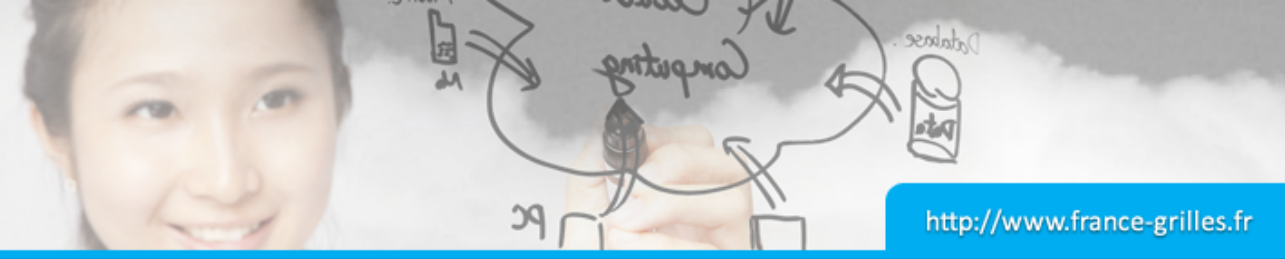
FRANCE GRILLES

Better support users

G. Romier

Relationship with user communities



http://www.france-grilles.fr

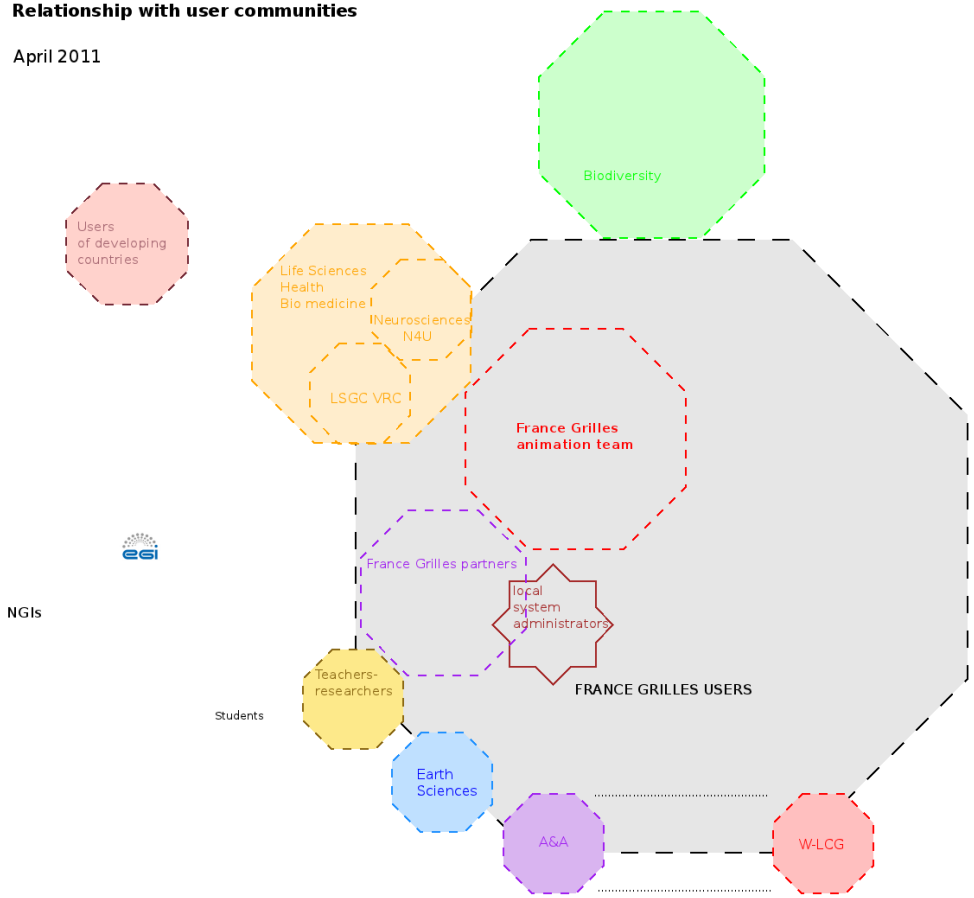
Table of content

- Introduction.
  - Where we were last year.
  - Where we are now.
  - Where we want to be.
- Where we are now and what was done this year?
  - Metrics, projects & status.
- Responses to the last IAC recommendations.
- Next steps.

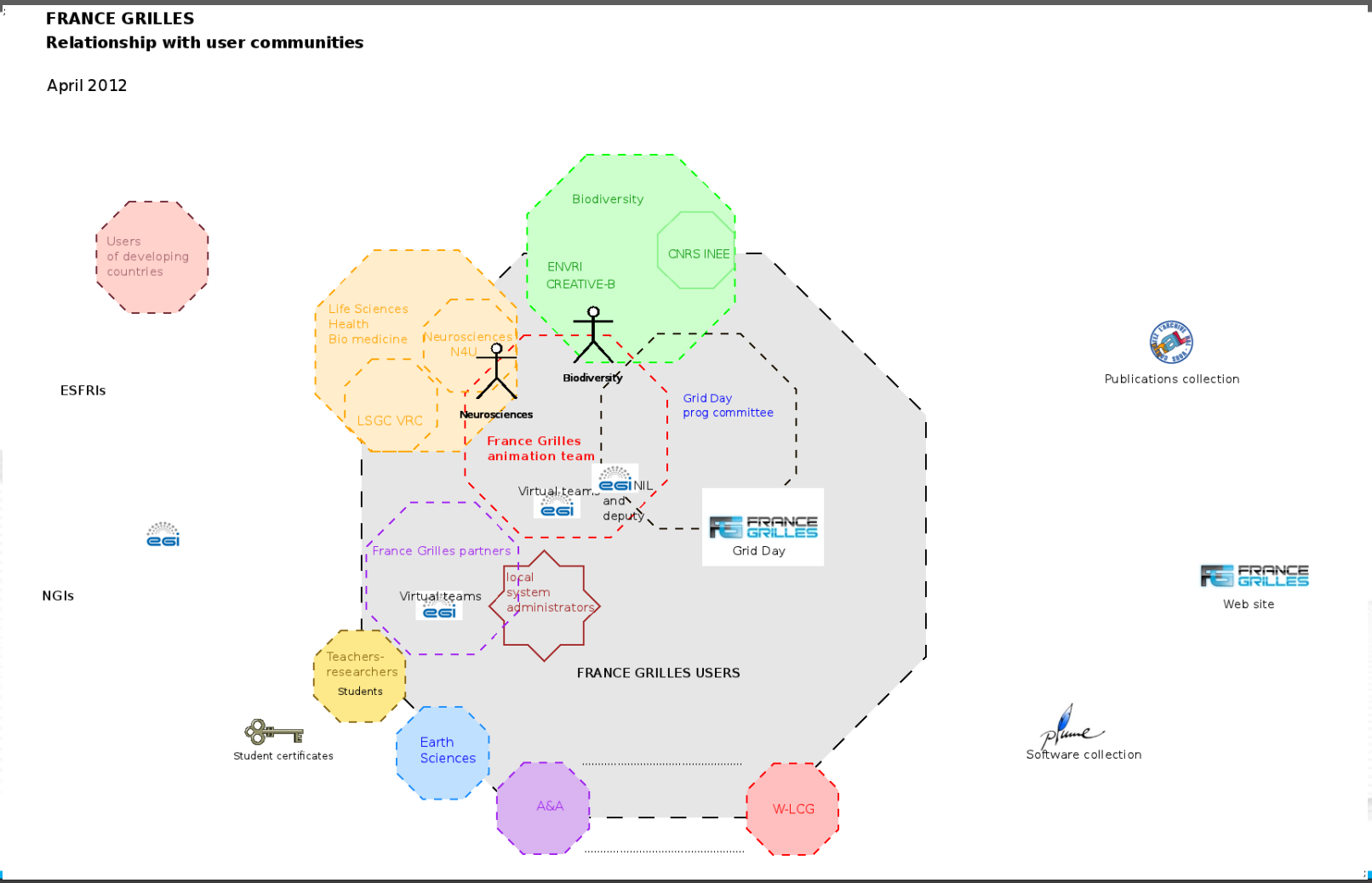
Introduction: where we were last year

FRANCE GRILLES
Relationship with user communities

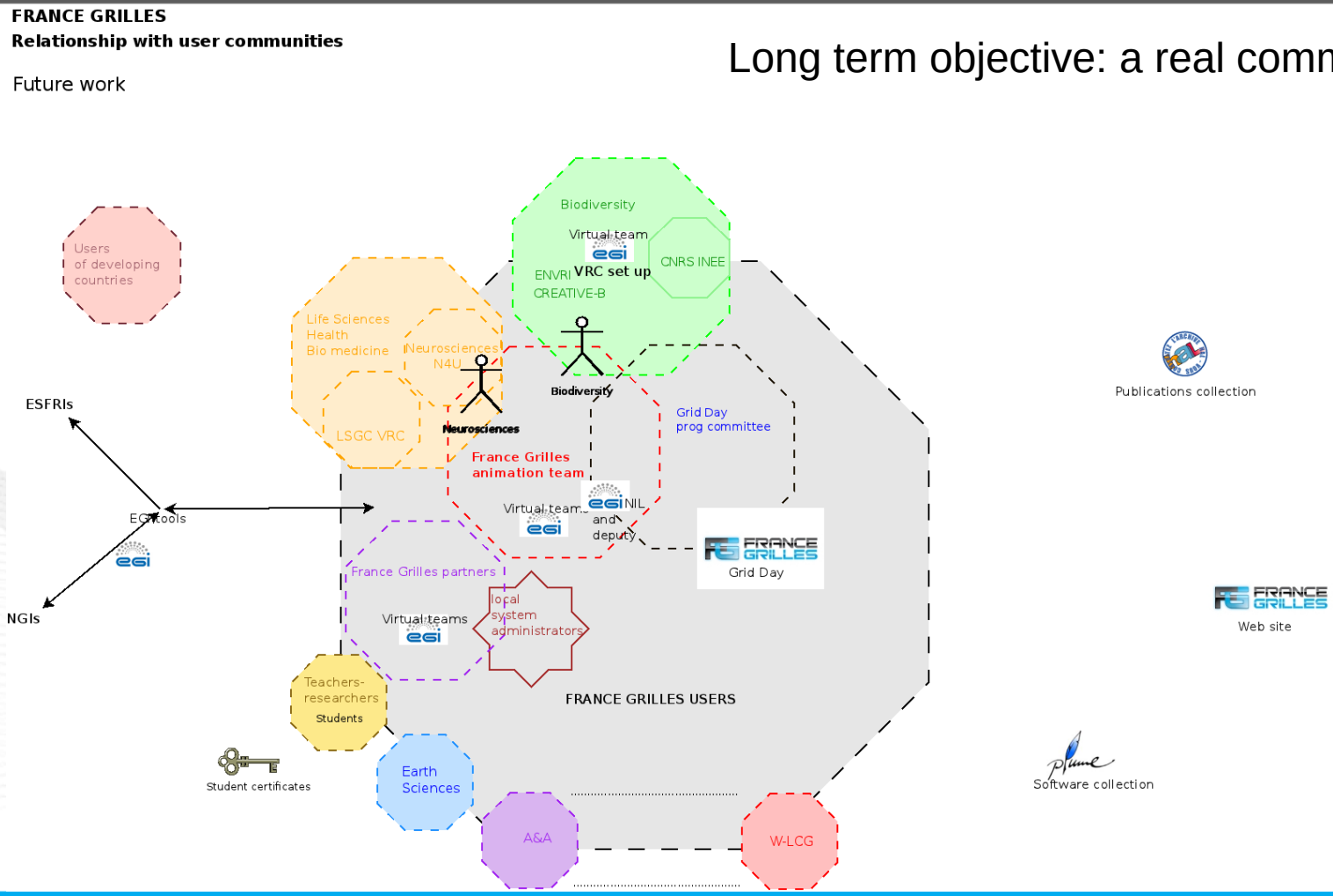
April 2011



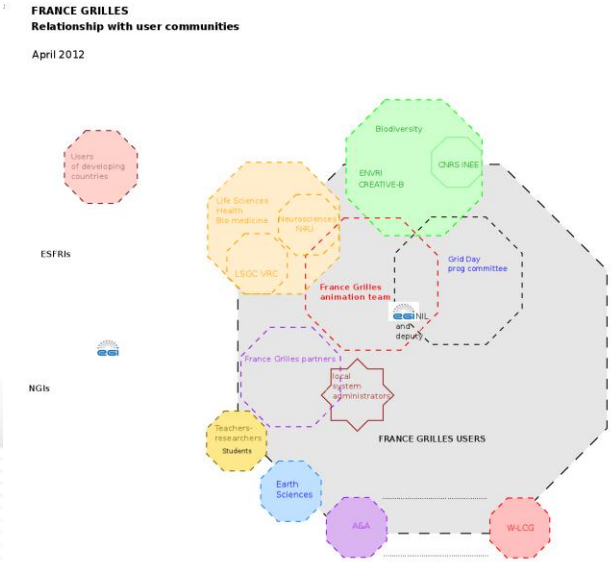
Introduction: where we are now



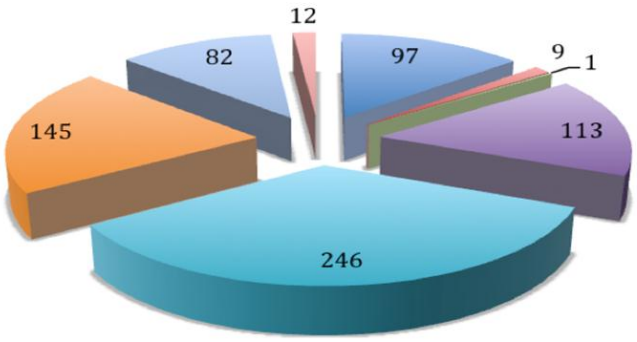
Introduction: where we want to be soon



And now? Where are we? Several metrics



- Who are our users?
- Which communities do they belong to?



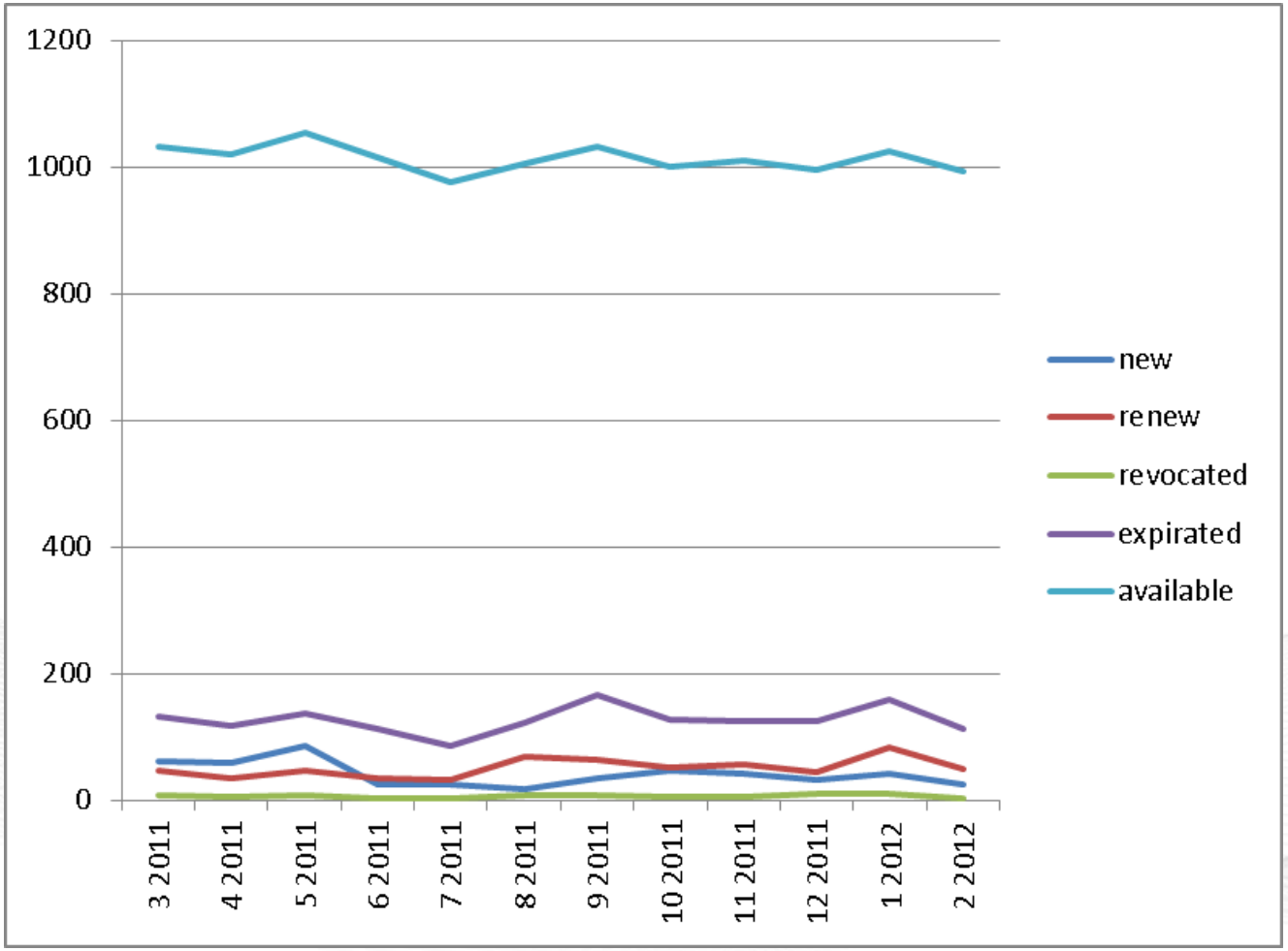
Per discipline distribution

How many national users do we have?

