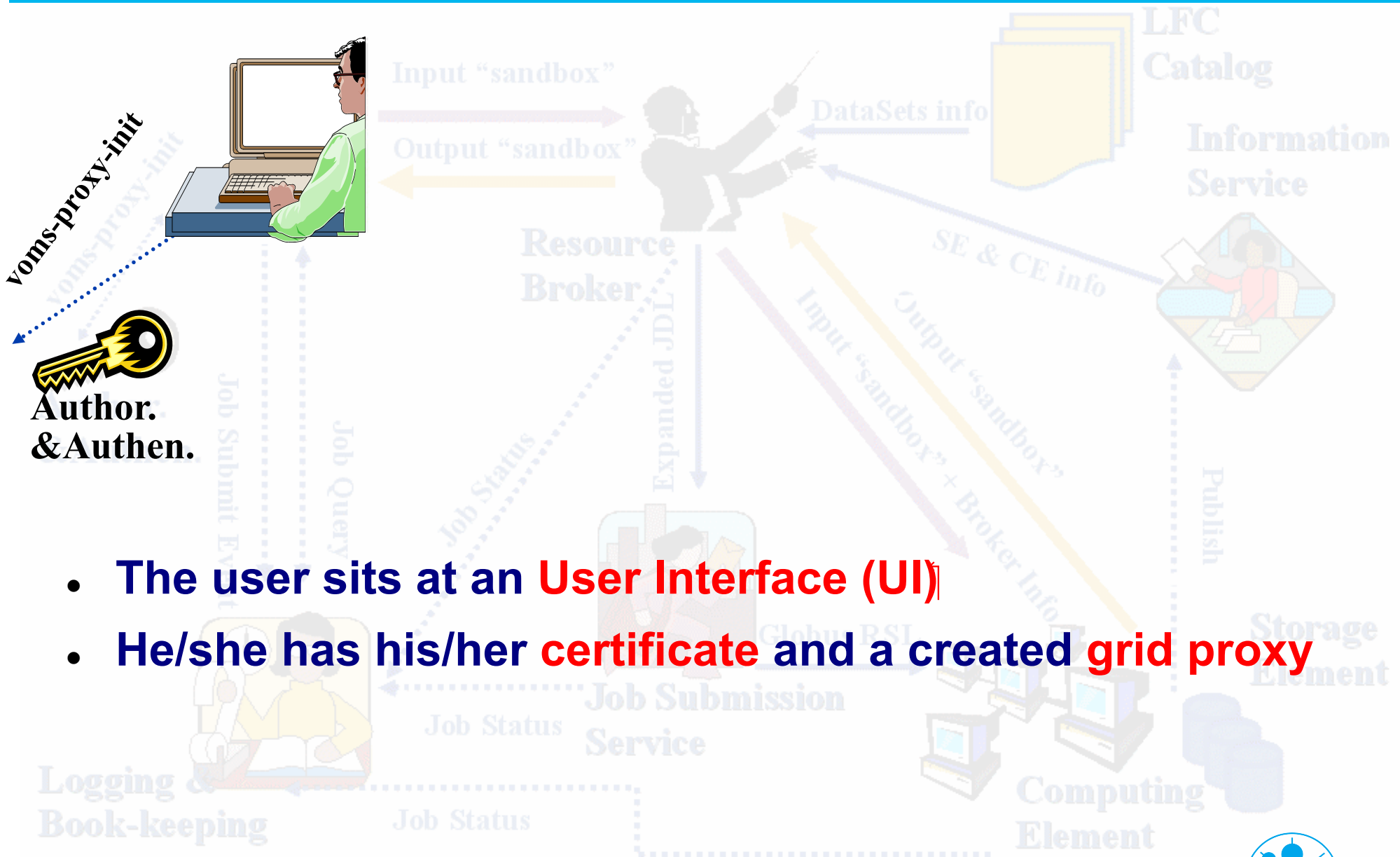
- **The user sits at an User Interface (UI)**
- **He/she has his/her certificate and a created grid proxy**

Input "sandbox"

Output "sandbox"

