Management of Simulation Productions in Auger using GRID Technology

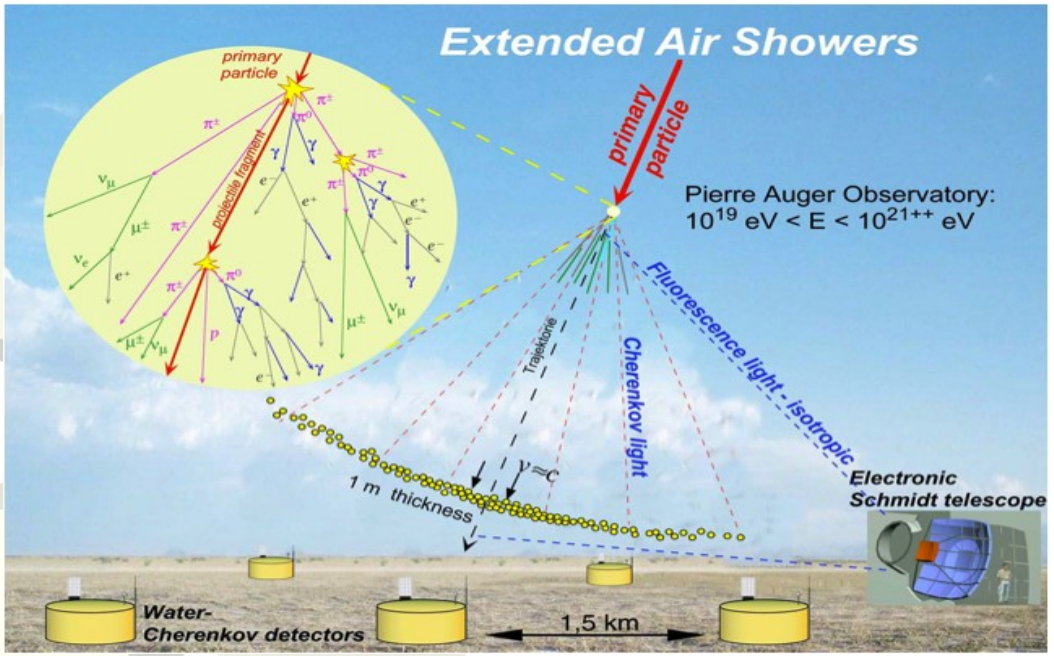
Jiri Chudoba[1], Julio Lozano[2],
Ginés Rubio[2], M.D. Serrano[2]
1 **FZU Prague**
2 **Universidad de Granada**

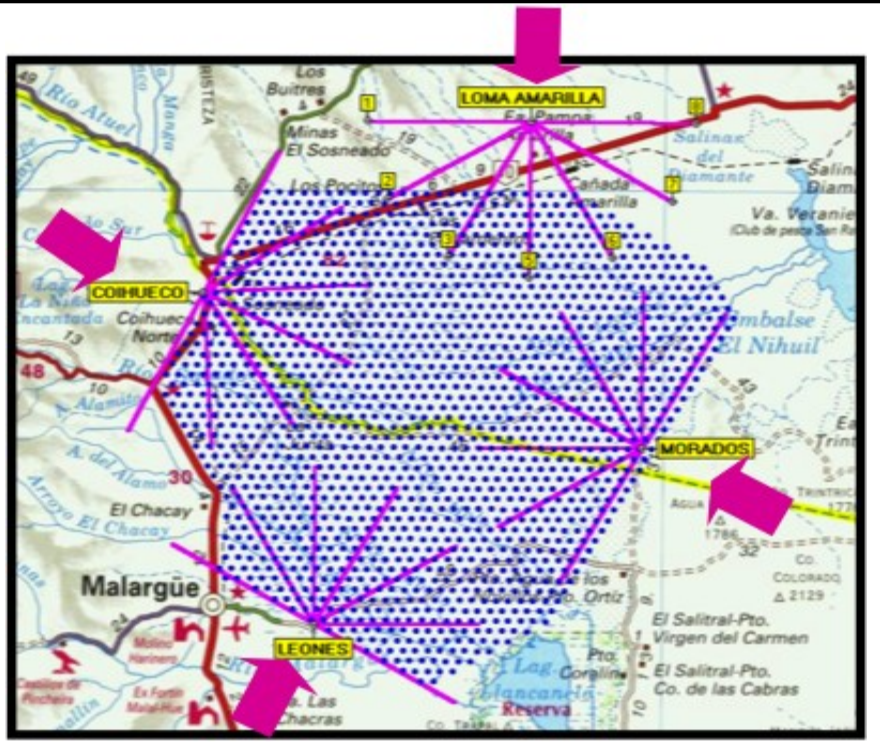
Pierre Auger Observatory

UHE Cosmic Rays Observatory :

Hybrid detector based on 2 different technologies:

- **SD (Surface Detector)** : array of ~1600 water tanks placed 1.5 km apart for a total area of ~3000 km^2
  - 3 PMTs transmitting wirelessly FADC traces
- **FD (Fluorescence Detector)** : 4 stations on terrain overlooking the area where the tanks are located.
  - 6 bays each with a fluorescence telescope covering 30 degrees in azimuth
    - 22x20 PMT array





Located near Malargüe, south of Mendoza in Argentina

- Blue dots : water tanks
- Pink arrows: location of telescope eyes
  - Pink lines: coverage in azimuth

Fluorescence detector measures longitudinal shower evolution

Surface detector obtains signals from particles reaching earth level

Auger Simulated Data: showers

Shower generation :

- CORSIKA and AIRES are the software packages that generate those kind of events. In official simulations we have only used CORSIKA

- A compilation tool lets the user decide on low (Fluka, Gheisha) and high (epos, QGSjetI/II) energy interaction models and some other options which do not change for a specific *library*. It requires only an input card (run number, energy, zenith angle, first interaction point, seeds, etc …) which is specific for each job.

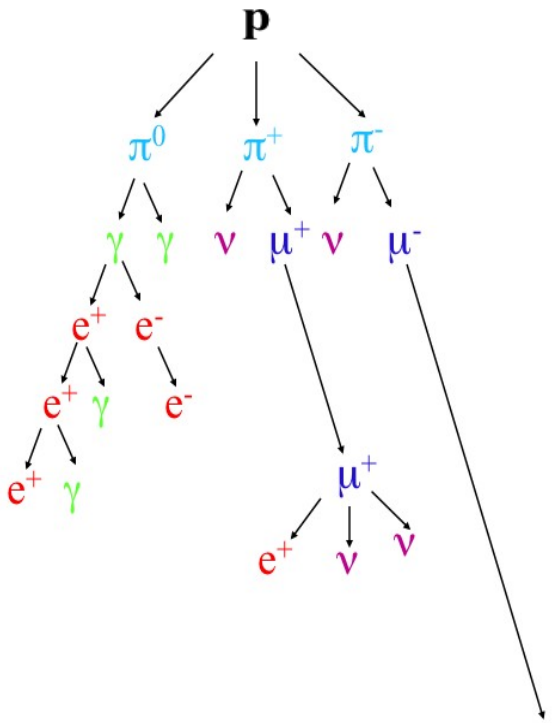
- CORSIKA package can be retrieved from Storage Elements where it is copied, but to increase efficiency on many sites it is available from the Software Area (specific repository where a Grid user having Software Manager Role can place software)