Number of GRID2-FR credentials

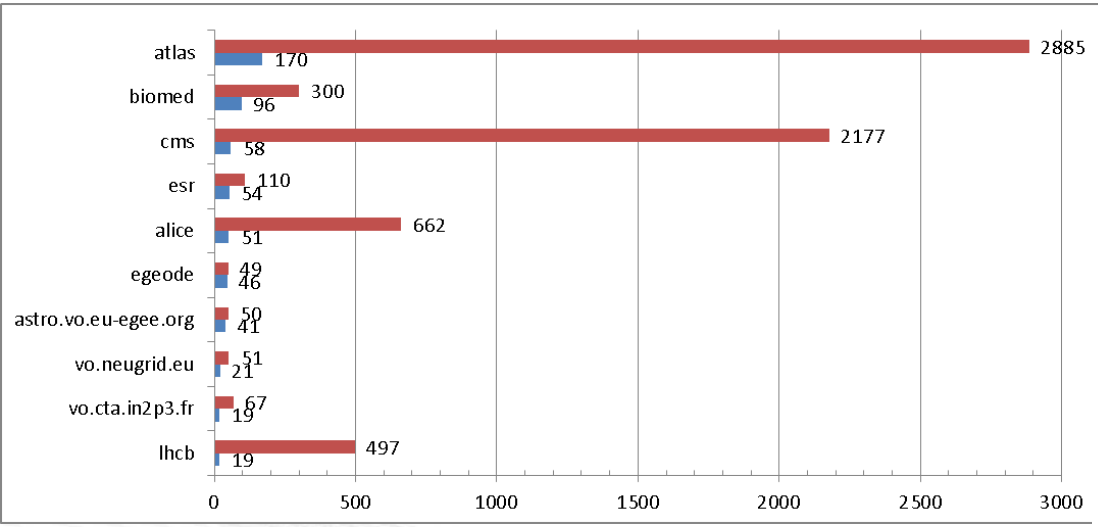
NB:

- some users use grid resources through a robot credential (e.g. 67 users use the Biomed robot).
- people may have credential without necessarily using the grid.



Where are most of them working?

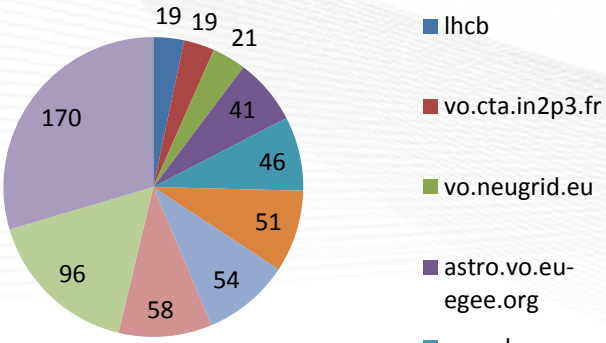
10 global VOs with most national users

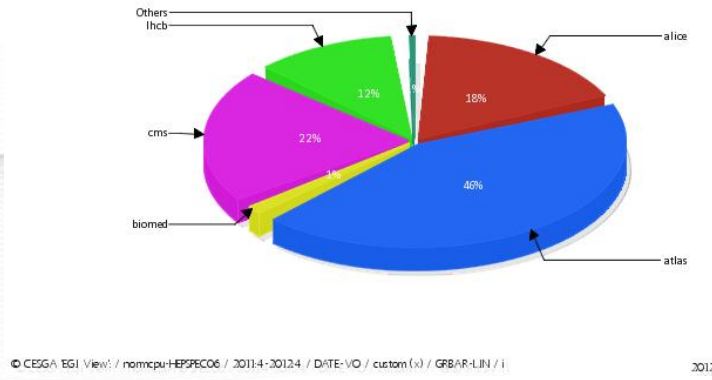


April 2012

Particle Physics: atlas, cms, lhcb
Nuclear Physics: alice
Life Sciences: biomed, neugrid
Earth Sciences: esr, egeode
Astronomy Astrophysics Astro-Particle Physics: astro, cta

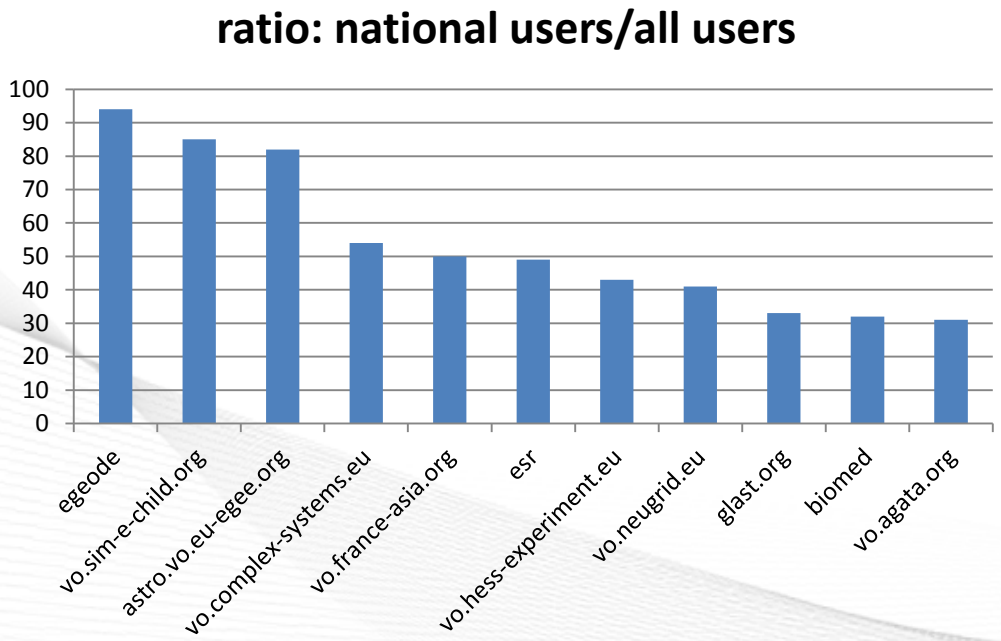
April2011 – April 2012 consumption





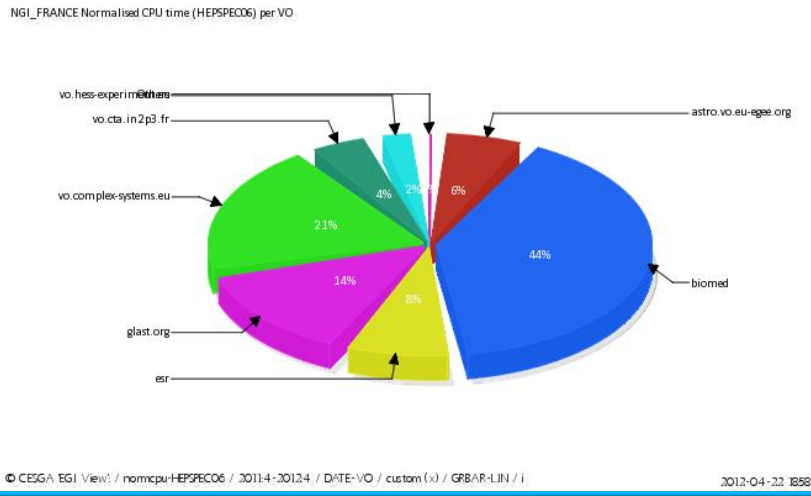
In which VOs are they proportionally most numerous?

global VOs where the ratio of national users is most important - April 2012



- **Life Sciences**: biomed, neugrid, sim-e-child
- **Earth Sciences**: esr, egeode
- **Astronomy Astrophysics Astro-Particle Physics**: agata, glast, hess, astro

April2011-April2012 consumption

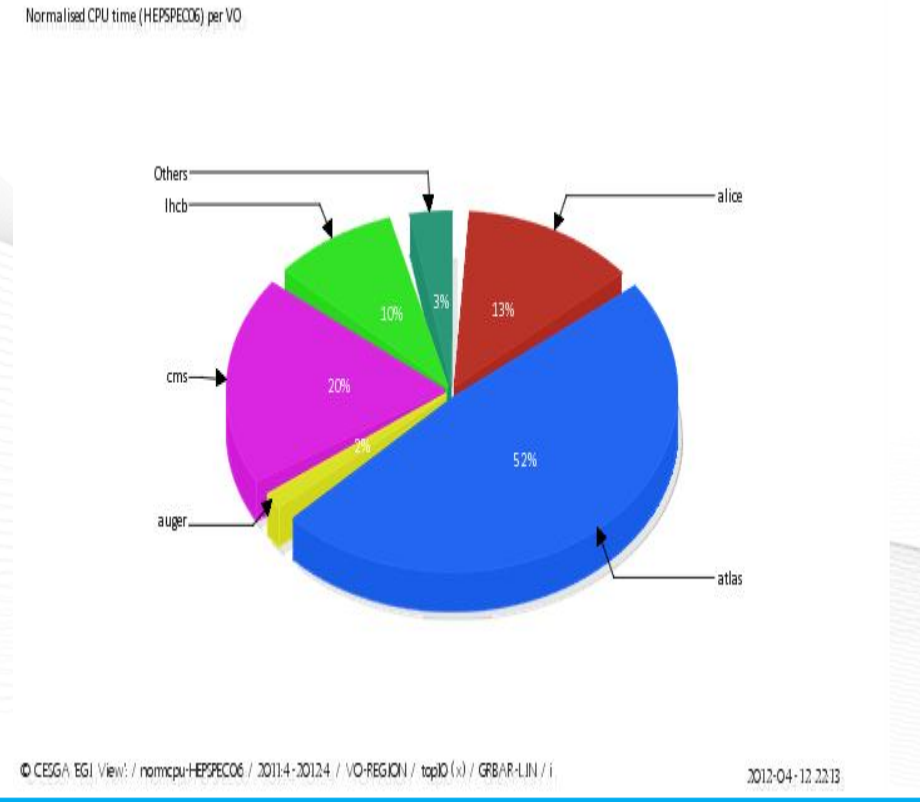


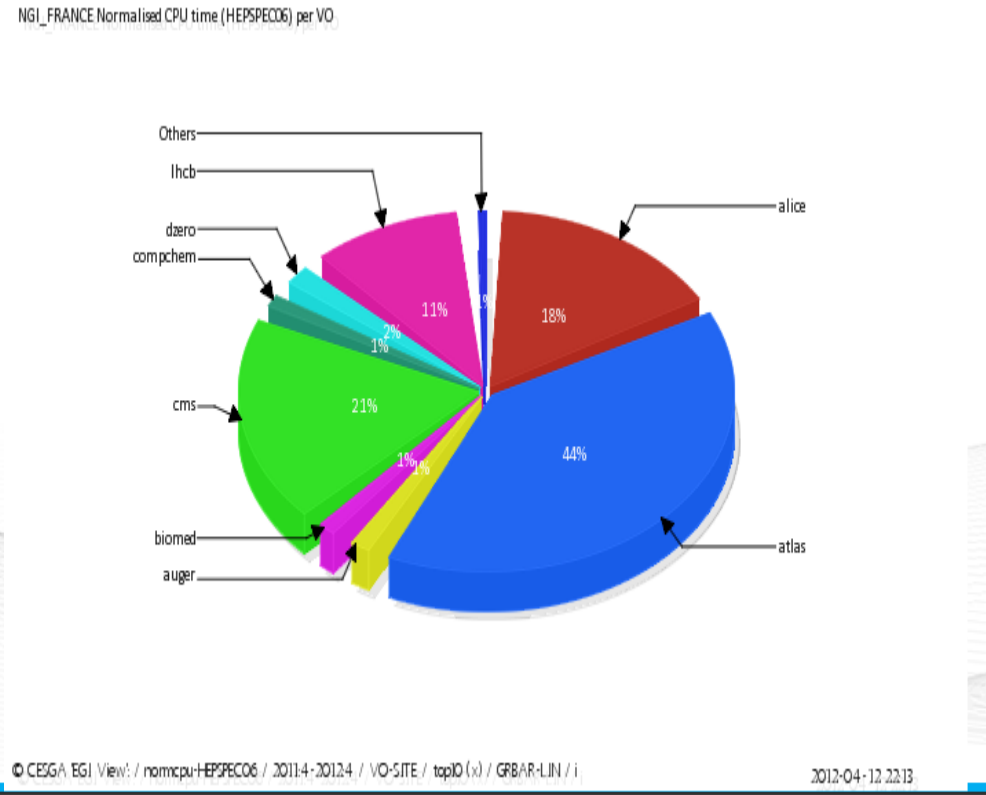
Who are the biggest consumers?
Normalised HEPSPEC06 CPU time per VO 4/2011 – 4/2012

World wide EGI VOs Top 10

NGI-France VOs top 10

Particle Physics: atlas, cms, lhcb, Dzero, **Nuclear Physics**: alice, **Life Sciences**: biomed, **Earth Sciences**: esr, **Astronomy Astrophysics Astro-Particle Physics**: auger, **Chemistry**: Compchem





Atlas and LHCb specificity

Atlas 2011, T1 and T2 sites

Jobs
user/all ~ 49.7%

CPU
user/all ~ 13.5%

WT user/all ~ 16.6%

Credit: E. Lançon

LHCb

Jobs
user/all ~ 41%

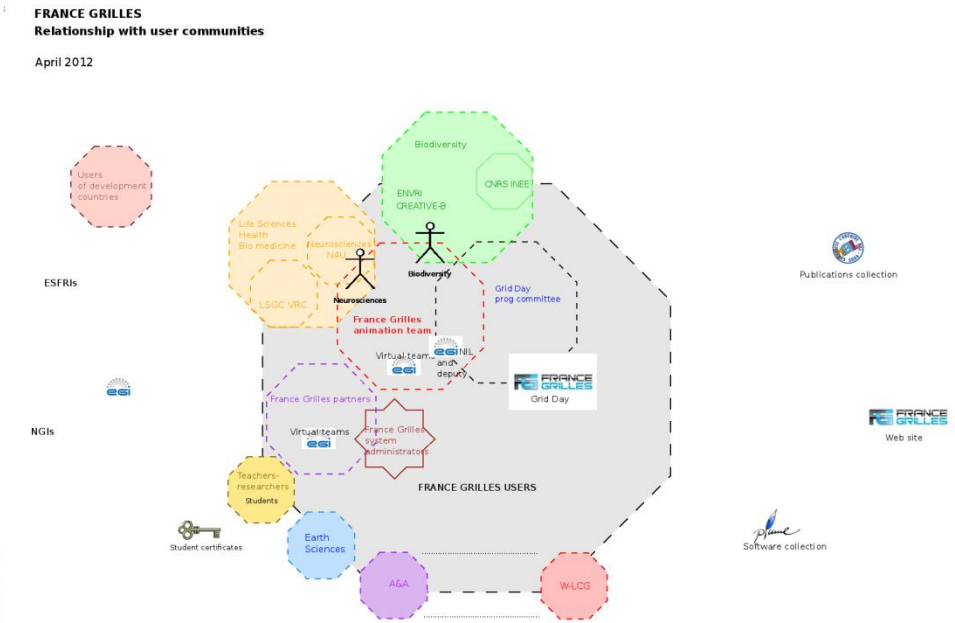
CPU
user/all ~ 10%

WT user/all ~ 13%

Credit: Andrei Tsaregorodtsev

What was done this year? activities targeted to the communities.