Resource
Broker (WMS)
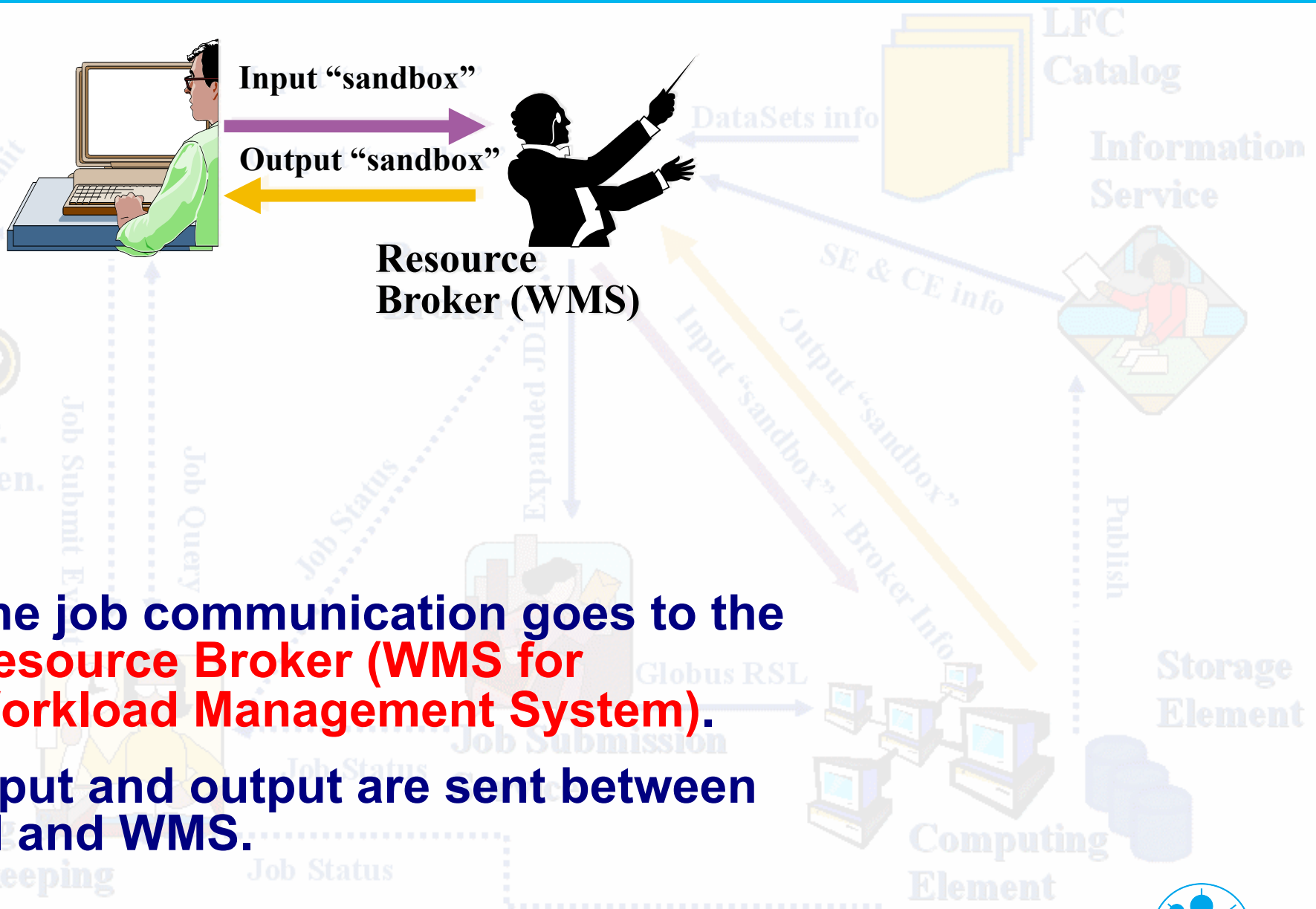
- **The job communication goes to the Resource Broker (WMS for Workload Management System).**

- **Input and output are sent between UI and WMS.**

Input "sandbox"

Output "sandbox"

DataSets info

Information Service

Resource Broker (WMS)