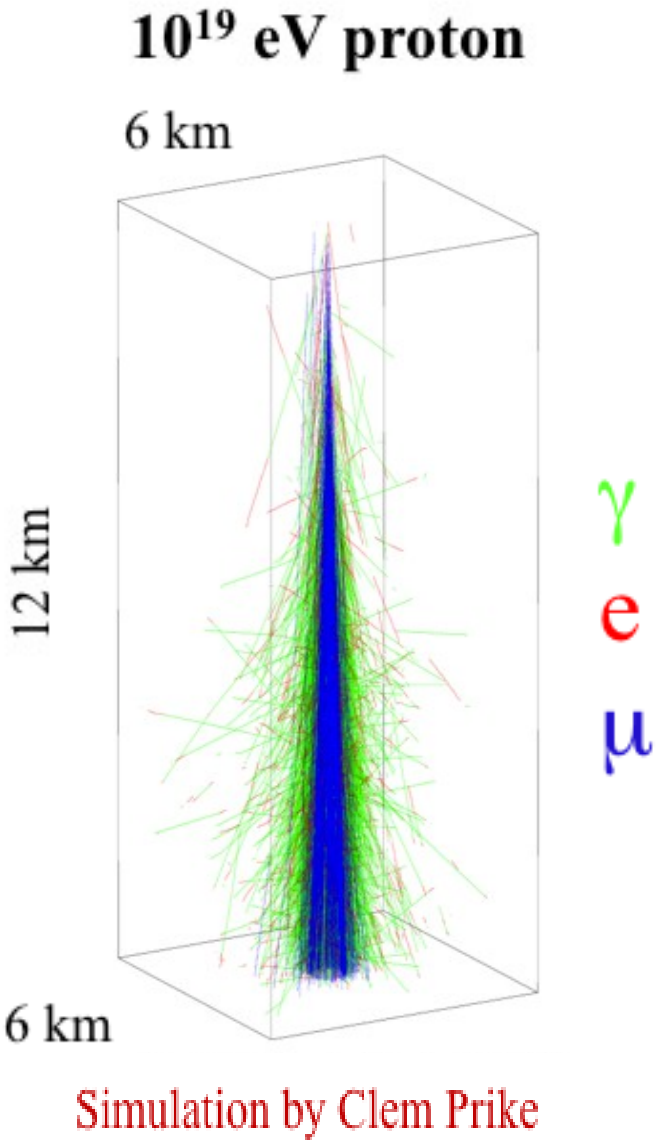
- Billions of particles being treated: *thinning* method is needed ! Shower 'particle' files collect all characteristics of particles at ground level . 'Longitudinal' files contain information on the longitudinal development of the shower and are only few MBs big at most.

Auger Simulated Data: showers

Shower example :

Cosmic ray primary interacts creating mostly secondary pions generating an electromagnetic shower and a big amount of muons





Auger Simulated Data: showers

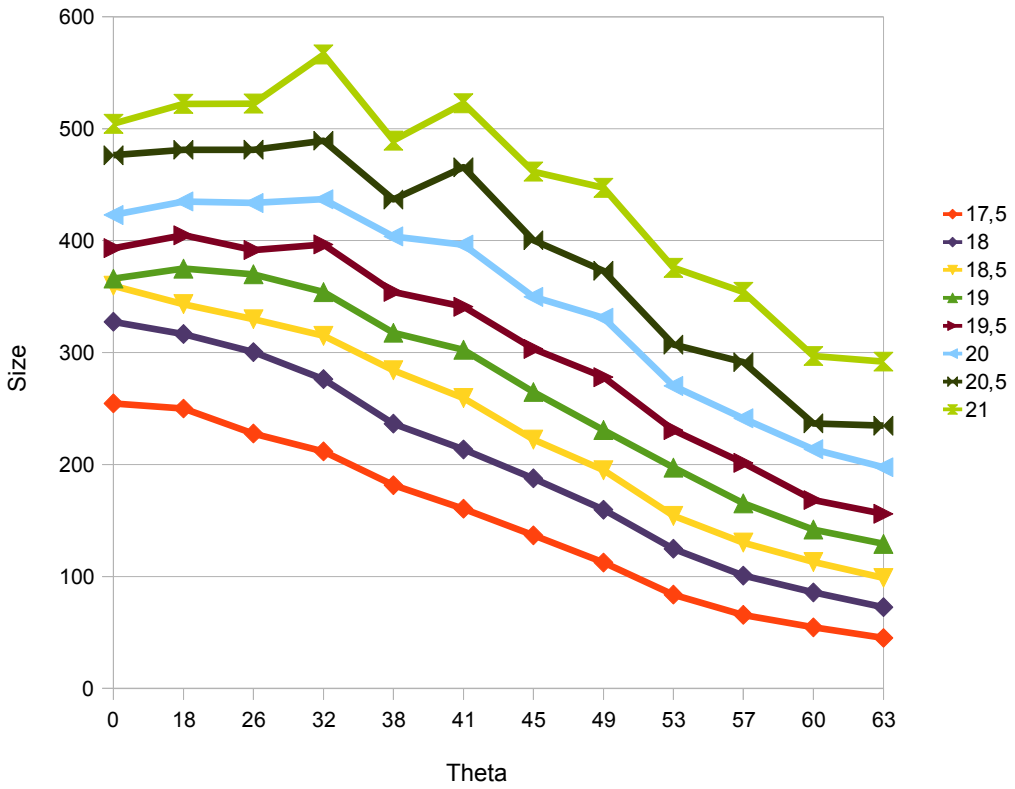
File size of ground particle files (epos high-energy model)

Even if we show smooth curves fluctuations are very big, both in file size and CPU time

Epos. Proton. Energy

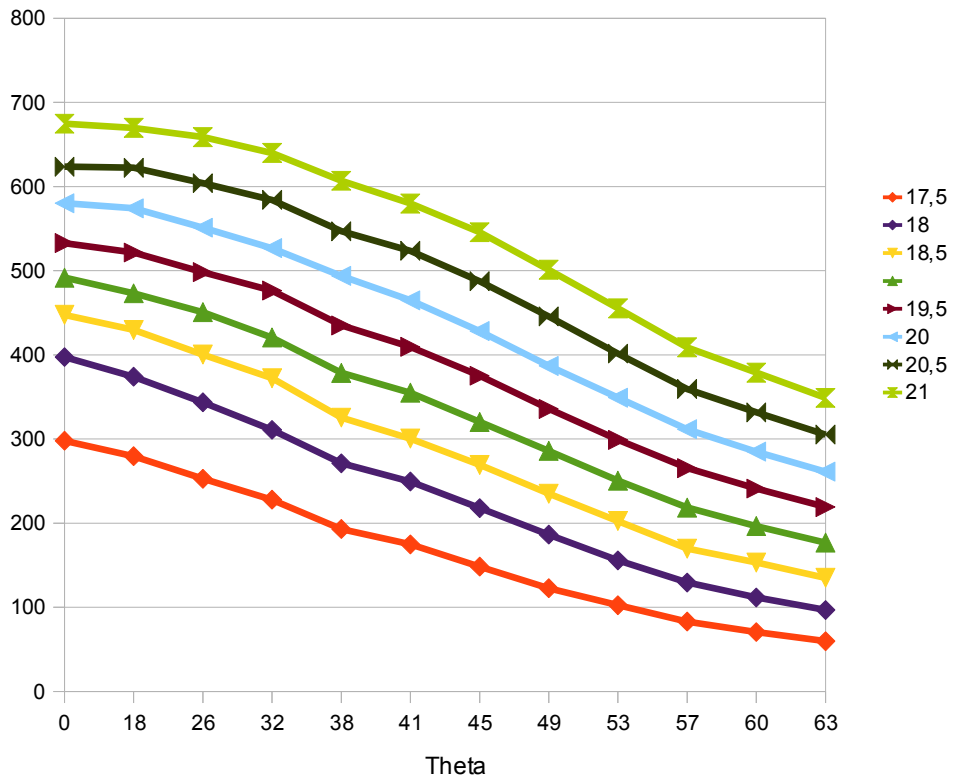
Size (MB)

energy expressed as log(E) (18 → EeV)