- International level: EGI.eu, N4U, ENVRI and Creative-B
- First French Grid Day.
- Scientific publications collection set up.
- First step of software dissemination.
- Training.
- What about students?
- How to reach new communities?
- Better welcome new users!
- Examples of new users.



International level: EGI.eu, N4U, ENVRI and Creative-B

- EGI.eu set up the NIL role and virtual teams: we do participate!
  - France Grilles is involved in several VTs:
    - Website content
    - Intelligence Collection & Analysis Process
    - Assessing the adoption of Federated Identity Providers within the EGI Community
    - Organise a high impact presence for EGI at EGU General Assembly 2012
- N4U project (NeuGrid follow up, neurosciences): an IdGC FTE is responsible for dissemination, exploitation, concertation activities, several French labs, organisms and one SME are also involved.
- Creative-B and ENVRI, biodiversity projects: an IdGC FTE is responsible for dissemination activity.

French Grid Day: Rencontres scientifiques France Grilles

- First edition September 2011, Lyon, co-located with EGI Technical Forum (that we co-organised). A success!
  - http://www.france-grilles.fr/Rencontres-scientifiques-2011
- Organisation set up:
  - program committee: (representatives of disciplines). We will build the France Grilles community around it!
  - 29 articles received / 10 oral presentations webcasted, now available from France Grilles web site / 11 posters and 5 demos on the FG booth.
  - 60 to 80 attended and 20 on-line (webcast)
  - scientific committee: awarded the price.
    - http://www.france-grilles.fr/Prix-France-Grilles-2011

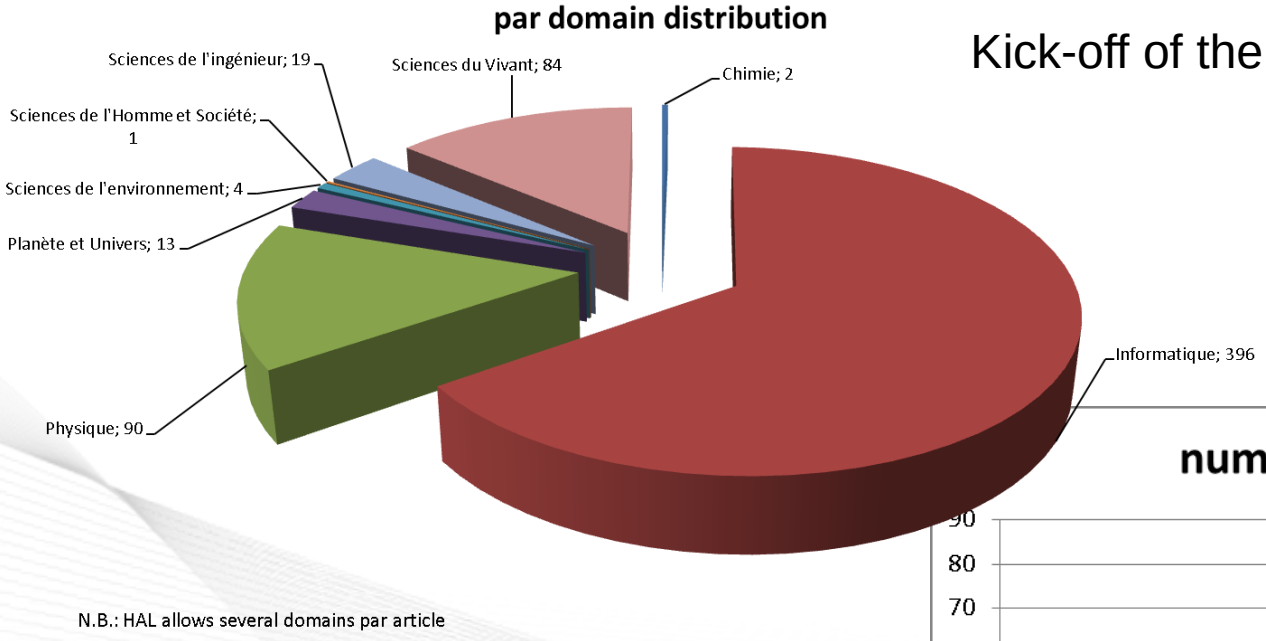
Collection of scientific publications

- France Grilles needs to
  - represent the scientific activity based on its infrastructure,
  - measure its impact,
  - justify to funding agencies the scientific use of grids for user communities,
  - establish statistics to evaluate the use of grids across time.
- Publications are the more visible result of research!
- Building a collection of publications: constraints
  - impossible to do manually,
  - authors are owners of their papers. Authors, communities and their scientific managers should be part of the project,
  - data should be publicly available for all in different ways adapted to all needs,
  - possibility of statistics is mandatory.

Scientific publications: our project

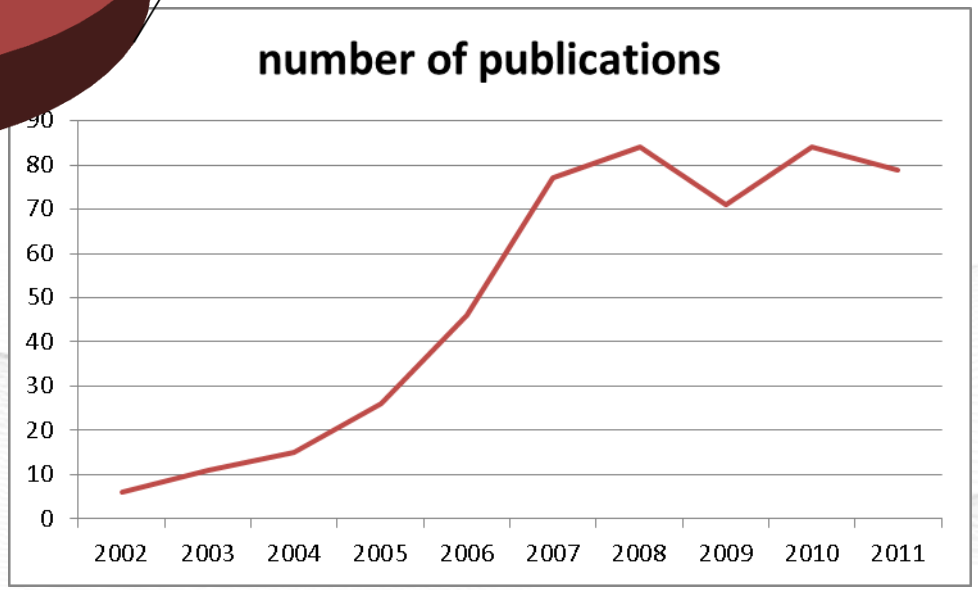
- The HAL platform is the tool build by CNRS CCSD and officially used by the France Grilles main partners to reference their publications.
- Principle: a specific tag “France-Grilles” is stamped (manually) on the publications. All tagged publications appear immediately in the collection.
- Difficulties:
  - a publication has to be in HAL to be stamped,
  - we need to inform and convince all researchers to put their publications in HAL,
  - finding the relevant publications is not easy.
- Help:
  - referencing the publications in HAL is mandatory in several organisms (as INRIA e.g.).

Scientific publications: current results, NOT representative



N.B.: HAL allows several domains par article

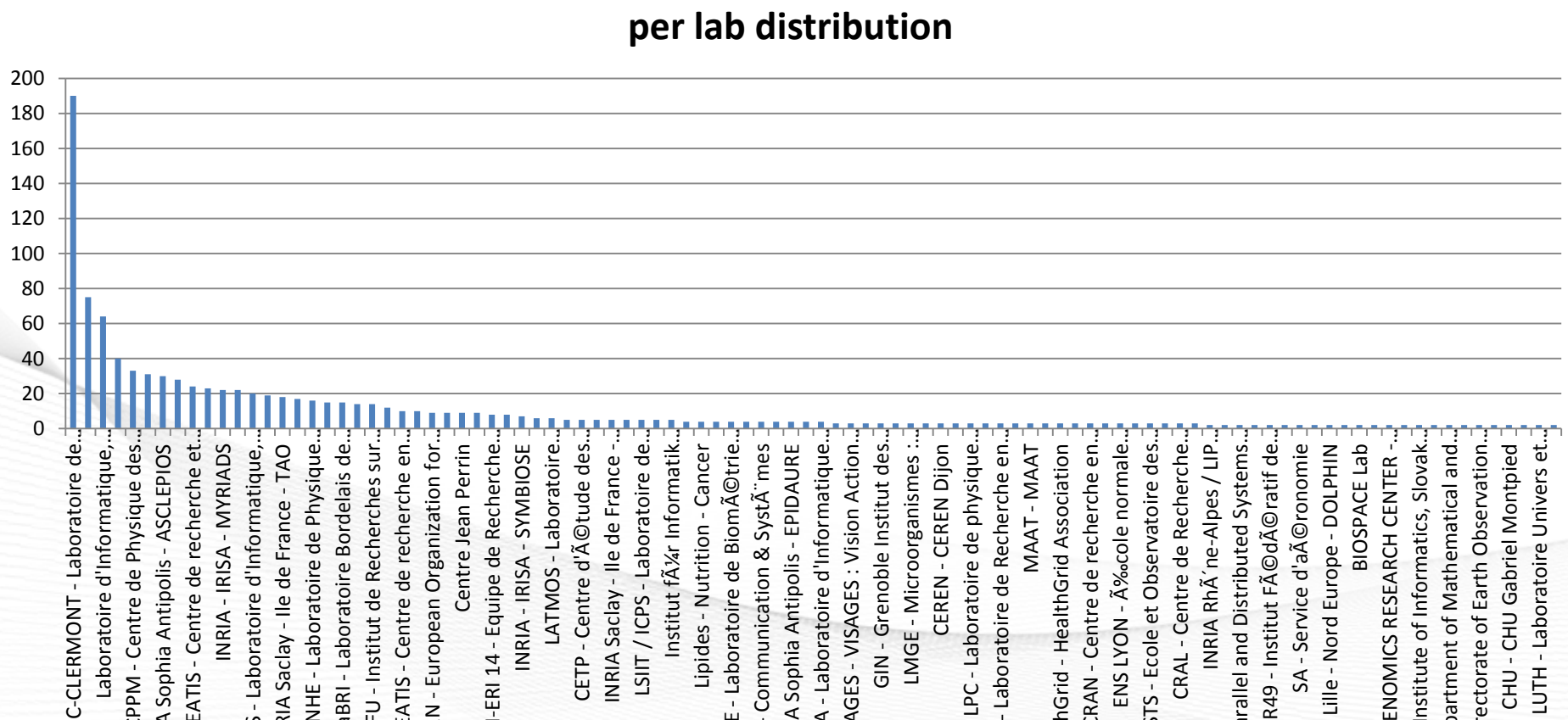
Kick-off of the collection: end of June 2011.



Evolution of the collection:

- January 250 articles and references.
- April 500 articles and references.

Scientific publications: just a beginning!



Software production

- Software production is another direct or indirect result of research.
- Software production should be of interest for other organisms, industry, potential partners...
- Publishing developed software is a way to reduce double work, to conduct software to a better level of quality and to find new collaborations...

Software production: a dissemination platform exists

- Plume platform was set up in 2007
  - now > 1000 software index cards (>80 in English)

  to **P**romote economica**L**, **U**seful and **M**aintained softwar**E** **F**or the Higher **E**ducation **A**nd **THE** **R**esearch communities.(understand FLOSS)
- Main target: all the French speaking research and higher education community

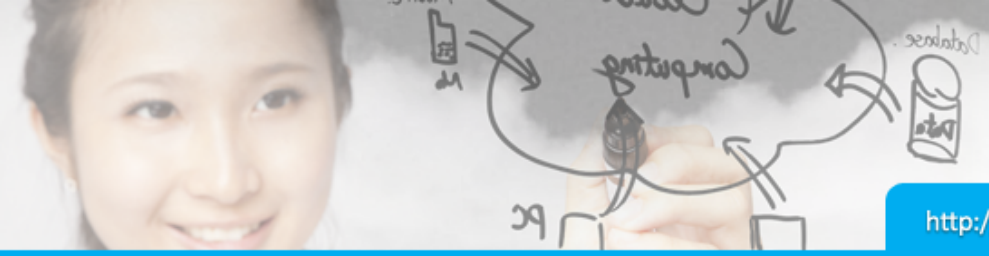
Plume is a good tool to disseminate software!

Software production: status of our project

- Grid and Cloud theme on Plume web site:
  - French “dissemination” cards published:
    - XtremWeb-HEP (LAL), OAR & CiGri (LIG), APCScheduler (APC), Paraloop (LIPM), HINTS (DSI CNRS),
  - English “dissemination” cards published,
    - XtremWeb-HEP (LAL), CiGri (LIG), Paraloop (LIPM), HINTS (DSI CNRS),
  - a few in work DIET, JSAGA, Lavoisier.
- We use also Plume to set up France Grilles catalogue of production software (see operation presentation).
  - iRODS, DIRAC, EGI Application Database French description cards e.g.
- We need to advertise and convince people to participate!

Training

- April 2010 - March 2011: 176 trainees / 18 days of training
  - of which 3 Dirac and 1 cloud.
- April 2011 – April 2012: 215 trainees / 25 days of training.
  - of which 52 trainees users and administrators / 9 days in Vietnam,
  - of which 2 cloud, 1 DIRAC, 1 Gatelab.
- Planned: 2 other tutorials in June.
- Questionnaire shows that training is important!



What about students?

- Following a few requests, France Grilles and GRID2-FR set up a procedure and a certificate category (specific OU) for academic students in the framework of a cursus.
- Almost ready!
- It will be available for the next school year.
  - work experiences within labs are considered as employment and still possible since long time.

Reaching out new users: development of our networks

- Through specific pages of our web site.
- Within user communities:
  - top-down and bottom-up at the same time when possible!
  - ESFRI and organisms with the help of EGI.eu and in collaboration with other NGIs (EGI.eu CRM will help).
- Through local administrators and professional networks:
  - as advisory people for computing in the labs.

Booths at conferences and networks contacts

- We attended JRES (French administrators conference, > 1000 attendees, many contacts).
- We are involved in the Plume technical team, in the RESINFO team, in the future database academic professional network.
- They all give us opportunities to advertise France Grilles!

Welcome new users!

- Most newcomers join existing communities.
  - They contact directly a site/VO and enter the grid!
  - We do not see them!
- Other newcomers are requesting access to the infrastructure through:
  - the FG web site (improved this year),
  - a GRID2-FR certificate request (information exchange established with GRID2-FR),
  - Word of mouth.

First steps to better welcome new users:

- Procedure set up to welcome new users in collaboration with technical direction:
  - a questionnaire to be filled in by new users
    - to describe the application, the team/lab, their needs.
  - a simple workflow to address the question: what is the best VO, how to give them support?
  - Discussion is internally recorded.
  - The procedure is regularly improved.

Better welcome users: our national VO.

- How to welcome new users if their discipline is not currently organised to work on the grid?
- Operations set up a national multidisciplinary VO vo.france-grilles.fr.
  - Dirac will be soon available.
- First users just welcomed
  - solid state physics and biodiversity.
- BUT we do not have correct user support now!

New users: emergent scientific communities on the grid

- Biodiversity, ecology:
  - CEFE laboratory (UMR 5175) - important biodiversity laboratory,
  - INRA.
- Solid state physics:
  - LPTM.
- The BEDOFIH project (European Database of High Frequency Financial Data):
  - Eurofidai.