SE & CE info

Publish
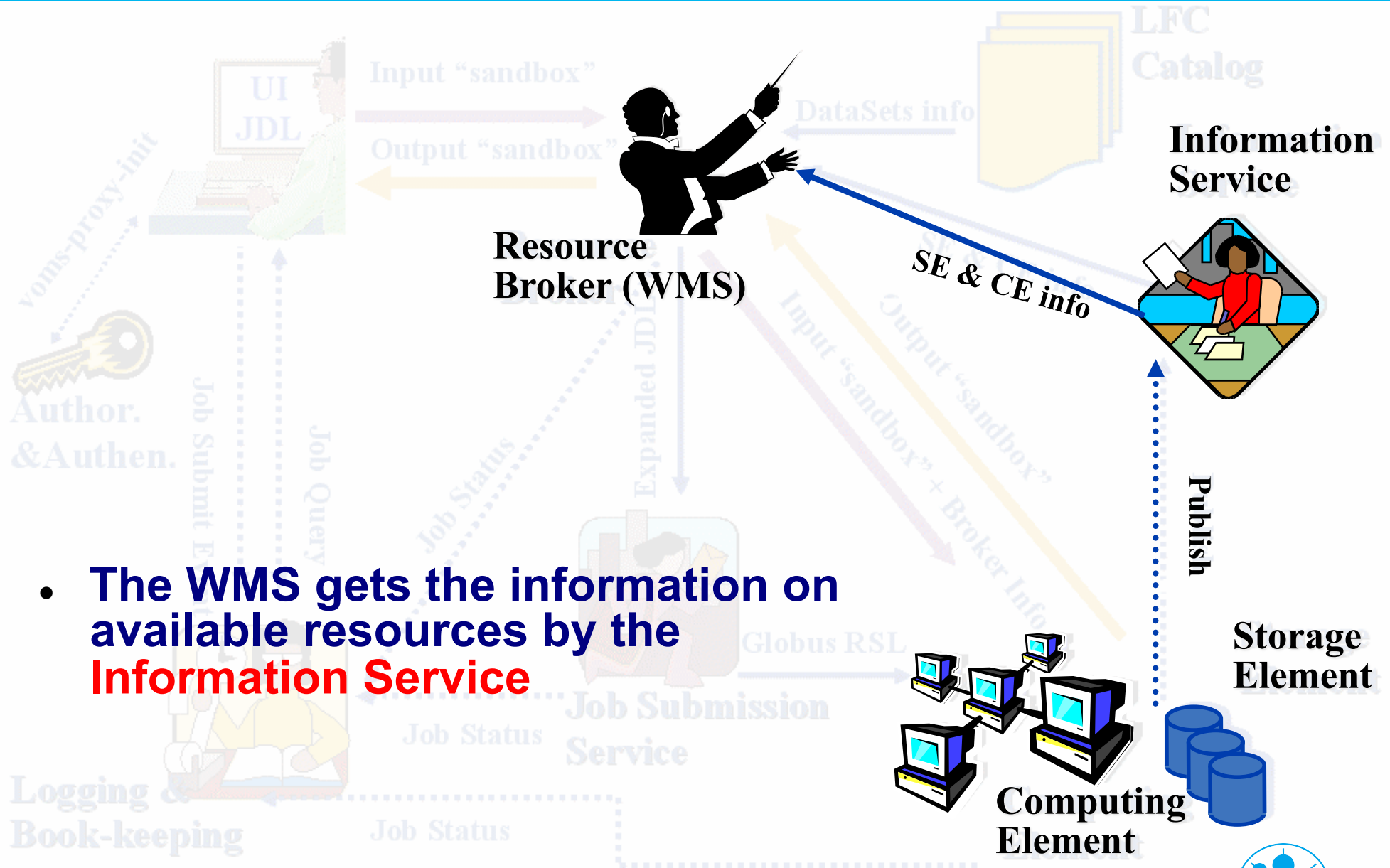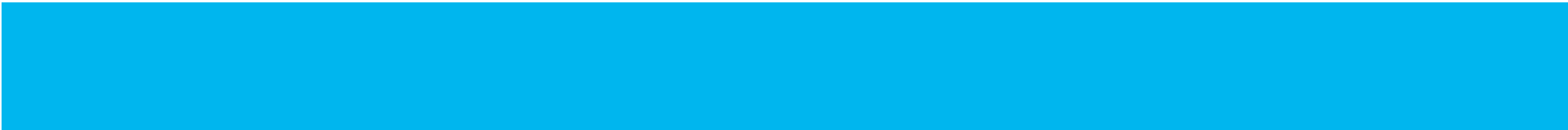
- **The WMS gets the information on available resources by the Information Service**

Storage Element

Computing Element

LFC Catalog

Input "sandbox"

DataSets info

Output "sandbox"

Resource Broker (WMS)

Input "sandbox" + Broker Info

Output "sandbox"

- **With this information the WMS distributes the jobs to the different Computing Elements (CE) at the grid sites**

- **The information on available data on the Storage Elements (SE) is provided by the LHC File Catalogue (LFC)**