Epos. Iron. Energy

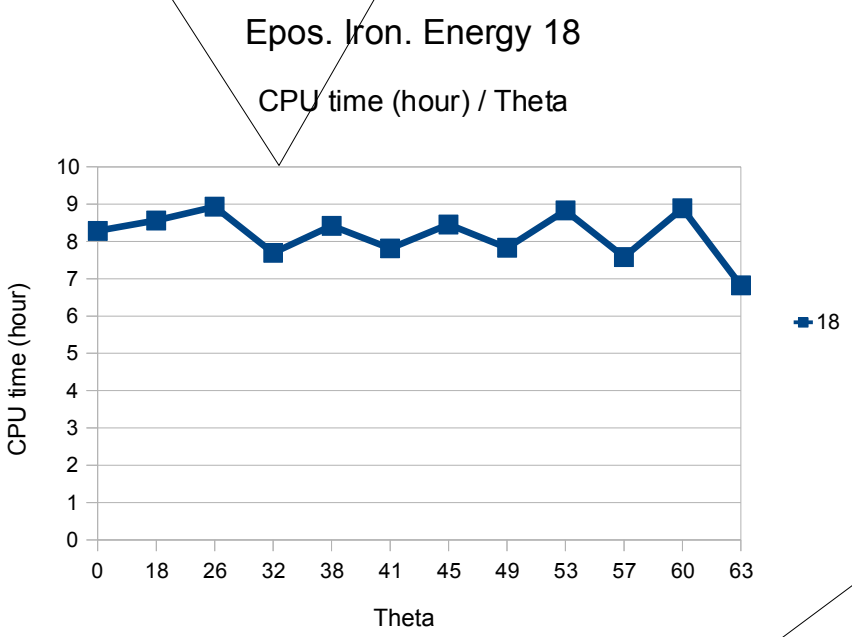
Size (MB)



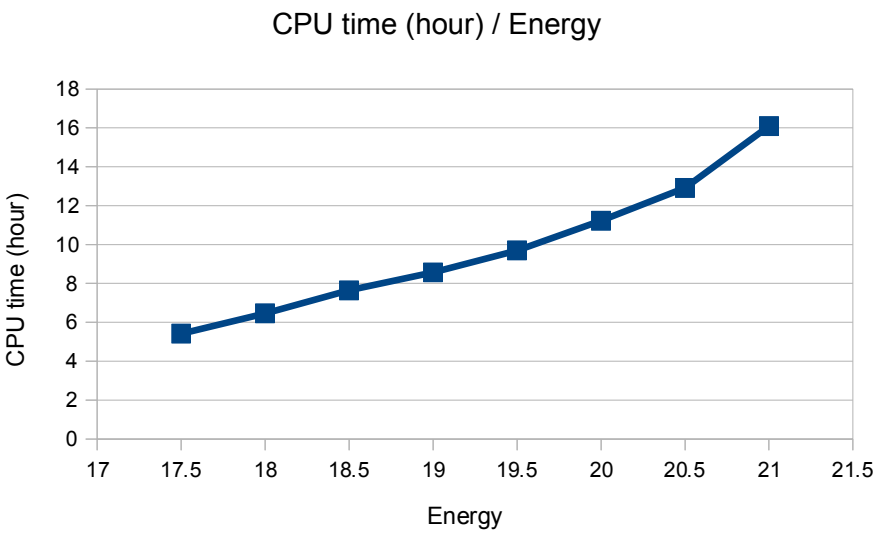
Auger Simulated Data: showers

CPU time (epos productions)

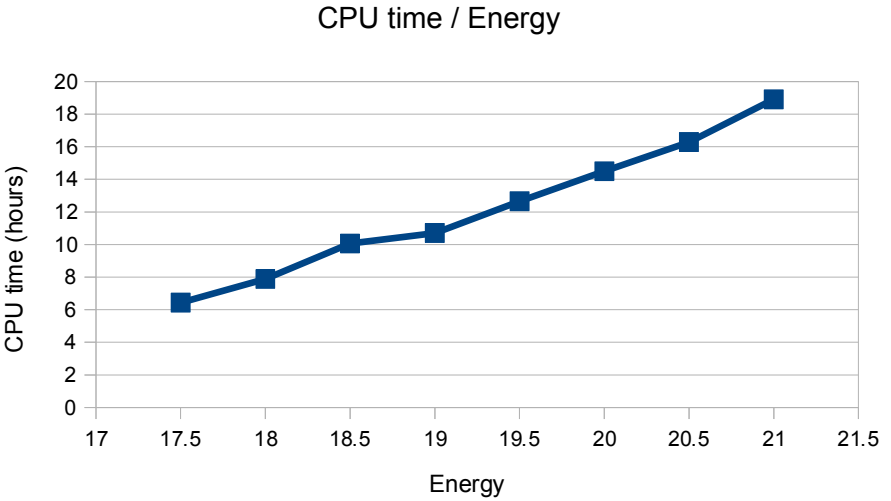
Almost no theta dependence up to 63^0



Epos. Proton.



Epos. Iron.



CPU time increases with energy and irons are slightly more CPU demanding

Auger Simulated Data: *Offline* event

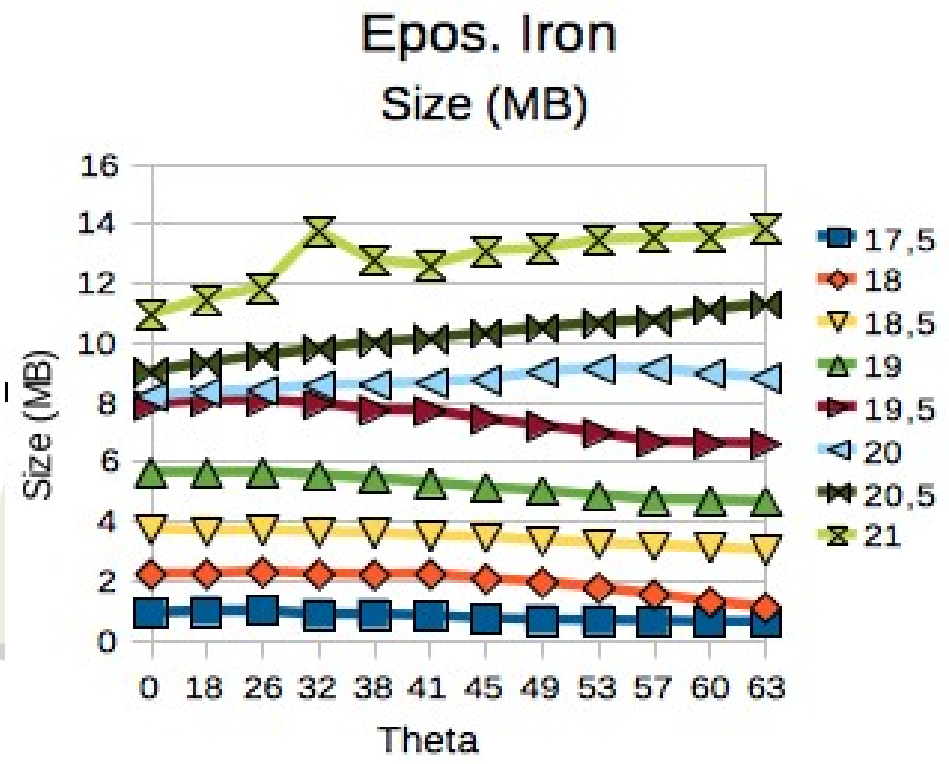
Detector simulation and reconstruction :