Biodiversity, the CEFE application example:

- CEFE laboratory (UMR 5175) - important biodiversity lab - EEE group : Evolutionary Epidemiology application
  - Evolution of pathogen virulence and host resistance
  - What is the impact of the spatial (or social) structure of populations on the evolution of disease severity?
  - Current project : evolution of host resistance in a bacteria-virus interaction.
    - can bacteria resist viral infection by committing suicide?
    - lab experiments + models + stochastic simulations
    - Grid used to compute probability that the resistant strain "wins“
    - Dirac was used to simplify job and data handling

Solid state physics

- A team fromLPTM (UMR8089 CNRS - Université de Cergy-Pontoise) laboratory will port a first solid state physics application to the grid
- They attended the last training session
- They just enter the national VO as first “real” users
- GRIF will support them

BEDOFIH project

- Initiated by EUROFIDAI (European Institute of Financial Data), to develop a European database of high frequency data.
- selected in the scope of future investments
- Such a database will allow France to develop innovative research in the current field of economics affected by the financial crisis.
- It collects and deals with several TeraOctets of monthly, financial information
- relies on the operational structures of France-Grilles and LPSC in Grenoble for this.

Scientific communities in France

- 89 VOs of which 27 national or local VOs organized by theme or around the sites and the labs.
- International well established scientific communities (W-LCG, VRCs such as LSGC, Earth Science VRC) have representatives in France.
- W-LCG: In 2011, the LHC part is 88 % of the French NGI CPU usage and the French infrastructure contribution represents 10% of the global consumption.
- France Grilles manages the LSGC (both scientific & technical mgt) and provides most human resources.
  - Perspectives
    - Improve 200+ users to 100+ sites liaison
    - Manage community data (scattered in 100s experiments)

Scientific communities in France

- Earth Sciences in France suffers lack of human resources to support its users. The French community could be a bigger user if there were several persons close to the researchers in charge of the gridification of the applications.
- User demands for using the grid in the complex-systems community are increasing, however the computational resources accessible to the VO are under-sized. Complex-systems are presently in process with France Grilles to get a support to increase the VO capacities, waiting for new investment.
- RENABI/GRISBI, a bioinformatics thematic French community,
- Grid Observatory and Green computing Observatory

Scientific communities: the local Vos around sites and labs

- Often related to laboratories or regional projects or dedicated to specific needs
  - 2 Astronomy, Astrophysics and Astro-Particle Physics
  - 1 Computer Science and Mathematics
  - 12 High-Energy Physics
  - 9 regional more or less multidisciplinary (mainly collaboration of institutes)
  - 1 training
  - Our national VO is one of them.

Responses to 2011 recommendations

- 8. Identify two non-HEP communities to each be supported by a dedicated full-time member of the IdG team.
- 9. Quantitatively analyse the current allocation of resources for user support and training.
- 10. Collect users' opinions on a regular basis.

8. Identify two non-HEP communities to each be supported by a dedicated full-time member of the IdG team.

- Two FTE were hired by IdGC in 2011.
- One is in charge of Biodiversity and responsible for ENVRI and Creative-B dissemination.
  - These projects will help to structure the biodiversity community.
  - He will be the leader of an EGI.eu VT which aim is to help to structure a Grid community through a VRC in collaboration with several other NGIs such as Spain or UK.
- One is in charge of neurosciences and responsible for N4U dissemination.
- They both take part in France Grilles work.

9. Quantitatively analyse the current allocation of resources for user support and training.

- User support and training organisation:
  - Since May 2011 a FTE is dedicated to the relationship with user communities.
  - Since 2008 a person is in charge of the organisation of training (partial time until September).
- Infrastructure:
  - France Grilles maintains a specific VO and grid resources for training. Two persons manage the VO, one is in charge of the credentials delivery .
- There is currently in France Grilles no allocated human resources for user support or for trainers. Sites or regional grids may offer support and/or training to their users
- Trainers:
  - The budget allocated is sufficient to organise the training sessions but not to pay trainers at market price level.
  - The administrative rules are complicated and we could not easily pay internal trainers.
  - Trainers work on a volunteer basis and specific training sessions are depending on the persons concerned. It can be very difficult!

9. User support

- Issue: France Grilles currently does not offer support to users, it is a barrier to the DCI adoption, people vanish after bad user experience.
  => We need to organise the user support!
  - NB: subject is user support and not operational support that exists.
- Context: Users use freely the infrastructure and the services, they are not required to participate even if they do not share resources.
- Users are consumers without paying! They do not care about their tool! We think they have to be concerned. We need to empower them!

9. User support as a common interest network?

- Question: could we set up a user support on a volunteer basis involving our users such as in FLOSS projects?
- Basis of this model:
  - involvement based on the common interest of the users for the success of the project, based on the share of knowledge and competences without immediate or direct counterpart
  - free engagement, no constraints, no payment but clear rules of participation
  - large part of contributions at best effort level - few people more active
  - organisation, tools, events and roadmap provided by the project

9. User support: what do we need?

- Short term:
  - seriously prepare the project and its dissemination before to launch it (discussions with users, how to interface it with operational support, possibility of prototyping...),
  - convince a volunteers' team to kick-off the project,
  - identify and set up an adapted user friendly collaborative tool ,
  - set up good documentation and FAQ (to be improved).
- Long term:
  - be attractive enough to gain and maintain a critical volume of contributions and effective support to the users.

10. Collect users opinions on a regular basis: questionnaire to our users

- Goal: to know the opinion of our users about France Grilles
- Constraints: no mean to address simply all our users, not bother researchers with a long questionnaire.
- Method:
  - ask a few contacts in the main communities to indicate 2 or 3 users (happy and less happy).
  - send the users a simple questionnaire
- The questions:
  - Who are you?
  - What was your interest for the grid? Did you get a tutorial?
  - Did the grid fulfill your needs? Main benefit? Did you get useful support?
  - Your proposal to improve the grid and the “user experience”? Do you use other infrastructure? (cloud, supercomputer...)
- To do: publish results on the web site and send them to concerned users.

10. Results of the questionnaire: lessons learnt

- Poor number of answers and difficult to obtain => limited interest
- half researchers/half engineers,
- all needed big resources (CPU, big data, lot of jobs),
- training is important, all found close support (lab or site),
- 3/7 found issues (<1/2 reliability or 1/7 data transfer problems),
  - know more about users experience => next survey goal
  - we have to improve users experience!
- 2/7 are/have been also HPC users => next Grid day will give more information
- 1/7 think are interested by cloud computing for the future.

Future work: new projects

- 2012 French Grid Day.
- New VTs proposed by user communities (biodiversity, data management by Biomed).
- A survey focused on user experience.
- User support organization.
- Ensure resources to communities.
- And of course continue all “in work” projects.

French Grid Day: Rencontres scientifiques France Grilles

- Second edition: Paris, 1st-3rd October 2012
- Co-organized with the HPC users meeting
  - same dates, same location.
- Reduce cost for participation! Limited travel budget this year!

How to ensure resources to the new communities? a policy

- New users/communities enter the grid mostly without having/sharing resources
- We need to set up clear rules:
  - France Grilles infrastructure is a multidisciplinary shared tool
  - France Grilles provides its users with a core set of services
  - users may bring resources and share them, large users are required to bring and share resources
  - operation people take care of the infrastructure at a high level of quality but they can not take care of all users
  - all users are invited to participate to the user support effort

Challenges and risks

- Main challenges of the last year:
  - Initiate projects mainly to disseminate “what” research is done with the help of the grid & first thoughts about a community.
- Main concrete challenges for next year:
  - Improve the representativeness of the data, show results,
  - first real steps to build a community of interest, users support kick-off.
- Main real challenge:
  - attract, convince, involve and keep people!
- Main risks:
  - Limited adherence to the projects from users, teams, labs, partners,
  - lack of trainers, difficulties to organise new sessions,
  - too many difficulties for new users/teams/communities (user experiences and users support).

Conclusion: priorities for next year

- Take active part in international activities (biodiversity, neurosciences, EGI.eu VTs and NIL job).
- 2nd Grid Day with HPC : must be a success!
- Get major progress in publications collection (number of labs or institute) => better show France Grilles scientific impact.
- Maintain level of users training despite difficulties.
- Better know users issues => new survey.
- Study, organise and kick-off collaborative user support.
- All actions to be able to better welcome new users!

- Thanks for your attention!

Scientific communities: the international well established

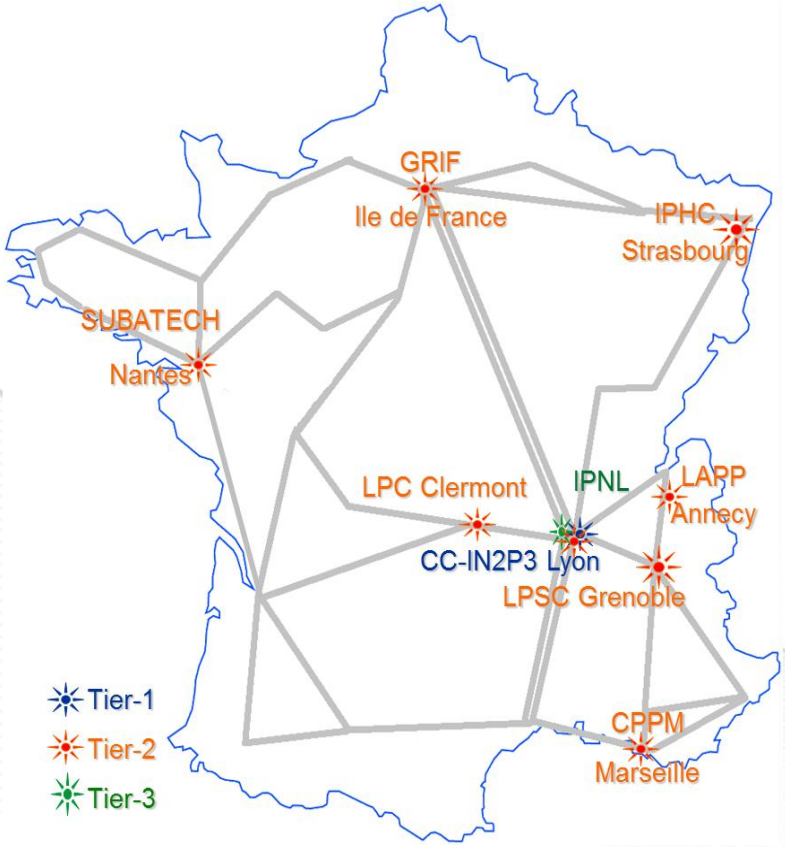
- LCG-France
- Life Sciences Grid Community (LSGC)
- Earth Sciences
- A&A
- Complex systems

Scientific communities: the international well established

- LGC-France: French part of the LCG community

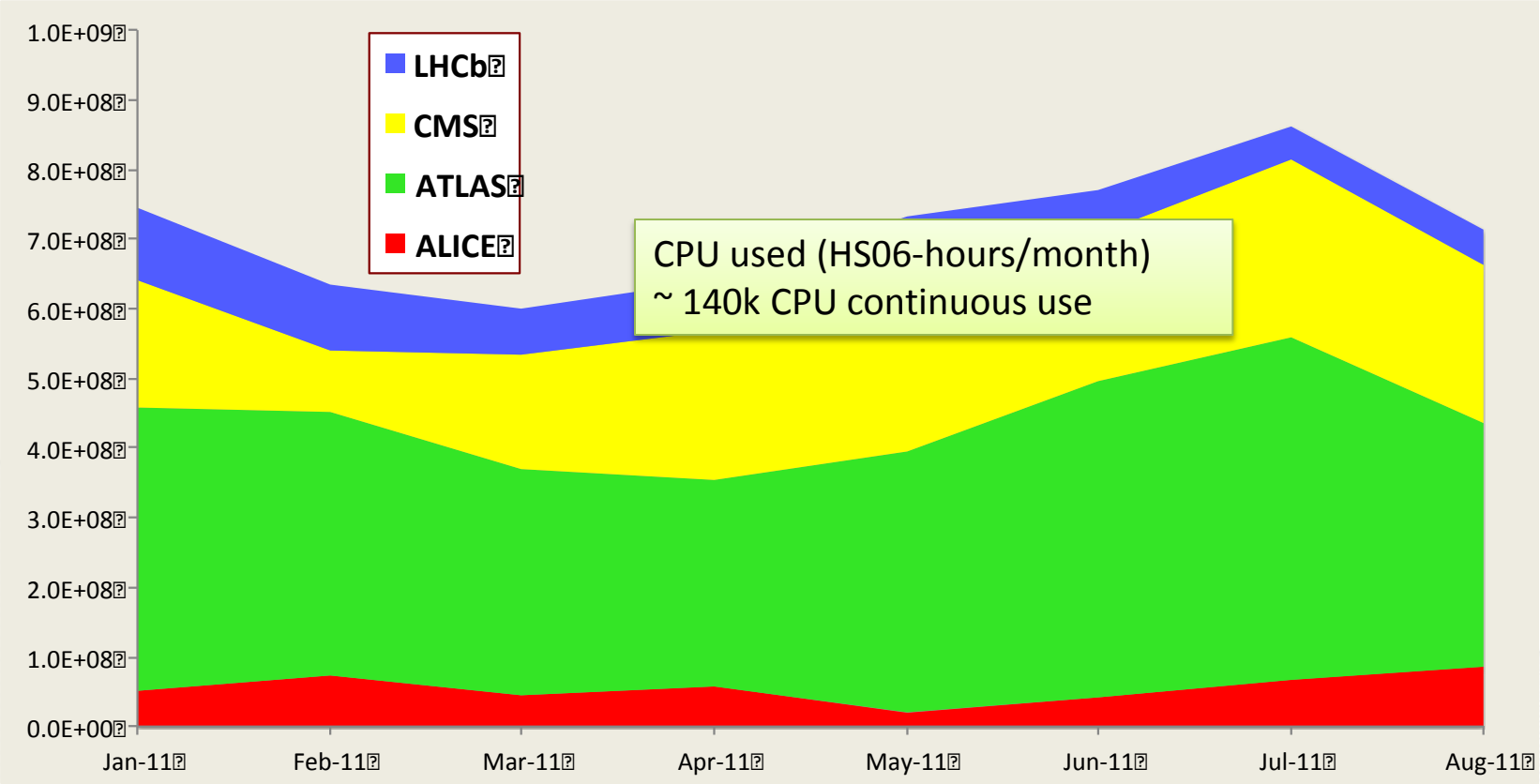
slides: Fairouz Malek and Frédérique Cholet

LCG-France Sites



Role	Site	ALICE	ATLAS	CMS	LHCb
Tier-1	IN2P3-CC	✓	✓	✓	✓
Tier-2	IN2P3-CC-T2 (AF)	✓	✓	✓	✓
	IN2P3-CPPM		✓		✓
	GRIF	✓	✓	✓	✓
	IN2P3-LPC	✓	✓		✓
	IN2P3-IPHC	✓		✓	
	IN2P3-LAPP		✓		✓
	IN2P3-LPSC	✓	✓		
	IN2P3-SUBATECH	✓			
Tier-3	IN2P3-IPNL	✓		✓	