Storage Element

Computing Element

DESY

gLite - Information System

> **The resources on the grid can be seen with**

> **lcg-infosites**

- ▪ Have a look at the available resources
  lcg-infosites --vo dech ce
- ▪ lcg-infosites --vo dech se
- ▪ lcg-infosites --vo dech wms
- ▪ lcg-infosites --vo dech lfc

- ▪ Also try other VOs.
- ▪ lcg-infosites --vo ops ce
- ▪ lcg-infosites --vo ops se
- ▪ lcg-infosites --vo lhcb lfc

Job Description Language

- **Job Description Language (JDL)**

  – The JDL is used to specify special needs of your job like necessary input-files/variables, generated output- and logging-information or requirements concerning the computing resources on remote sites.

  ```
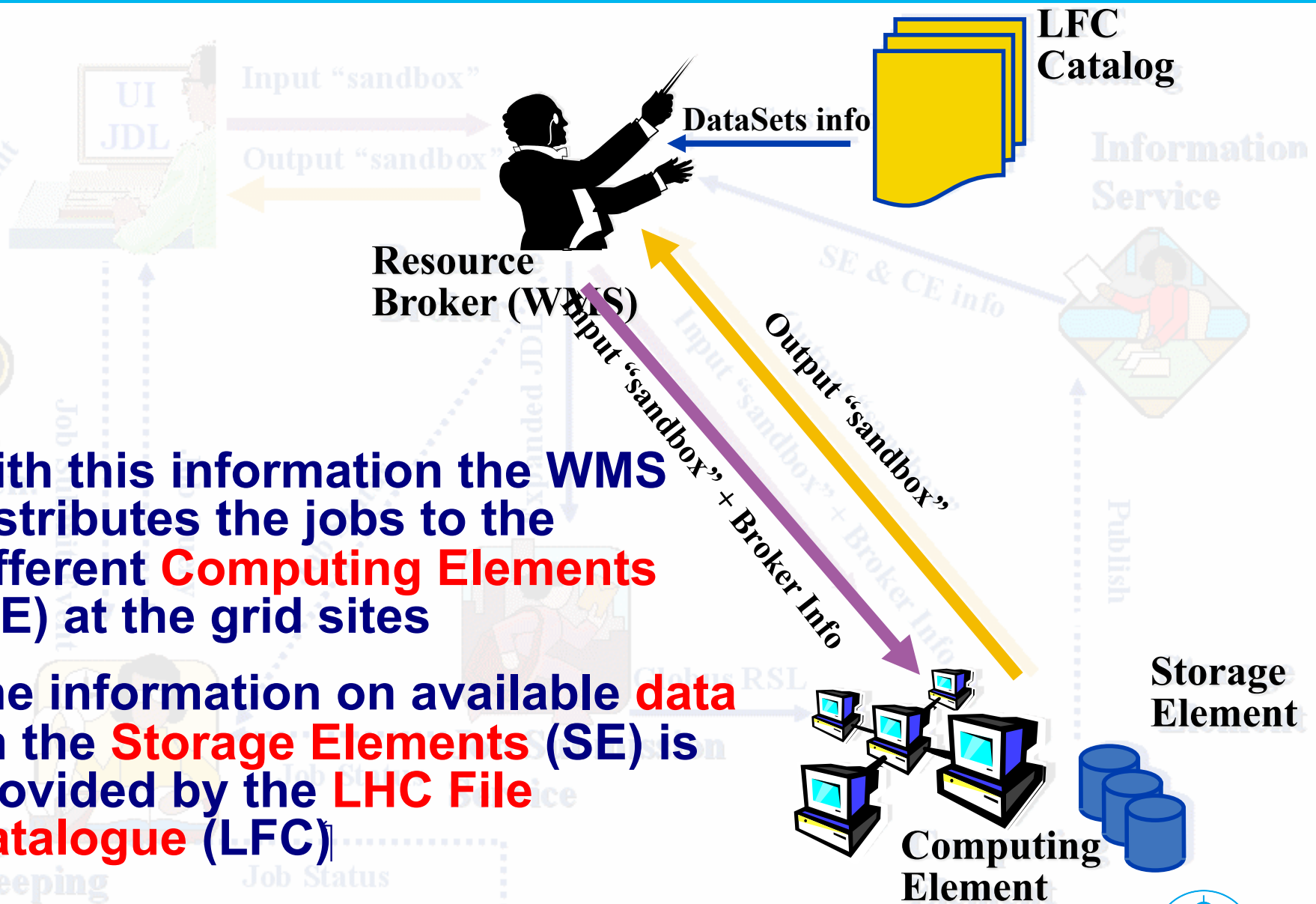  Executable = "bin/script";
  Arguments = "list of arguments";

  StdOutput = "std.out";
  StdError = "std.err";

  InputSandbox = {"input", "files", ...};
  OutputSandbox = {"std.out", "std.err", "output.files", ...};

  VirtualOrganisation = "dech";
  RetryCount = 2;
  ```