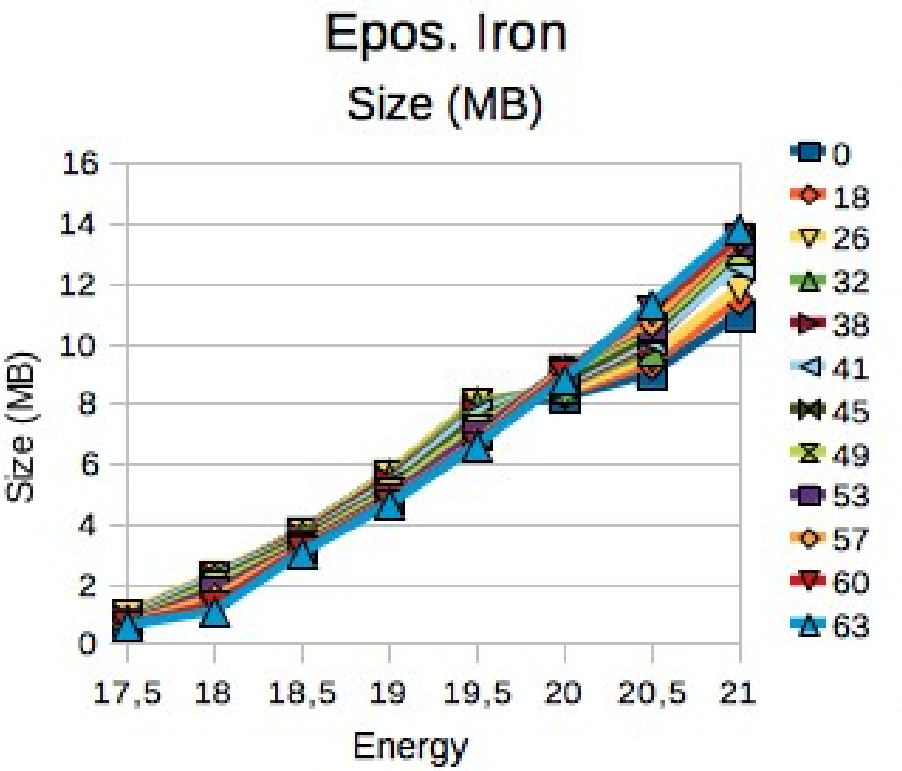
- Auger registers :
  - Cerenkov light of particles remaining at ground level as secondaries from the shower created by an UHE cosmic ray primary
  - Fluorescence light emitted while those secondaries are traveling through the atmosphere
- *OffLine* (DPA) package simulates the response of the SD and FD
- Modular package driven via xml configuration files determining sequence of modules to be used and running parameters
- File sizes are usually of order 10s of MBs
- Needs previously generated shower files → showers have to be kept on Grid
- Software package is heavy and has many dependencies → software compiled and installed on Software Areas

Auger Simulated Data: *Offline* event

File size of output data

Files included are one with detector simulation data and another one with a summary of the event reconstruction

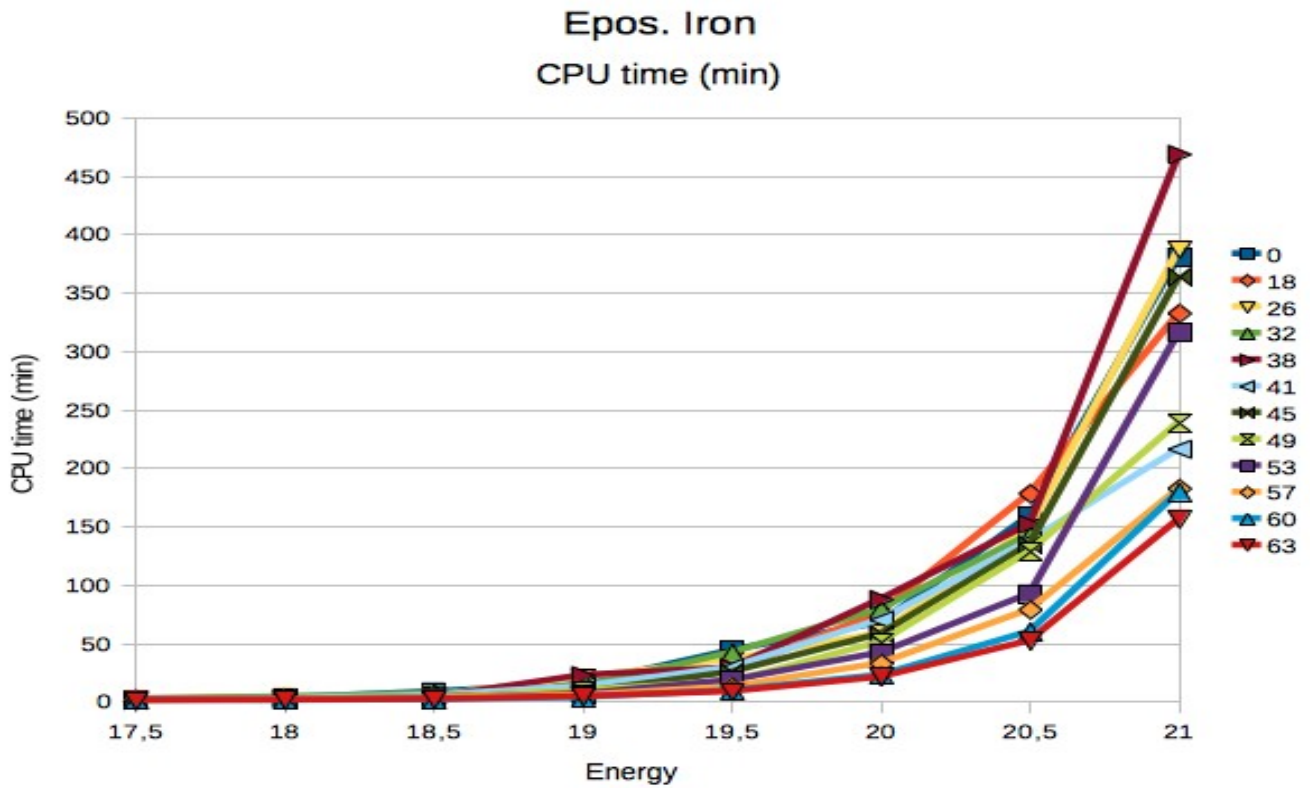




Auger Simulated Data: *Offline* event

CPU time (Offline job)

Each *Offline* job involves the simulation of 5 times the same input shower (changing seeds)



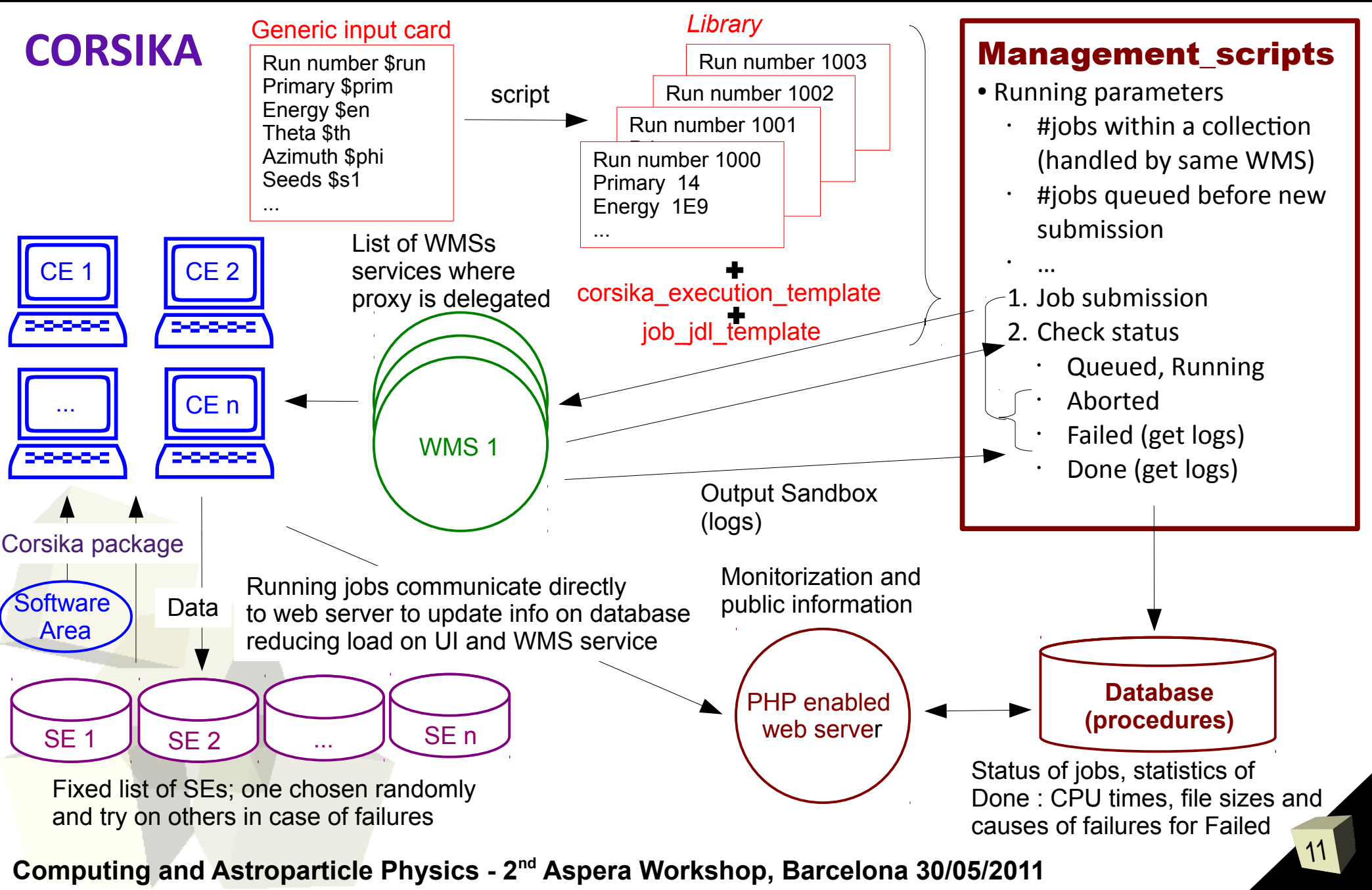
CPU time increases with energy and depends on the zenith angle in a non-linear way