WLCG CPU usage

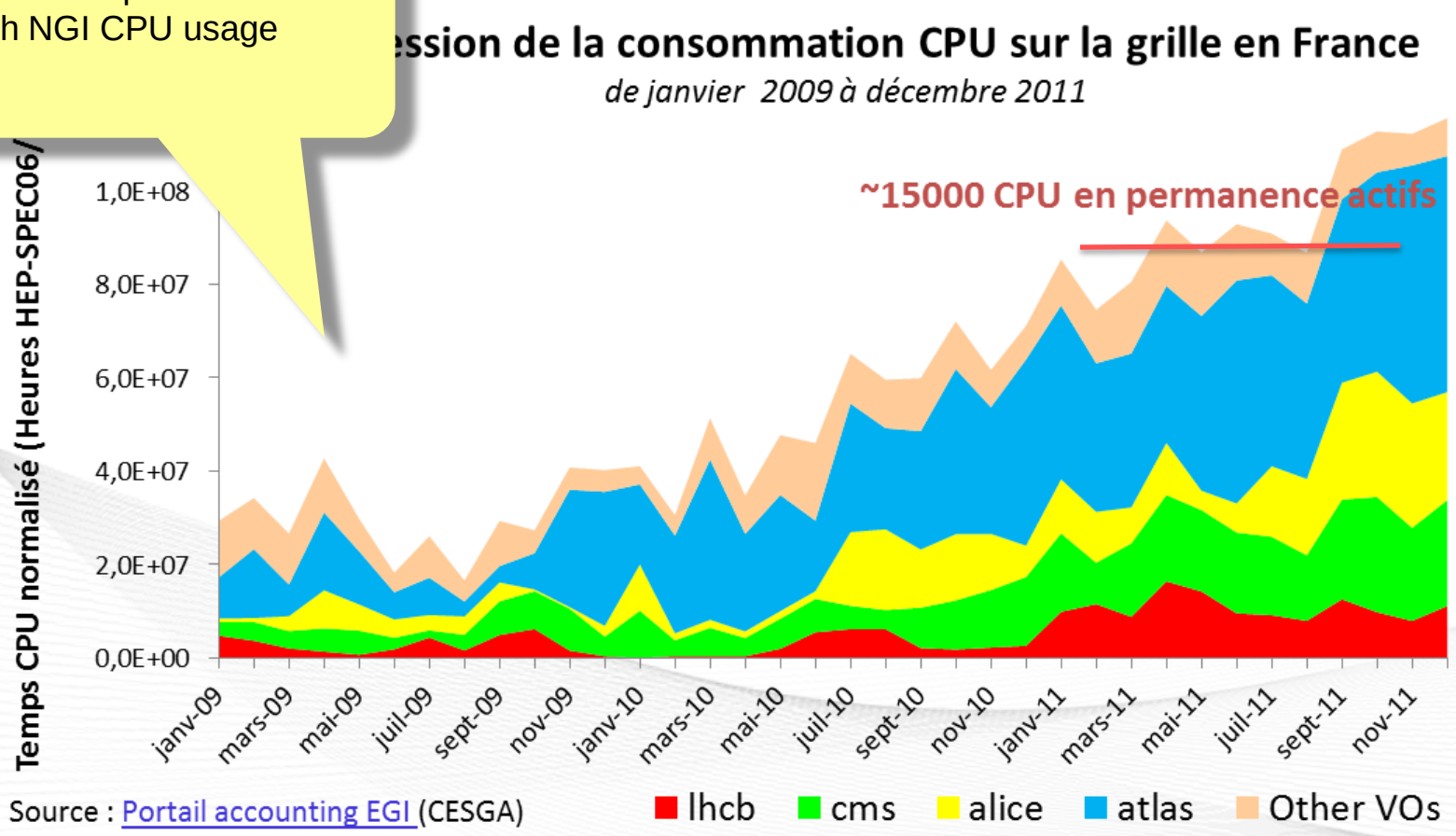


from WLCG project status report , CRRB 18/10/11 (Ian Bird)

F.Malek/F.Chollet
53

Grid CPU usage in France

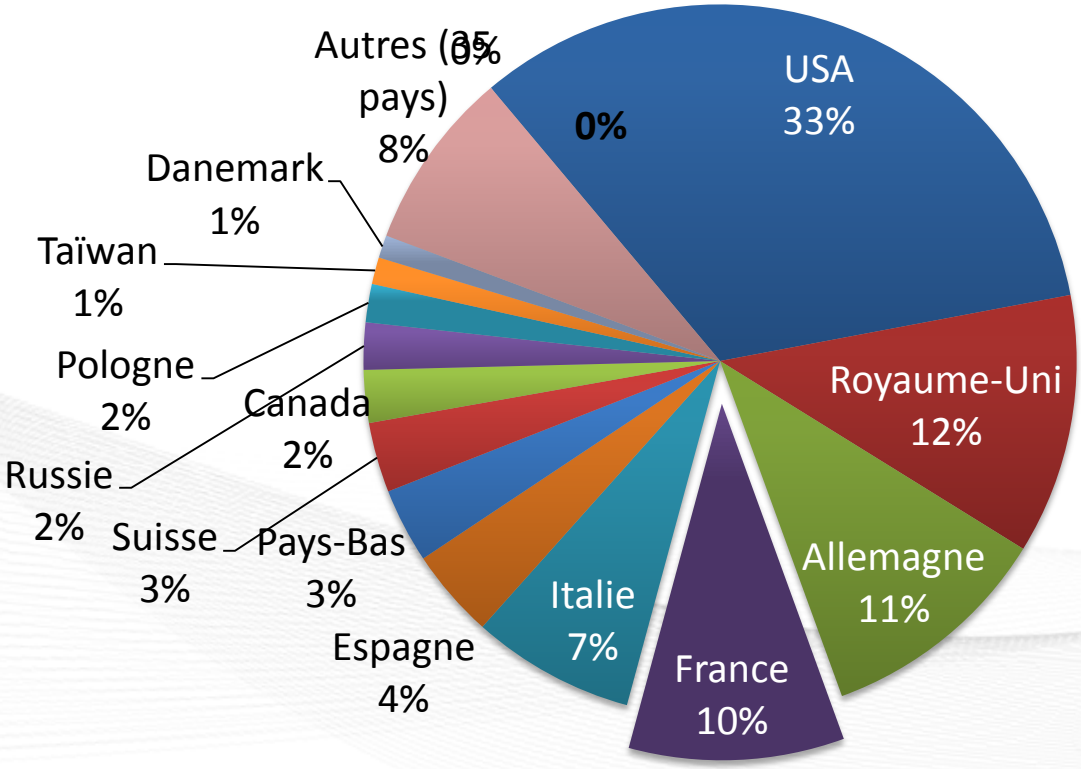
In 2011, the LHC part is 88 % of the French NGI CPU usage



F.Malek/F.Chollet
54

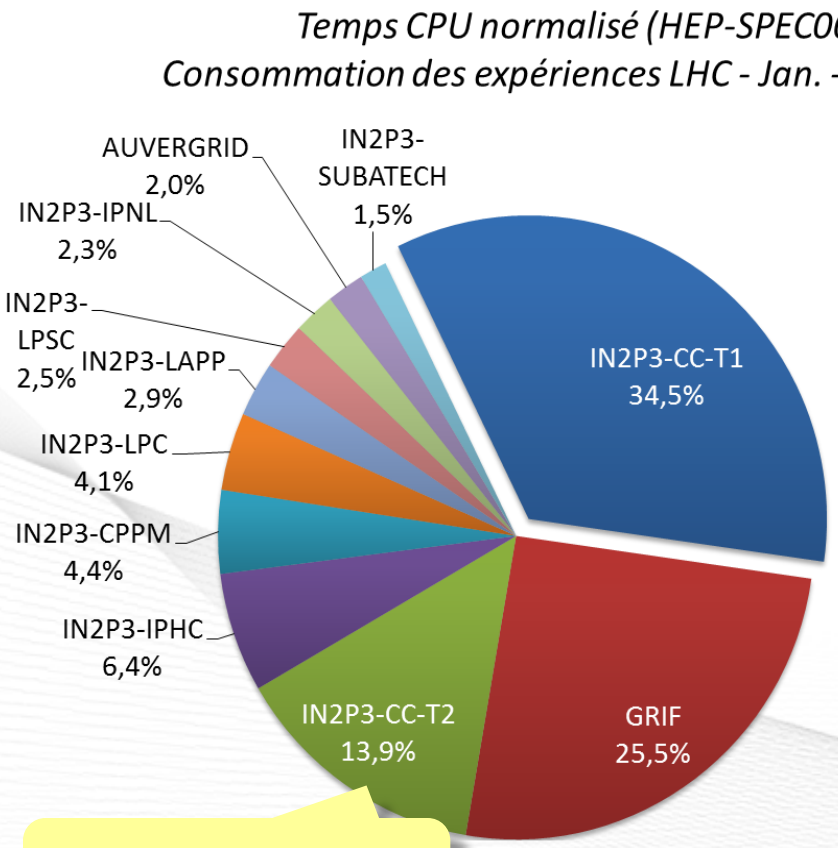
LHC CPU in France
Contribution per country

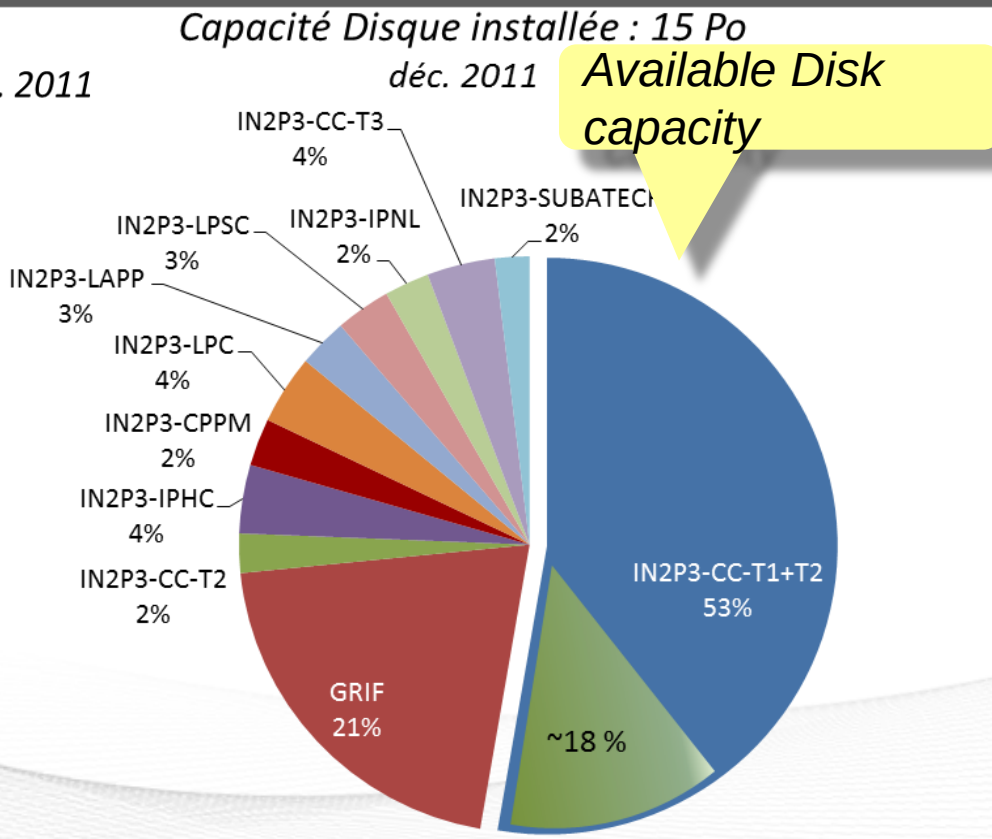
Temps CPU normalisé (HEP-SPEC06)
Consommation des 4 expériences LHC - Janv. -Déc. 2011



F.Malek/F.Chollet
55

C resources usage in France
Contribution per site



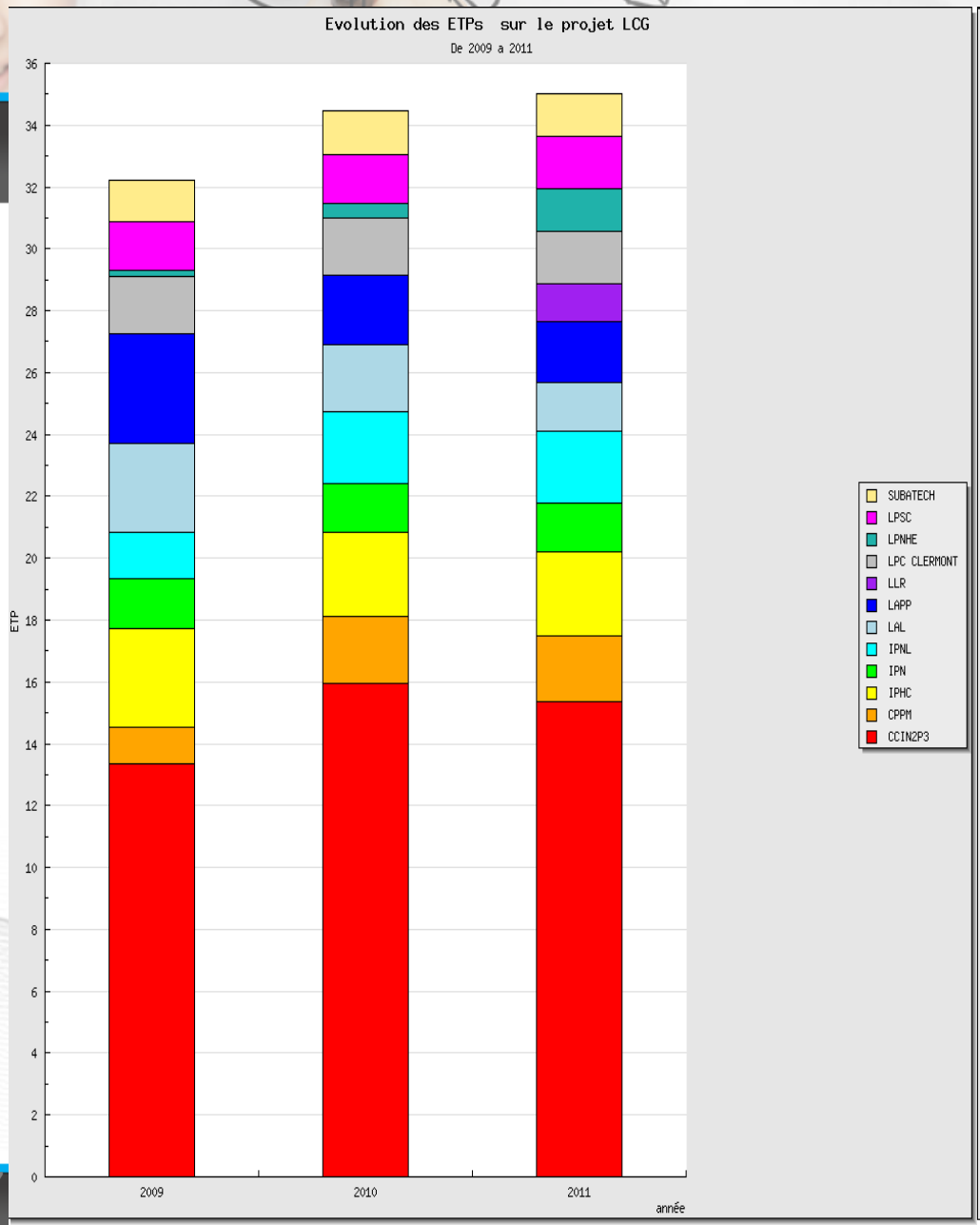


Sources : http://gstat.egi.eu/gstat/summary/EGI_NGI/NGI_FRANCE/
https://grid.in2p3.fr/LCGFrAccounting/