  – In the InputSandbox wildcards can be used.
  – **… or have a look at $HOME/job/computation.jdl used for the MC jobs!**

Topics

> **Preliminaries**
  - Networking

> **Security**
  - Virtual Organizations
  - Authentification
  - Certificates
  - Grid Proxy

> **Grid User Support**

> **Data Management**
  - Overview and LFC
  - LFN and PFN
  - Copy, Register, Replicate, Delete

- **gLite**
  - Job Workflow
  - Workload Management System
  - Job Description Language

- **Job Handling**
  - A "real life" example

- **General recommendations**

Job Handling – WMS

> **Job Handling commands begin with glite-wms-job-**
> - `glite-wms-job-list-match -a --rank myjob1.jdl`
> - `glite-wms-job-submit -a myjob1.jdl`
> - `glite-wms-job-submit <-r RESOURCE> -a myjob1.jdl`
>   - RESOURCE is a specific CE, like
>   grid-ce5.desy.de:2119/jobmanager-lcgpbs-default
> - `glite-wms-job-submit -o JobID.txt -a myjob1.jdl`

> **WMS stands for "Workload Management System", and is the new name for the product previously called "Resource Broker". I use the term "RB" instead of "WMS" to denominate the functionality, always meant is the new product "WMS"**

> **The output of the submit-command will contain something like**
>   **https://grid-lb1.desy.de:9000/CbIqayj-TIoubOTOvx6qIQ**

> **You will need this JobID for other commands!**
> - Might use a JobID.txt file

Hands on

> **Get the job status:**
  - `glite-wms-job-status [-v 1] <JobID>`
  - Parameters:
    - `<JobID>` Specifies the Job URL (https…)
      - *Also via* -i JobID.txt
    - `-v <0|1|2|3>` The higher the verbosity level, the more information you get (e.g. about rescheduled jobs)
  - The command can have several job ids
    - `glite-wms-job-status <job_id_1> ... <job_id_n>`
  - To get information on all submitted jobs there are several options
    - `--all`
    - `--from/--to [MM:DD:]hh:mm[:[CC]YY]`
    - `--status <status> or --exclude <status>`

Hands on

> **In case something goes wrong, your job can be cancelled via:**
  - `glite-wms-job-cancel <jobID>`

> **Get job output**
  - `glite-wms-job-output <jobID>`

  - Standard output directory:

    `/tmp/jobOutput/<username>_<glite_WMS_job_id>`

The "real life" example

> **Look at $HOME**
  - computation.jdl
  - script.sh
  - timeout.sh

> **This job takes up to a few minutes**
  - Submit the job to the Grid

> **Rather short job, but shows how physicists work**
> **Some jobs have been sent to the Grid already:**
  - Look at $HOME/jobs.id

Mean Job duration on WNs (with error bands)
(Production grid-ce5, ATLAS VO) Y-Axis is seconds
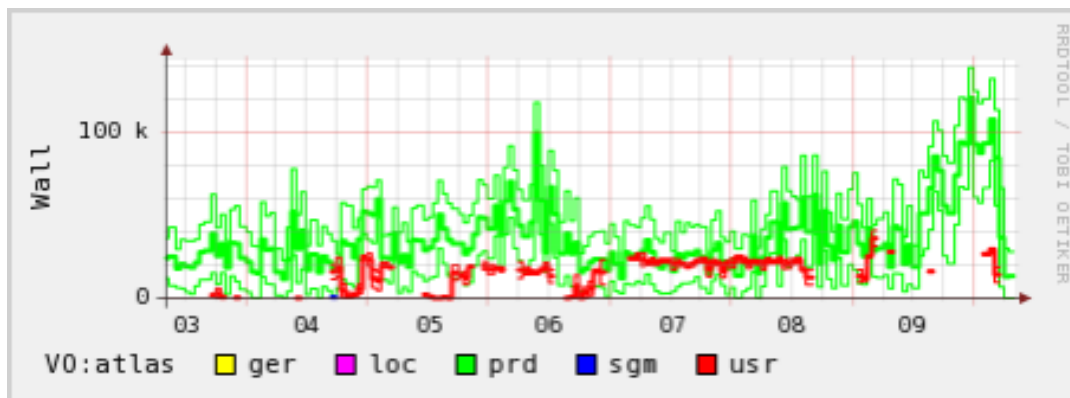
> **Download all data files (Already done previously)**
> **Get the values you are interested in (e.g. grep … > …)**
> - Or use the file /home/common/simulations_all.txt
> **Compute the mean value (e.g. awk or tr+sed+bc)**
> **Do the analysis**
> - E.g. produce pairs of old_value-Mean new_value-Mean per line (e.g. awk)
> - Compare number of "+… -…" and "+… +…" pairs
> - What do you see?