Small differences between primaries (p, Fe)

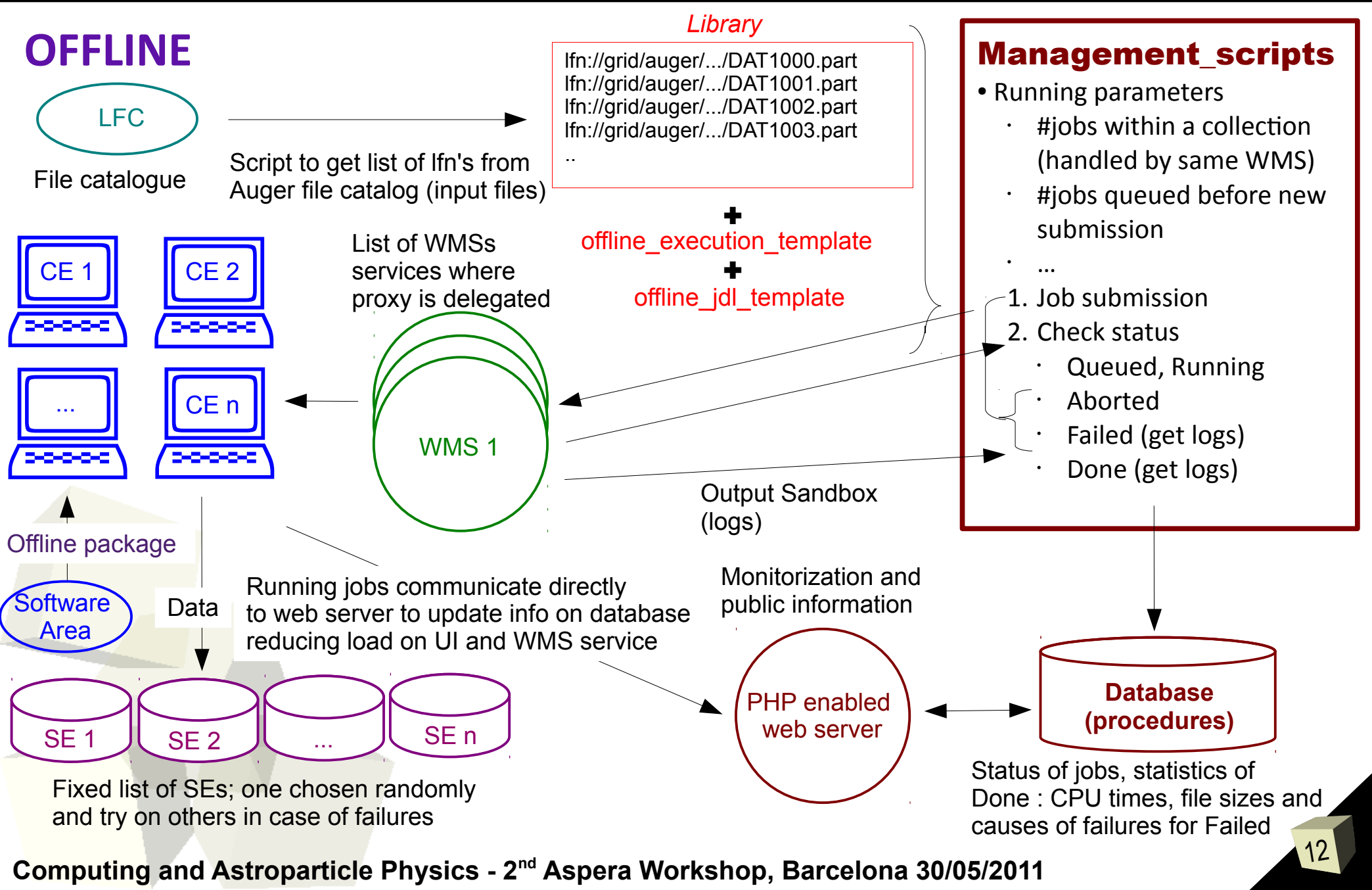
Grid Technology

- Grid Technology allows us to perform massive productions on limited time scales:
  - Computing infrastructure with tens of thousands of CPUs to execute jobs
  - Storage sites with tens of TBs to place output files
- Glite (Grid middleware) provides user commands to handle job management, output retrieval, etc …
- To avoid too much manual intervention we have developed a set of scripts (bash and python) wrapping up previous scripts written by collaborators at Prague to automate all job production aspects
  - Corsika and *Offline* productions have each theirs set of scripts