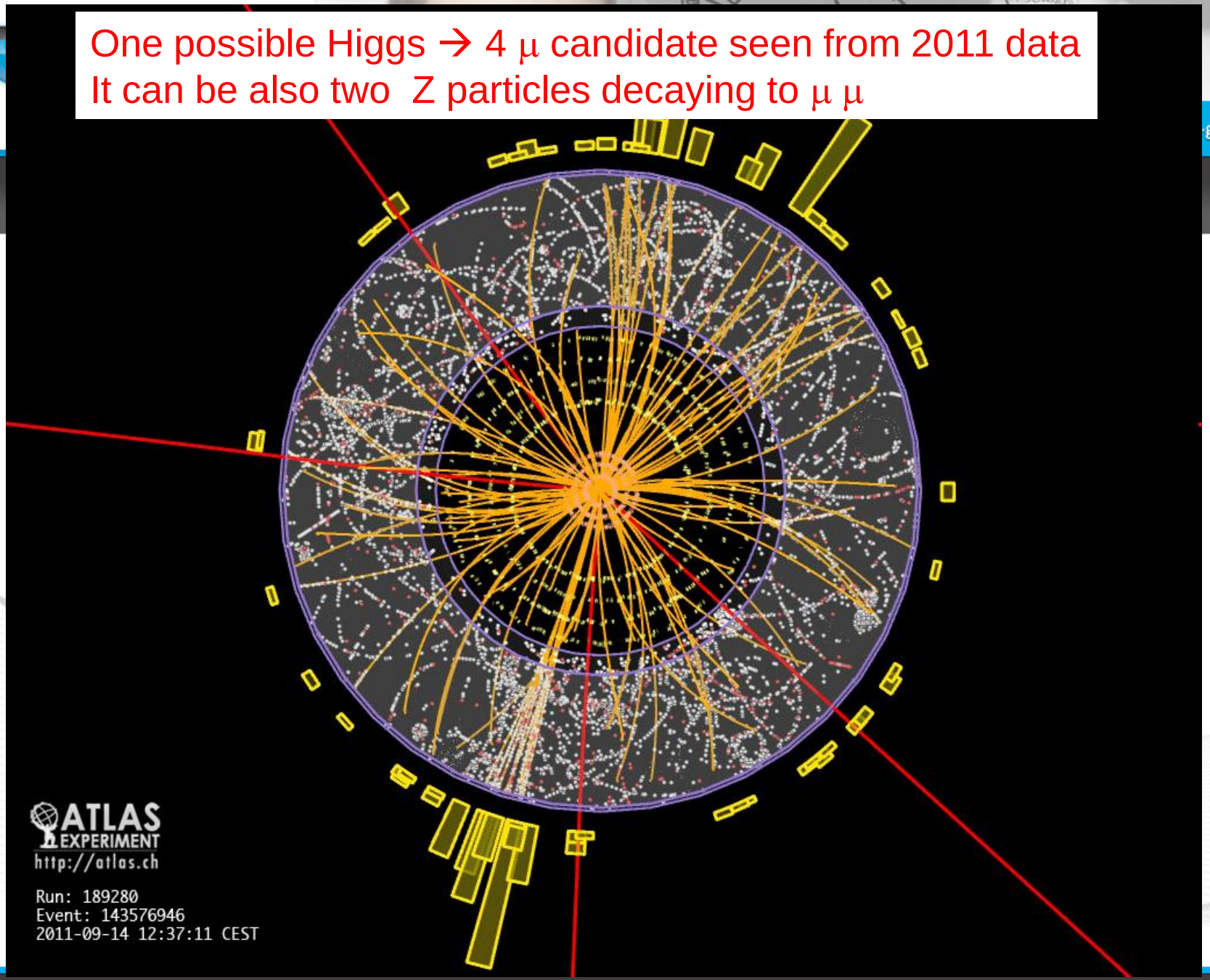
F.Malek/F.Chollet
56

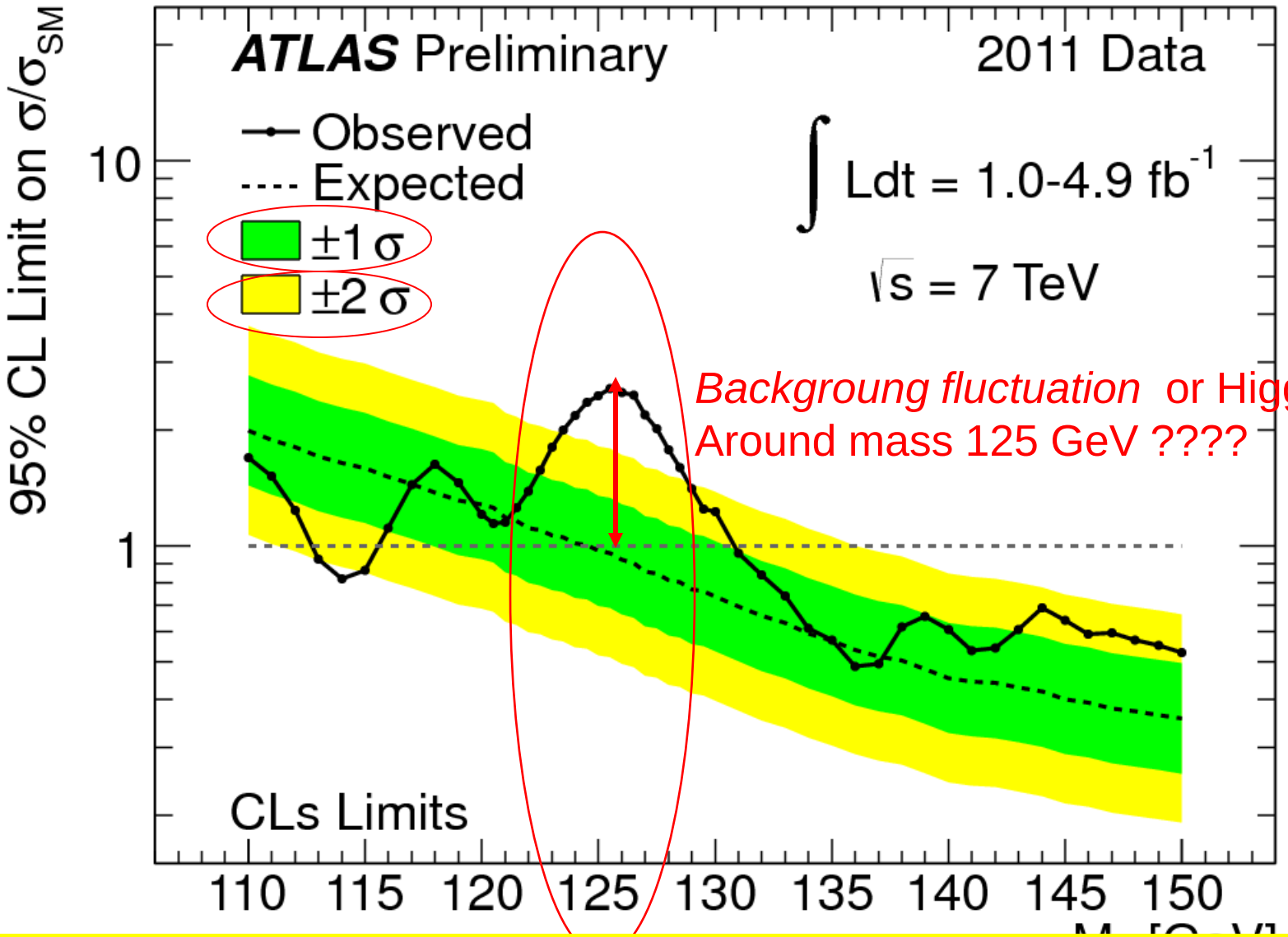
RH IN2P3 HR resources extracted from ISIS DB

In 2011 : ~34 ETPs were declared by IN2P3 Labs (where 40 % is by CCin2p3).



One possible Higgs → 4 μ candidate seen from 2011 data
It can be also two Z particles decaying to μ μ





The answer to that question is possible only if the measurement is done with a enough statistics and precision→ (> 3 σ) → results not before end of 2012

Scientific Results and Publications

- Each of the 4 LHC experiments has already published more than 200 scientific articles since 2010.

- Results on the existence of the Higgs at mass 125 GeV or the discovery of any other nature will not be possible until enough statistics and precise measurements are done. We believe some answers can be provided at the end of year 2012.

Scientific communities: the international well established

- Life Sciences Grid Community (LSGC): a community where national users are particularly active

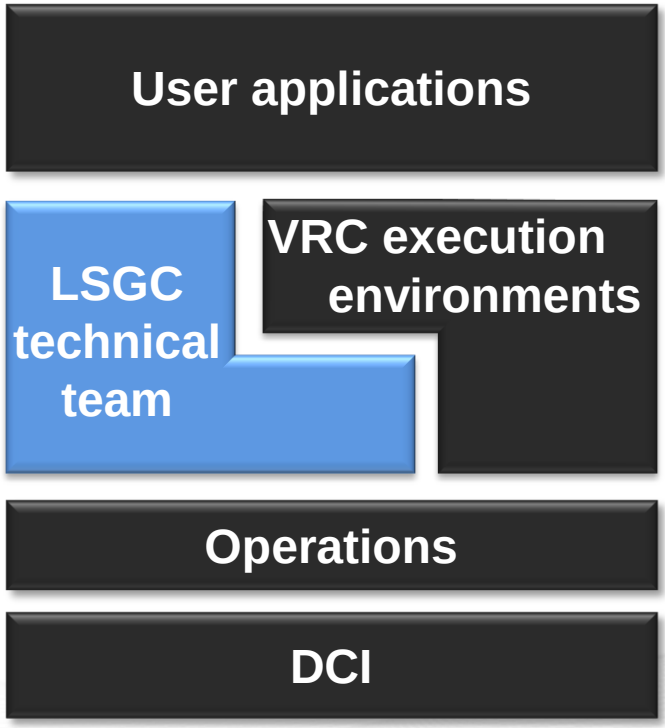
Slides: Johan Montagnat

Life Sciences Grid Community (LSGC)

- LSGC is a broad spectrum community
  - Life ScienceS deserves plural, largest community beyond LHC experiments
  - Multi-grids (although a large fraction is an EGI Virtual Research Community – implemented through 5 VOs)
  - France Grilles manages the LSGC (both scientific & technical mgt) and provides most human resources
- Had to face HealthGrid association ending
  - Proved to be a resilient, loosely coupled organization

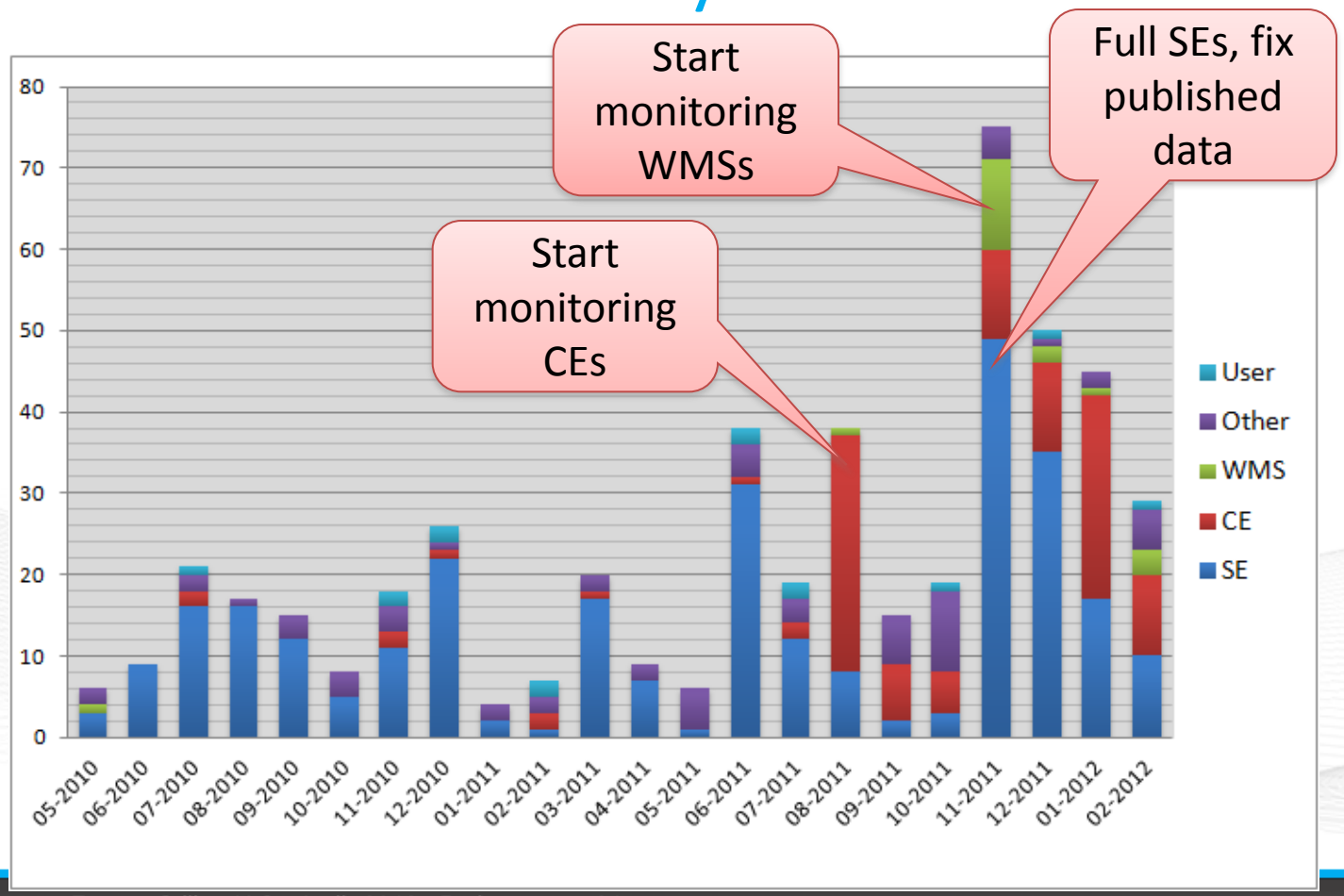
LSGC achievements

- LSGC Technical team
  - Improve resource reliability (VO-wise monitoring)
  - Liaise with resource providers and EGI (100+ resource centers)
- Resources consumption in 2011
  - 4.2 Mjobs = 1600 norm. CPU years
  - 2 PB of data
  - 200+ active users (lot of small experiments)



LSGC connection with operations

- ~500 GGUS tickets handled last year
  - Currently cleaning up 8 years of activity (left users and obsolete data; site resources configuration and availability…)



LSGC activity

- Too many “small” experiments to report
  - *omics, HT sequencing, structural biology, system biology, medical imaging, neurodegenerative diseases, models and simulations….
- Stringent requirements and community-driven environments
  - Data-driven workflows, distributed databases, knowledge repositories and inference, long-term data archival, grid pilot-jobs and clouds (virtual environments) exploitation

LSGC needs and perspective

- Technical management
  - Proved to be a key point, mandatory between end-users and operations
  - Currently evaluating existing and missing tools for proper user community management
- Perspectives
  - Improve 200+ users to 100+ sites liaison
  - Manage community data (scattered in 100s experiments)

Scientific communities: the international well established

- Earth Sciences
- slides Monique Petitdidier

Earth Science [1]

- Earth Science research is intrinsically a Global undertaking that addresses both fundamental understanding of the Earth's evolution, and augmented societal issues concerning natural hazards, energy resources, environmental changes and national security.
- Monitoring networks and satellite observations are continuously transmitting a wealth of data integrated in loosely coupled international data infrastructures with standard DMS and formats
- **ES data intensive applications**
  - Data analysis of large volume and large number of distributed data sets
  - Data simulation, inversion and assimilation
  - Data integration and parallel data mining; Large scale synthetic data analysis;
  - Scenario and stochastic analysis (parametric methods).
- **Challenges and Needs**
  - A service-oriented architecture and a platform of services and tools across Data, Grid, Cloud and HPC infrastructures

Earth Science [2]

- **Scientific domains**
- **ESR**: *Atmospheric sciences (interface with satellite observations); Earthquake and Seismology (interface EGI/EPOS); Hydrology; Climate (interface EGI/ESG);*
  EGEODE: *seismic and geophysics exploration using Geocluster on the Grid provided by CGG-Veritas – update of Geocluster porting*
- **European dimension**
- **Earth Science Virtual Research Community** ***(EGI VRC, EGI-Inspire): Earth Science (M. Petitdidier, IPSL & H. Schwichtenberg, SCAI/Fraunhofer)***
- **ESFRI-PP projects**: **EPOS, EMSO …**
- **FP7 Initiatives:** *Infrastructure:VERCE (CNRS-INSU), DRIHM; ENVRI capacity call; EnviroGrid*
- **International dimension**
- ***Collaboration with other ES Grid activities***: *in* ***Europe****, in* ***Asia*** *(Taiwan, Vietnam),* ***Latin America*** *(GISELA through IPGP)*

Earth Science – National dimension

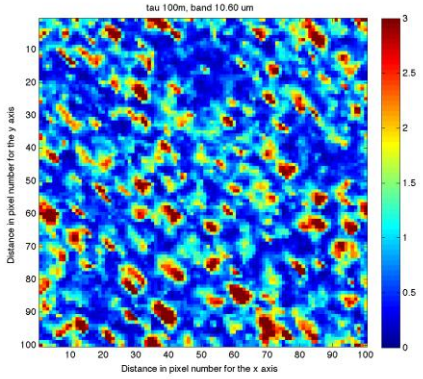
- 2 Sites : IPGP&IPSL at IPGP, BRGM & access to most of the French sites and many European sites
- VO:
  - ESR VO: IPSL, IPGP, LOA, LGEI, BRGM ….
  - EGEODE:stand-by for the users of Geocluster - GeoAzur, IPGS-EOST, ISTEP, UPMC-Sysiphe, ENS-Paris, IPGP
- Partners
  - Few persons in each laboratory and Institute
  - Need more Grid people in ES laboratories or Institutes to support new teams and applications
  - Need an engineer devoted to dissemination, contact with European Grid activities, technical support and survey in replacement of Dr Monique Petitdidier (IPSL)

EARTH SCIENCE NEW APPLICATIONS

- Radiative Transfer Model for Cloud (LOA, Lille)

- Interpretation of POLDER/Parasol satellite observations
- Monte Carlo Method (12days% 3years with a 8 processors computer)
- A paper to be finalized and a PhD in preparation

Brightness Temperature simulated with a 100m-resolution (T. Fauchez)



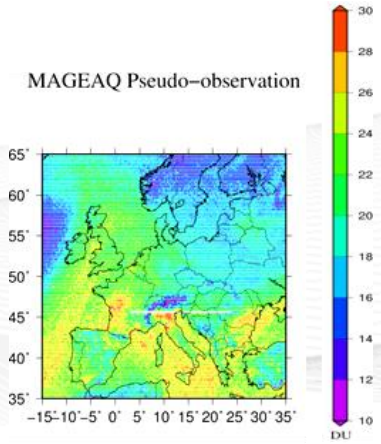
GeoQAIR (IPSL/LISA, Paris)

•Development of a OSSE (Observing system simulation Experiment): Design and validation of a new satellite instrument concept for air pollution observations

•Pseudo-observation database from 4 previous satellite instruments: 45.000.000pixels from 4 instruments i.e. around 50O.000CPUs –quite done

•1 paper published

0-6km ozone partial column simulated for one of the two future geostationnary instruments : MAGEAQ – to air quality.