> **For the advanced:**
> - Try to do reproduce the distribution functi
> - With whatever tool you like
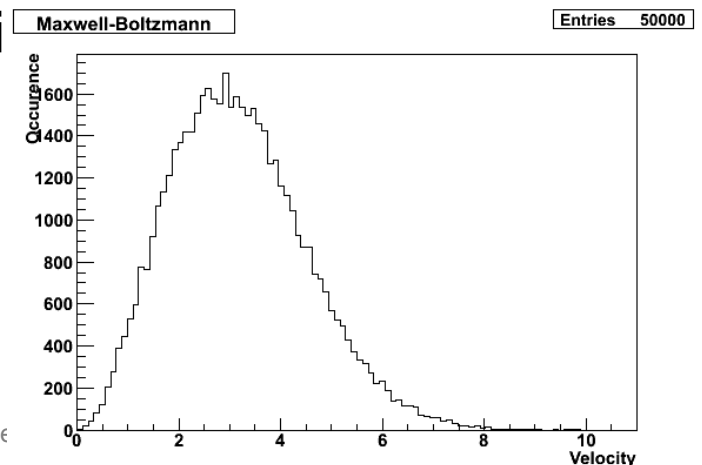
Commands needed for the analysis

> **Compute mean value**

> **Analysis**
- E.g. create pairs of "OldValue CurrentValue"
- Then e.g. compare with mean value

> **The mean value ("Mittelwert") is the wrong measure here: Instead one should speak about the median**

  - The media is the value exactly in the middle of the sample, having the same number of particles with smaller than with larger velocities
  - The median is smaller than the mean value for this distribution (~3.1)
  - Repeating the previous analysis using the median will have the expected result: The first value does not influence the second value

> **Remember this when you play Roulette:**

  - The ball has no memory! Really!

- **Preliminaries**
  - Networking
- **Security**
  - Virtual Organizations
  - Authentification
  - Certificates
  - Grid Proxy
- **Grid User Support**
- **Data Management**
  - Overview and LFC
  - LFN and PFN
  - Copy, Register, Replicate, Delete

- **gLite**
  - Job Workflow
  - Workload Management System
  - Job Description Language
- **Job Handling**
  - A "real life" example
- **General recommendations**

General recommendations

> **Make your job robust!**
  - Everything can fail:
    - Worker node defective (HD, RAM, Network, installed SW,…)
    - Data transfer from/to remote fails for unknown reasons
    - Some actions takes forever
  - Envelope every action with a watchdog (e.g. kill the action after X minutes if the command has not returned a signal of its own)
    - Poor-mans solution using bash, more advanced sig-handlers: Perl…
  - Have a "plan B":
    - If data transfer to site X fails after 5 minutes, try site Y, then site Z
  - Your job should be completely autonomous
> **Make job handling robust!**
  - You cannot treat 100 jobs/day by hand
  - Build a robot with automated job creation, submission, output retrieval and resubmission
> **Rely on standards, and use them!**
  - If you build a solution for one particular site, your job will only run there. You will give away much larger resources

General recommendations (2)

> **File handling**
  - Have high-level tools that rely on the standards
  - Make data transfer from site-to-site failsafe
  - Automatic retransmission

> **Have a plan on how to manage data**
  - Who can write data?
  - Who can access it?
  - Who can delete it?
  - When can it be deleted?
  - Owners come and go - the data persists. How do you react?
  - What happens if data gets lost?

> **Make sure Catalogue and SE are consistent**
  - How can one enforce this?

Literature – Further Information

> **gLite 3 User Guide:**
> **https://edms.cern.ch/file/722398/gLite-3-UserGuide.pdf**

> **Job Description Language attributes specification:**
> **https://edms.cern.ch/file/590869/1/EGEE-JRA1-TEC-590869-JDL-Attributes-v0-8.pdf**

> **DESY Grid pages:**
> **http://grid.desy.de/**