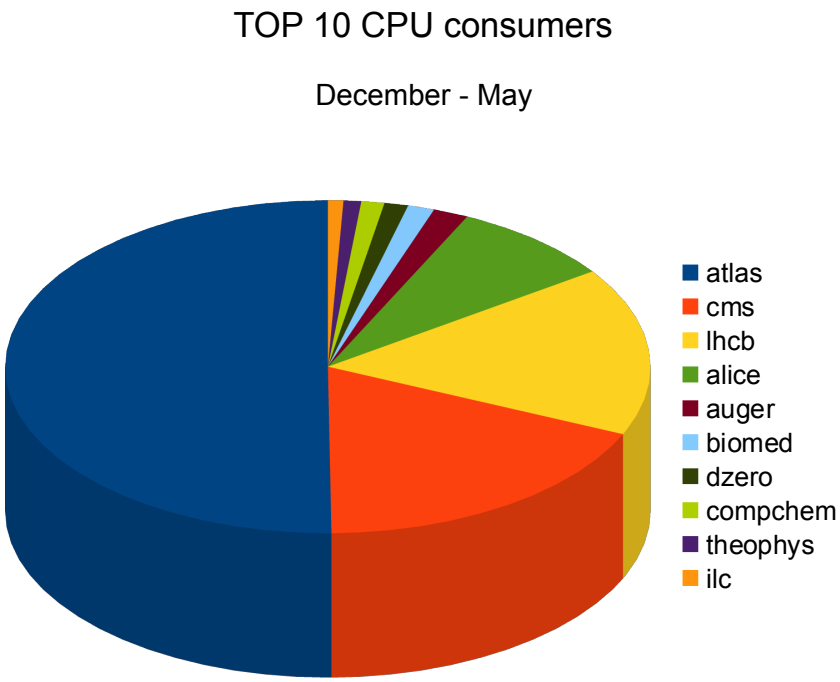
Grid Technology: implementation

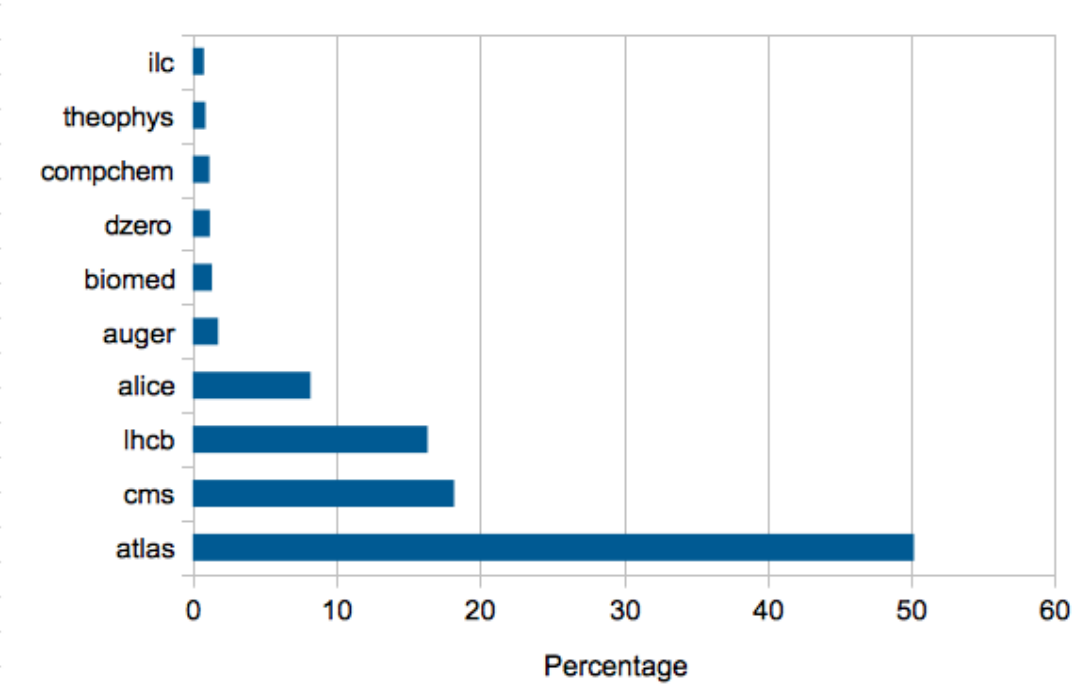


Grid Technology: implementation



Auger VO among top ten CPU consumers
(EGEE only)

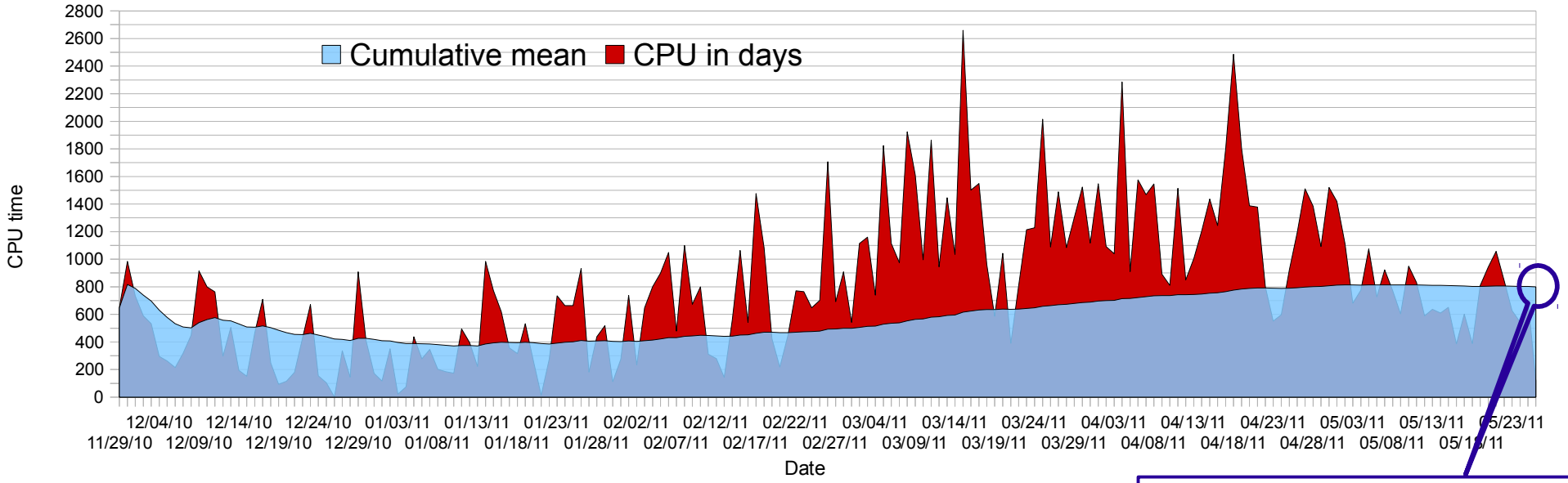




Auger consumed ~2% of total CPU time
(relatively small set of accessible sites)

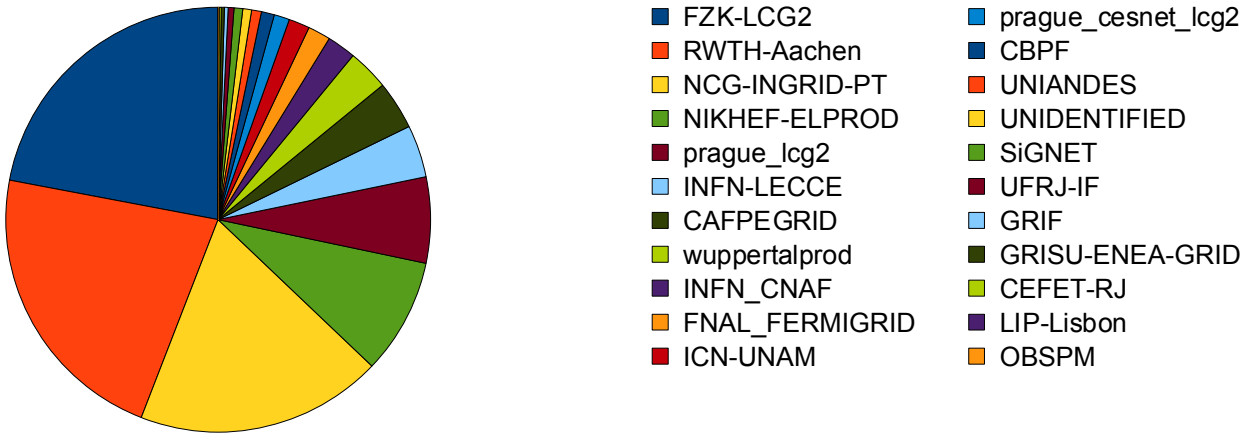
Grid usage: CPU time

Cumulative and daily CPU time (walltime)