EARTH SCIENCE APPLICATION - new issues

- Interface ESG (Earth System Grid) & EGI (IPSL, Paris)
  - ESGF: database of CMIP5 climate model outputs provided by worlwide teams for the next IPCC (International Panel on Climate Change)
  - Goal: Use EGI for parametric, regional climate and impact studies,Extreme event statistics…
  - Testbed: Climate simulation: MPI application, Data retrieved from ESGF
- New issues
  - Synchro-data: CMIP5 Access tool – due to its performance largely used by Climate teams
  - Credential_translation_service: to overcome the difference of security protocols between ESGF & EGI
  - A&A problem addressed to EGI & ESGF in collaboration with P. Kershaw (STFC, UK)
    - Participation & contribution to the workshops on federated identity systems for scientific collaborations
  - Impact on new climate projects

Earth Science

- To face the data intensive challenges Grid seems to be a good candidate but not effectively adopted in Earth Science:
  - In France Few teams (scientists + Grid engineers) then few permanent involvement
  - Lack of local expertise & information, difficult to use, non-unified security protocols
  - European and International collaboration in ES GRID mainly based only on best efforts => difficult for common works & developments, even information
- ES Scientists used HPC, Cloud or Grid according to the stage of their work
  - Need: architecture hiding complexity of the underlying infrastructures through mechanisms to manage automatic mapping of credentials, data access adapters and to call heterogeneous security protocols
  - Cf. : Climate application, EU Projects: VERCE, DRIHM, EnviroGrids
- DISSEMINATION
  - European Geosciences Union: Grid, Cloud & HPC sessions, Booth
  - Seminar organized by CNES, several Grid centres and projects
  - Intensive computation and Grid days organized in Egypt and RDC during a school on Space weather

Scientific communities: the international well established

- Astrophysics & Astroparticles:
- slides Giovanni Lamanna
- slides Franck Lepetit

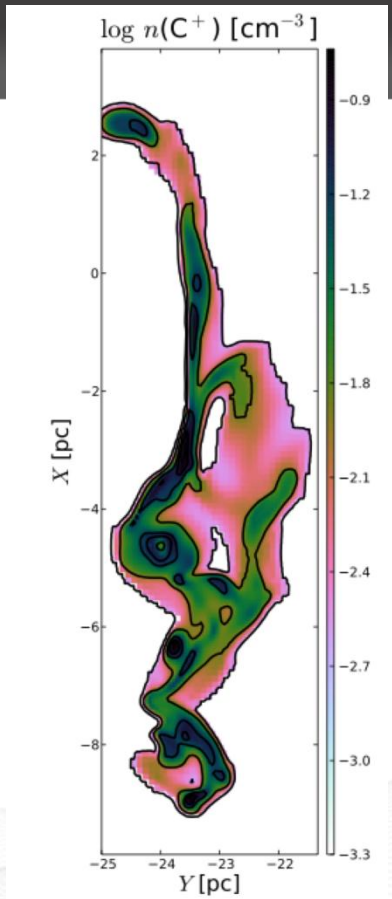
A&A Community

Specificities of A&A applications on Grid

- Embarassingly parallel applications
- Heterogeneous applications

Example of applications in 2011 - 2012

- Simulations of atmospheric cascades for CTA project
- Modeling of chemical structure in the interstellar Medium
  - post-treatment of large simulations
- Study of orbits stability in the Earth neighborhood for spatial missions



A&A Community

Grid is potentially useful for several astrophysical applications

- Exploration of parameter space
- Online simulation services (Virtual Observatory)
- Some massive data analysis

Up to now few users because of

- historical reasons :
  use of national computing infrastructure (IDRIS, CINES, CCRT)

- difficulties to link Grid and Virtual Observatory

- no strong user community

Scientific communities: the international well established

- Complex Systems

Slides: Romain Reuillon

Requirements

- Grid is used to solve inverse problems:
  - Model calibration
  - Sensitivity analysis
  - Systems optimisation
  - Study of dynamic properties
  - Reconstruction of dynamic from 3D + times images
- Applications in divers scientific domains: geography, social-sciences, biology, food, neurosciences...

Example

- Calibration of an agent based model of inter-cities geography.
- Which values of the input parameters of the models leads to physically realistic situations.
- Resolution using evolutionary algorithms:
  - Evaluation of a set of parameters for the model => 30 replications of the model or 30 minutes CPU
  - Solving the problem requires at least 100 000 evaluations
  - 2000 evaluations are executed at the same time on the grid when a result comes back a new job is submitted.

Means

- Usage of vo.complex-systems.eu and of the workflow management system to experiment on models: OpenMOLE.
- User demands for using the grid in the complex-systems community are increasing, however the computational resources accessible to the VO are under-sized.
- The LAL cluster which should provide computational power for the VO doesn't accepts jobs anymore (0 jobs running)
- We are presently in process with France Grilles to get a support to increase the VO capacities, waiting for new investments.

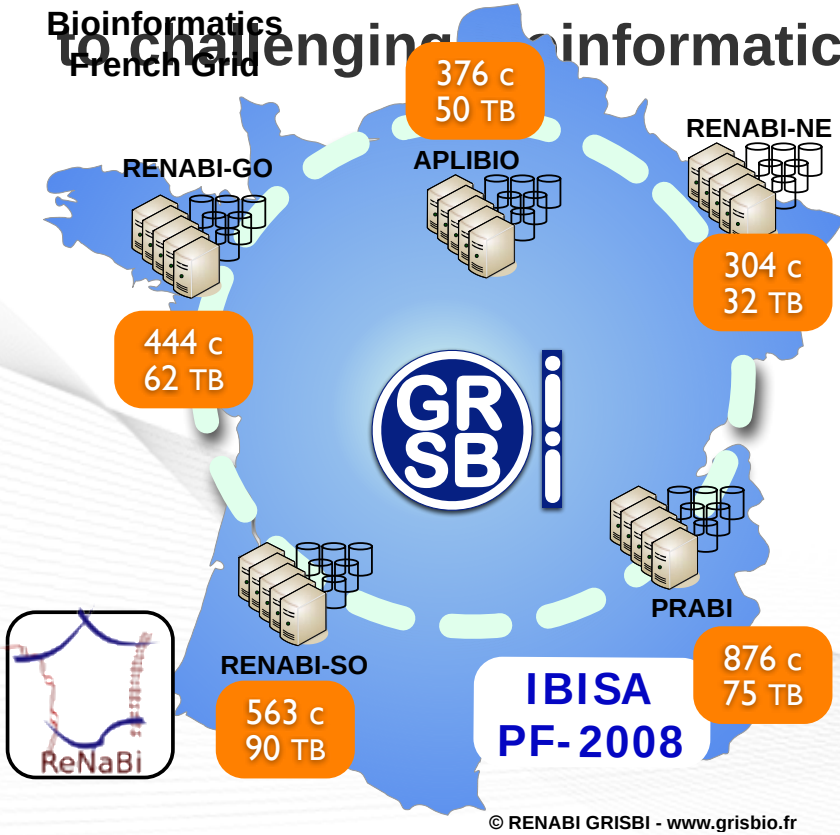
RENABI/GRISBI

- A thematic French community

Slides Christophe Blanchet

French RENABI GRISBI

Federate Life Science community and provide with IT answers to challenging bioinformatics applications

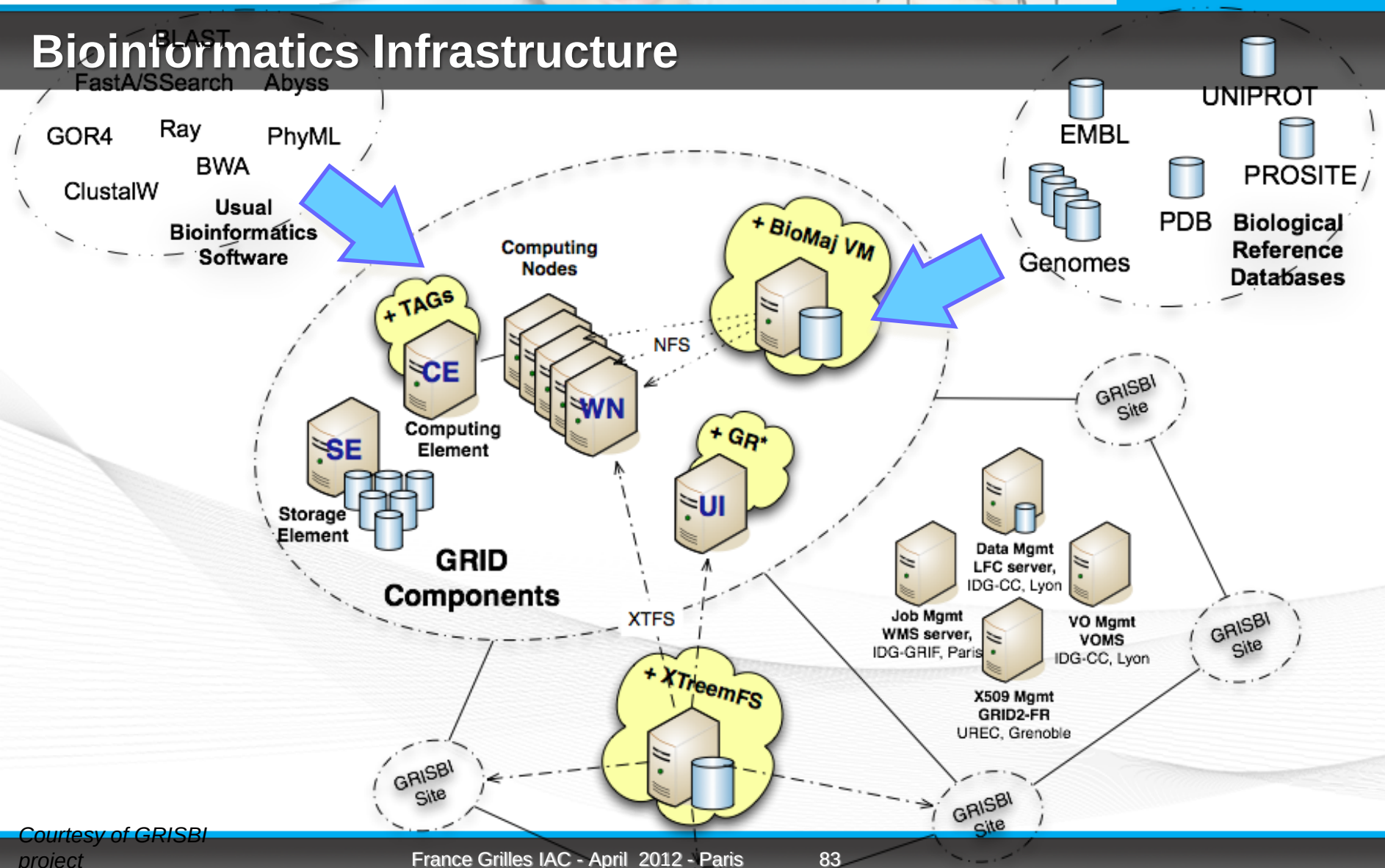


- Working group on technological aspects:
  e.g. gLite, DIET, GridWay, BioMaj, ActiveCircle, Caringo, HDFS, XtreemFS, dCache, …
- Build a distributed infrastructure
  - Above the regional centres of RENABI (5)
  - Based on the bioinformatics platforms (9)
    - approved as national RIO / IBISA
    - ~90 registered members
  - Computing resources
    - in PFs **2600 cores, 310 TB** storage
    - in grid **860 cores, 26TB** storage
- Financial support by **RENABI , IBISA 2008-2011, IDG 2009-2010**
- In collaboration with the national IT infrastructures: IDG, GENCI, Grid5000, regional computing centres

www.grisbio.fr

Courtesy of GRISBI project

Bioinformatics Infrastructure



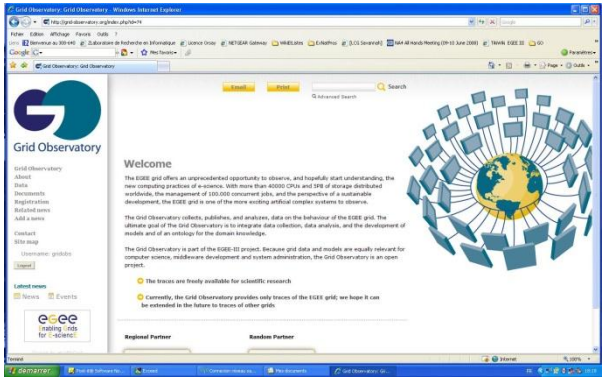
Courtesy of GRISBI project

Grid Observatory, Green Observatory

- Slides Cécile Germain

The Grid Observatory

- Grid traces collection portal
  - Started in EGEE-III, online since October 2008
  - Digital curation of the behavioural data of the EGI grid: job traffic, file access, middleware, energy usage
- Recent supports
  - EGI-inspire, INRIA (ADT), CNRS (PEPS), COST Action IC0804

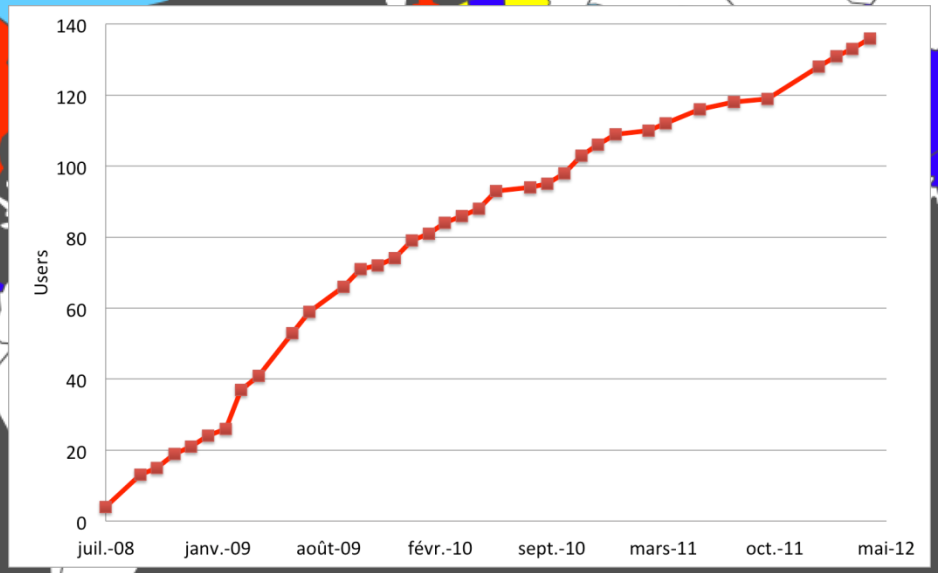


www.grid-observatory.org

- Research-production link
  - Complex systems description
  - Models, optimization, Autonomics