~800 CPU*days averaged over ~6 months

CPU (in days) by site

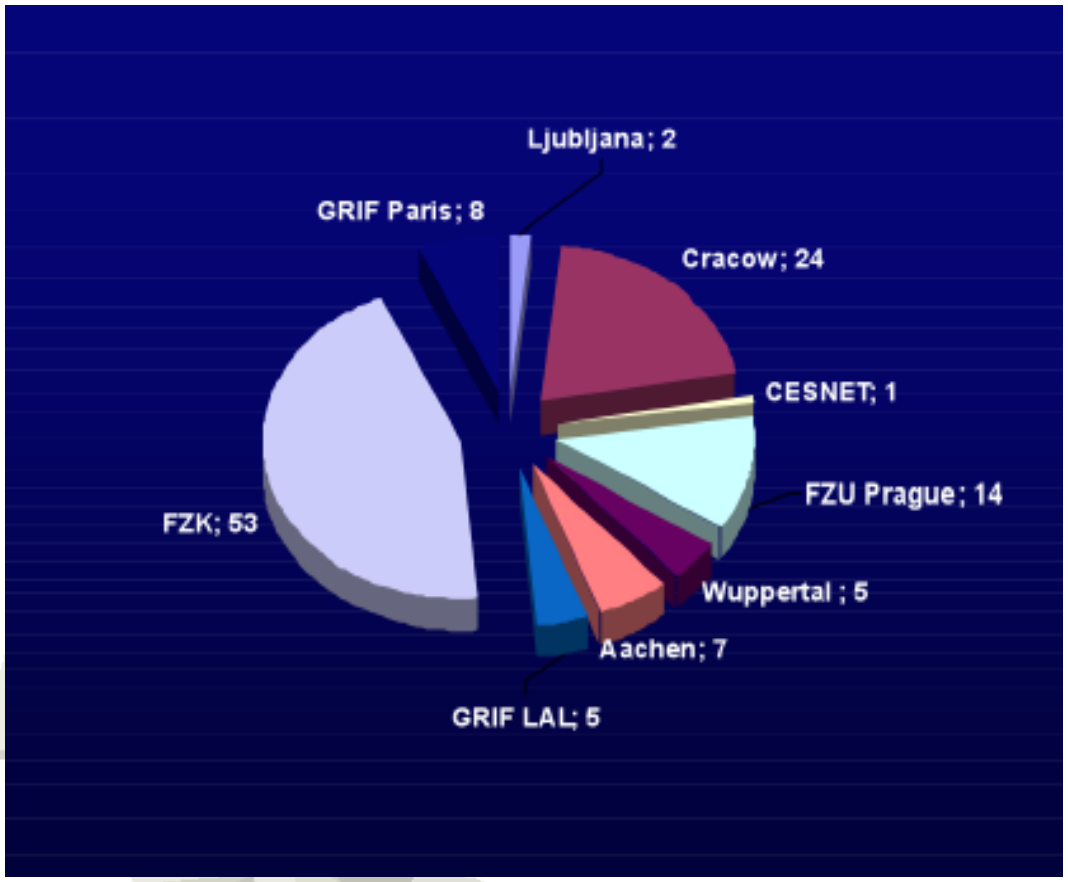


- 4 sites handle about ¾ of our jobs
- Most relevant contribution from european sites but also sizeable contribution from Latinamerican sites

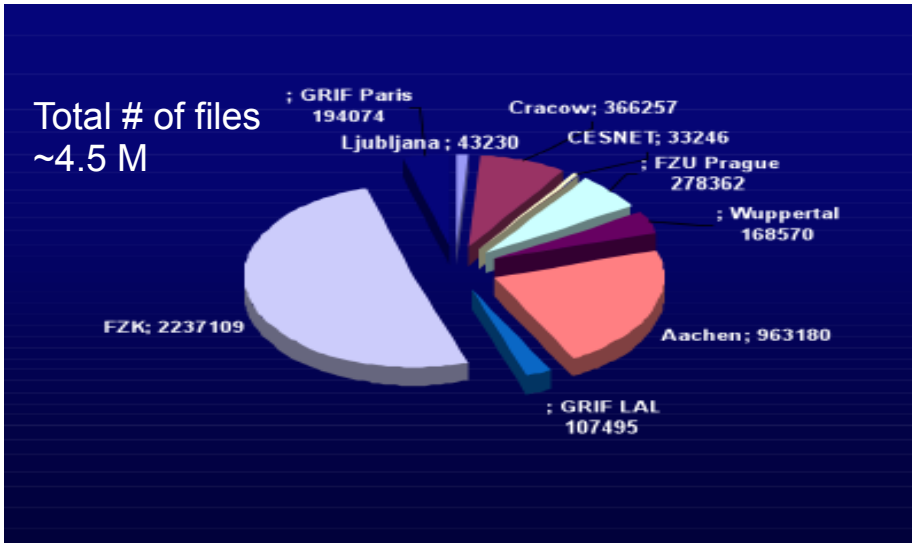
Grid usage: storage space

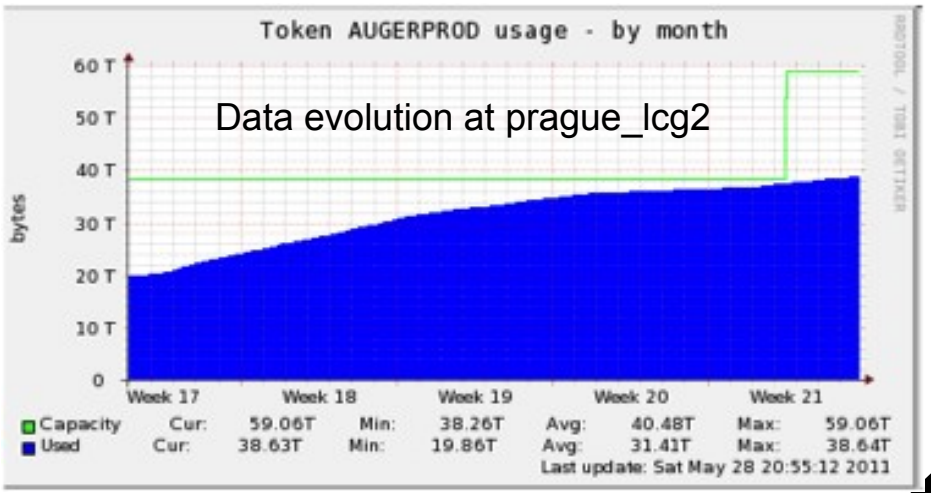
Disk space on Storage Elements

The total amount of disk space consumed is **139 TB** (**92 TB** just in the last 6 months !)



Distribution of storage space per SE site in TBs

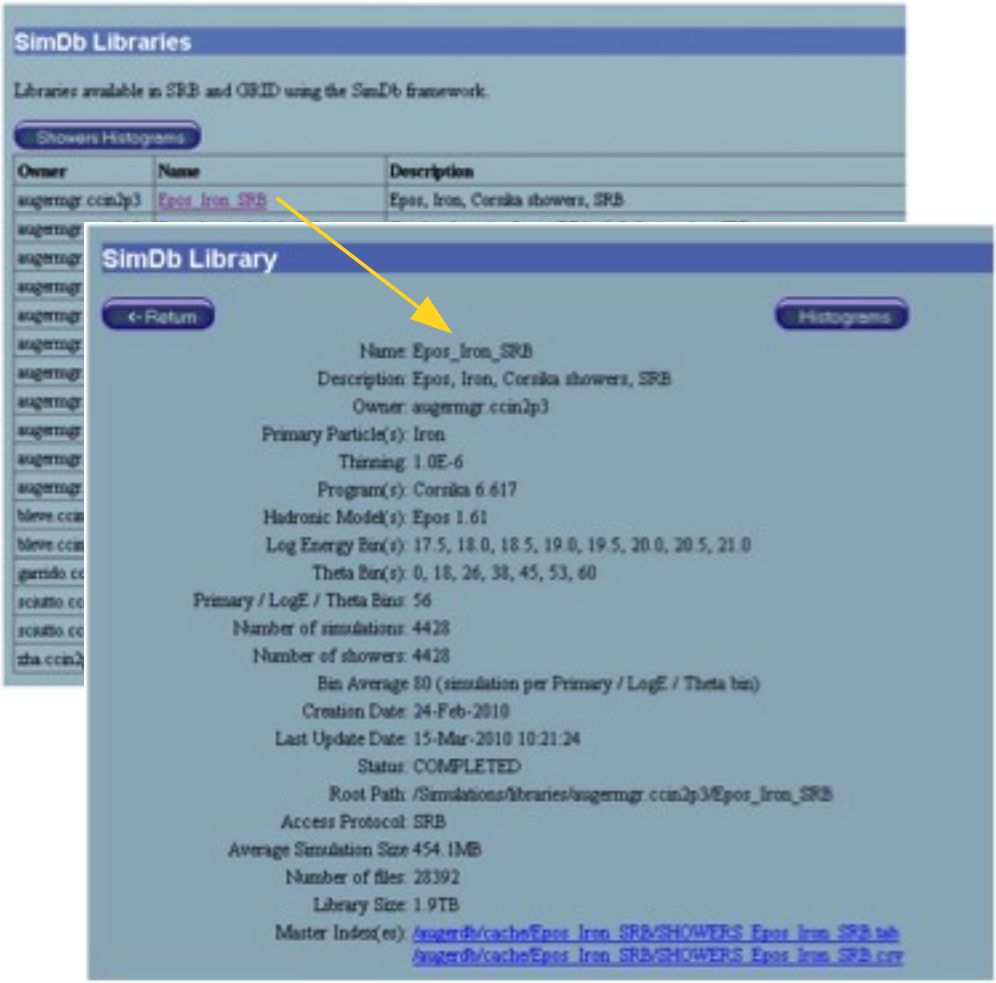




User data access

- Space on disk at SEs could become a poblem but
  - Data will remain on SEs only until Offline processing finished
  - SRB (Storage Resource Broker) used to migrate data to our reference center (CC-Lyon) while additional tools fill a database with metadata information (event indexation).

- Members of our collaboration not having Grid certificates access the data using:
  - SimDB/AugerDB; provide command line tools (SRB client tools) and dynamically generated pages for data selection and Retrieval



More information on this can be found in last year's presentation by Jean-Noel Albert (LAL)

Using Grid: experience

Positive aspects (outweight negative aspects in any case):

- Well stablished and accepted technology which is 'worldwide' available: fair amount of resources at one's disposal (computing power and storage space)
- User software provides adequate means for job handling in an easy way
- Diverse tools to help users in case of troubles; ticketing system, messaging system to make site downtimes public, etc ...

Shortcomings :

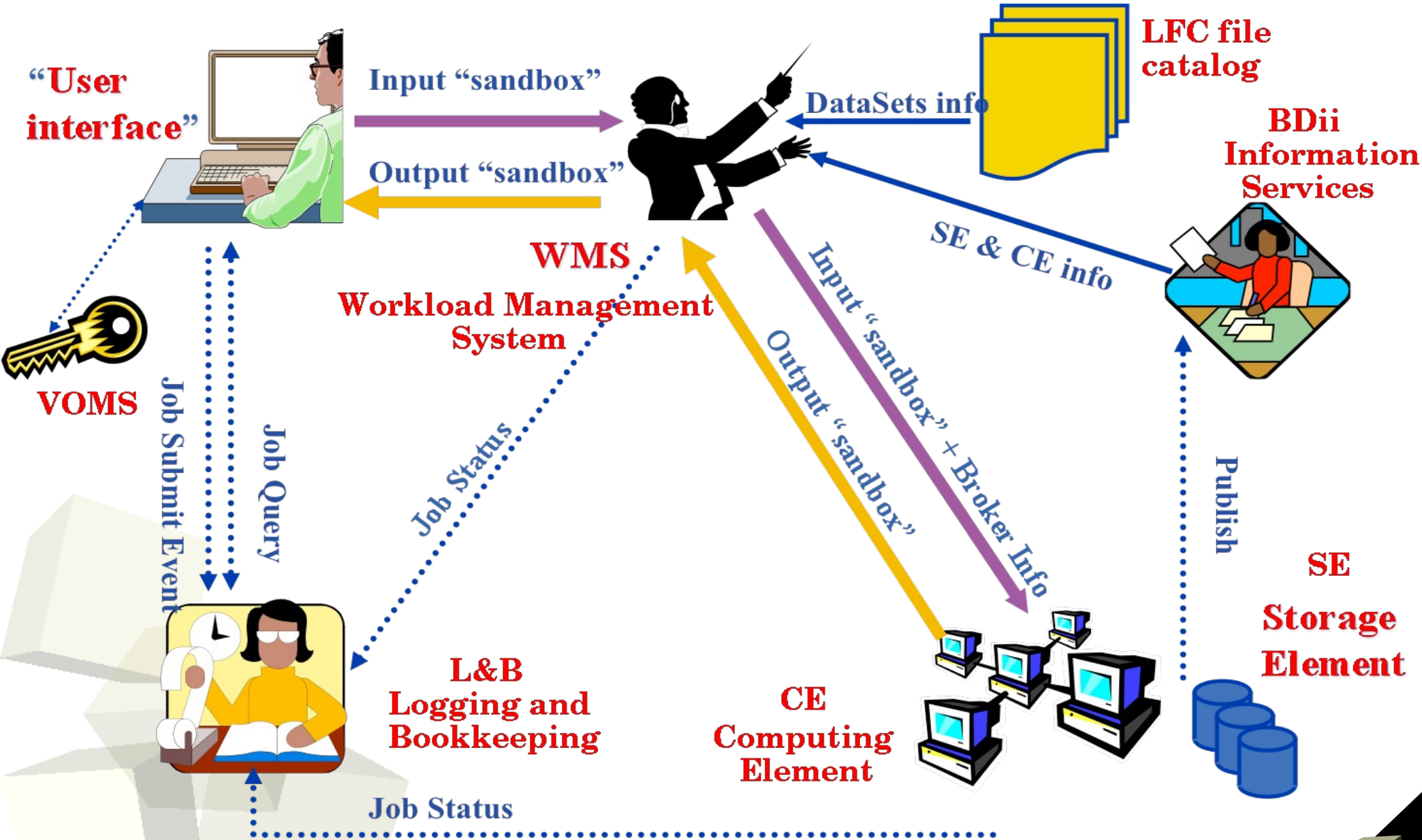
- High competition for resources; difficult to understand how to do it in a fair but efficient way
- Information Systems on sites don't provide enough and reliable information
- Uncertainty on amount of resources 'dedicated' to our VO (prioritized CPU usage by tweaking 'fairshare' queue parameters and allocated storage space on SEs)
  - ➔ 'Assigned' disk space surpassed: painful file migrations from one to another SE
- Frequent technical problems with diverse services; WMS in particular which is key in the submission process, but also VOMS and site configuration changes affect authorizations for job submission
- Loss of files; fortunately does not happen frequently
- In spite of valid ticketing system, direct contact with administrators is sometimes missed
- Downtimes are often reported once is too late to take any actions
- Grid infrastructure is like a living being and it's almost impossible to estimate times for producing given amounts of data due to big daily fluctuations (Global Grid usage, site troubles, ...)
- Amount of sites which can be accessed depends on how big a collaboration you are. Right to access resources has to be negotiated.

Summary

- **Auger** is a big collaboration in need of high processing power due to the characteristics of the events we have to simulate involving billions of particles
  - Complex software which needs site installation
  - Shower files of 100s of MB; *Offline* outputs are smaller
  - Jobs may require up to several days of computation
- **GRID technology** is mature enough to be profitted from, even by small teams … but it seems to need polishing
- We present a fairly simple solution for the **automation of production tasks** which has helped us to increase much our job production rate
- Anyway, **Grid is a complex and evolving system**; 'stability' is not guaranteed (see daily CPU time consumption ...)

BACKUP

Grid overview



Resources

- Computing Element CPUs

Total	33008
FZK	13616
GRIF	4335
Prague lcg2	2903
Nikhef	2440
Aachen	2488
Signet IJS	1162
NCG-Pt	1064
Wuppertal	928
Lecce	176
UniAndes	216
CBPF	344
UNAM	58
UFRJ	912
LIP	532
Prague Cesnet	80
Grisu-Enea	95
CNAF	1659

Mostly shared with, e.g., LHC Collaborations

Number of 'dedicated' CPUs appears in slide 7

Resources

- Dedicated resources:

Gathered by Jiri:

Site name	CPUs (jobslots)	SE Disk Space [TB]	Note
FZK	100	50	
Aachen	85	15	
Wuppertal	90	25	update from 07.2010
Lyon			resources are used locally, storage via SRB
GRIF APC (Paris)	5		
GRIF LAL (Paris)	15		
Prague FZU	25	8	local usage is included in the CPU capacity
Prague CESNET	16	1	actual CPU use is often maximum 72 jobslots
Lisabon LIP	2	0	
NIKHEF	20	0	
IJS	10	0	
UNAM	4	0	

372 CPUs 99 TBs

Grid usage: storage space

Disk space on Storage Elements since end 11/2010

The total amount of disk space consumed by the last productions is **92 TB** (all output files combined)