Usage statistics

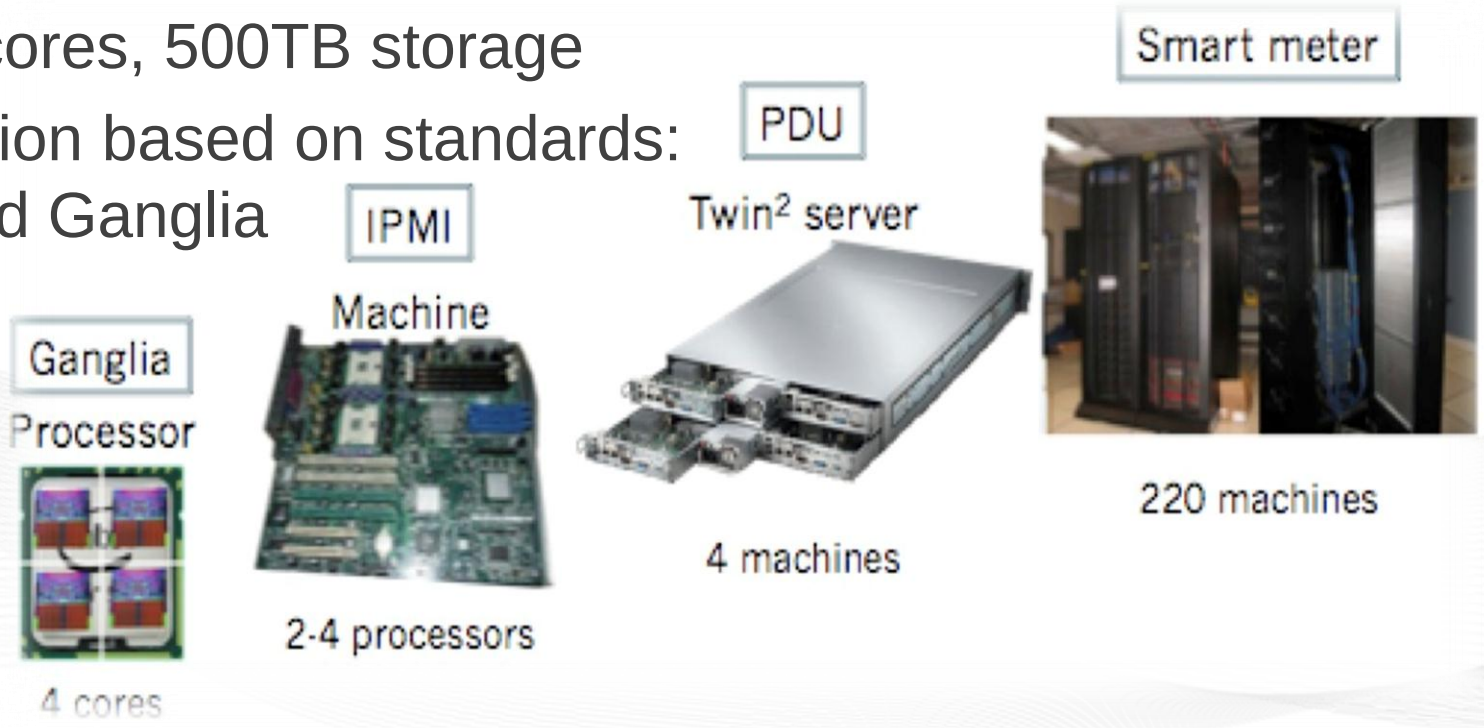
30+ international peer-reviewed computer science publications



The Green Computing Observatory

Extensive traces of energy consumption from LAL site

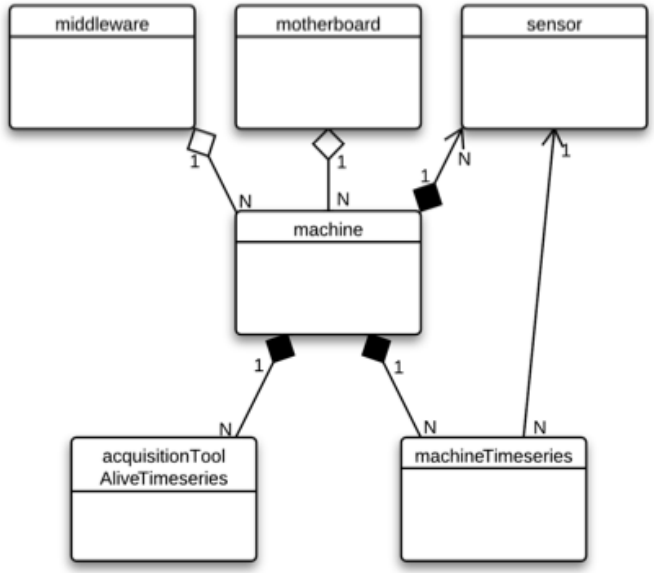
- 2200+ cores, 500TB storage
- Acquisition based on standards: IPMI and Ganglia



The Green Computing Observatory

Extensive traces of energy consumption from LAL site

- 2200+ cores, 500TB storage
- Acquisition based on standards: IPMI and Ganglia
- Integration based on a new ontology of IT systems measurements, including virtual machines
- The XML-formatted traces are published by the GO portal

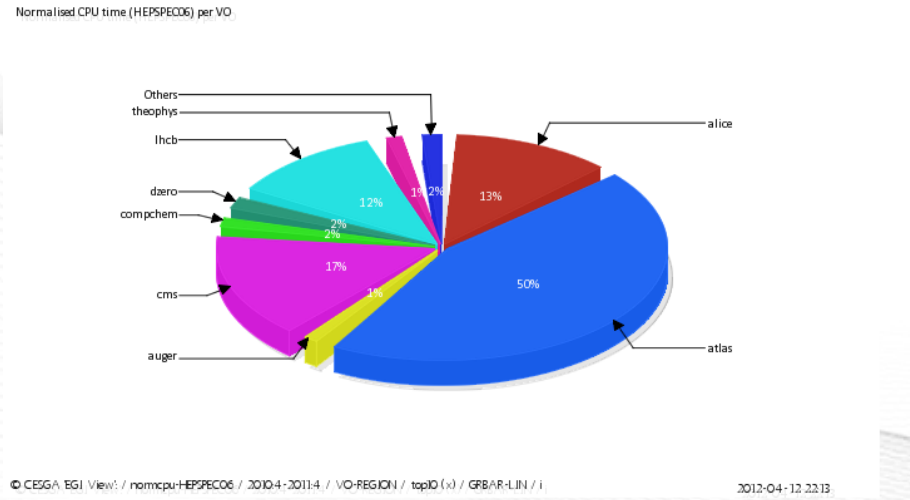


Additional slides

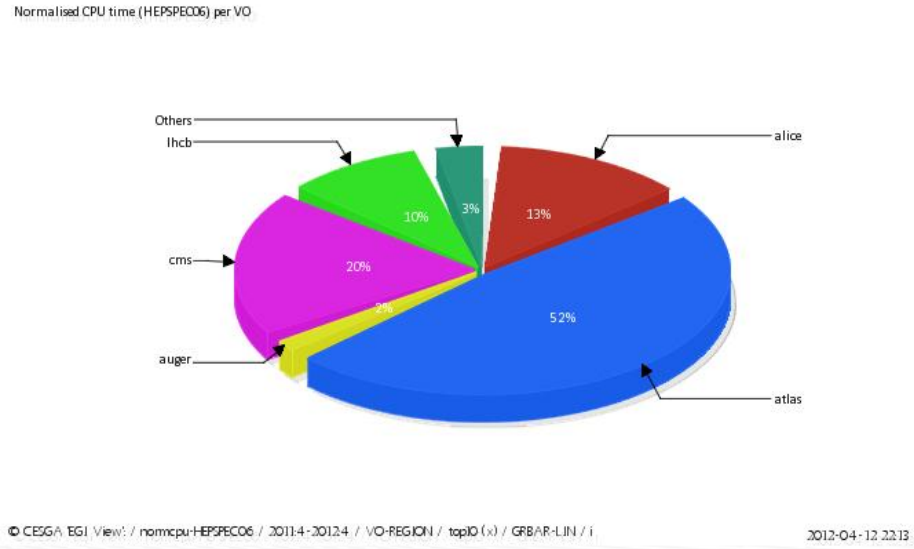
- additional metrics
- details

World wide EGI CPU resources consumed per-VO

April 2010- April 2011



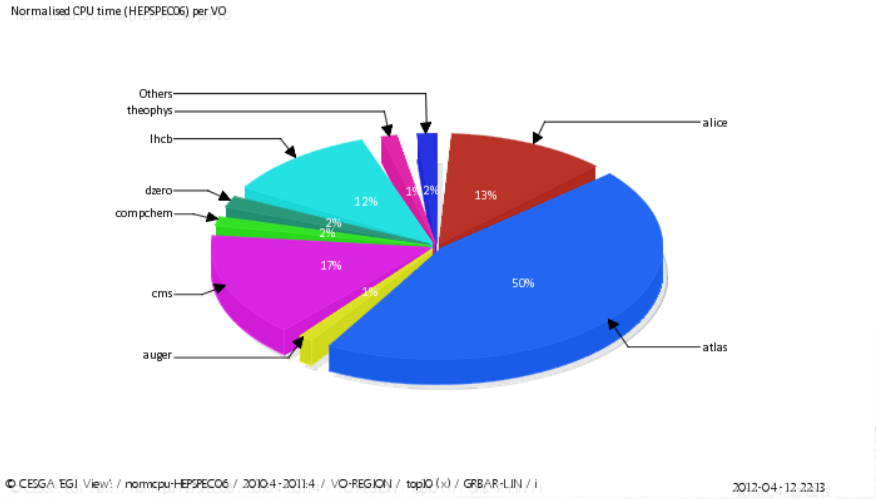
April 2011- April 2012



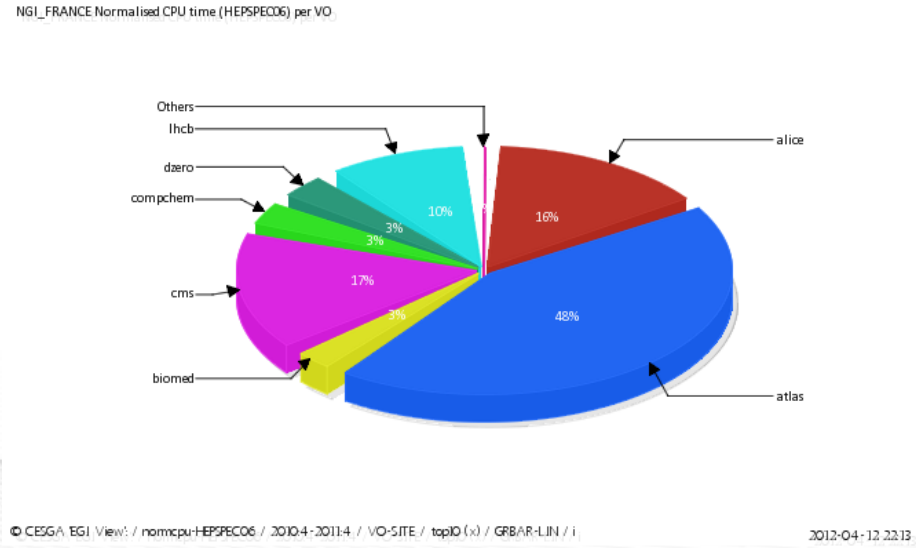
Pie Charts showing the share in Normalised CPU time (HEPSPEC06) per VO (only information about **TOP 10 -ordered by CPU time- VOs** is returned).

Normalised HEPSPEC06 CPU time per VO 4/2010 – 4/2011

World wide EGI Top 10

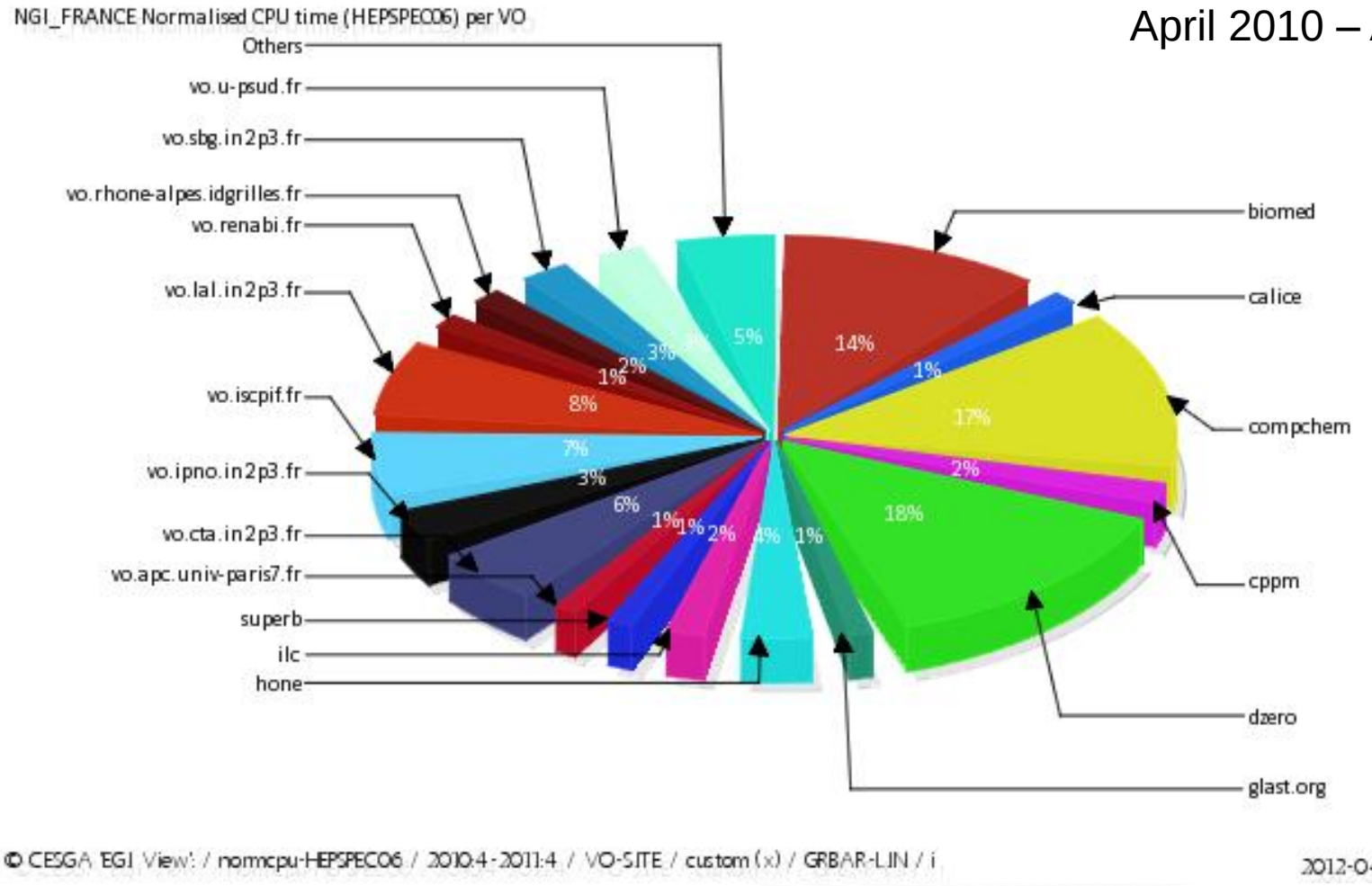


NGI-France top 10



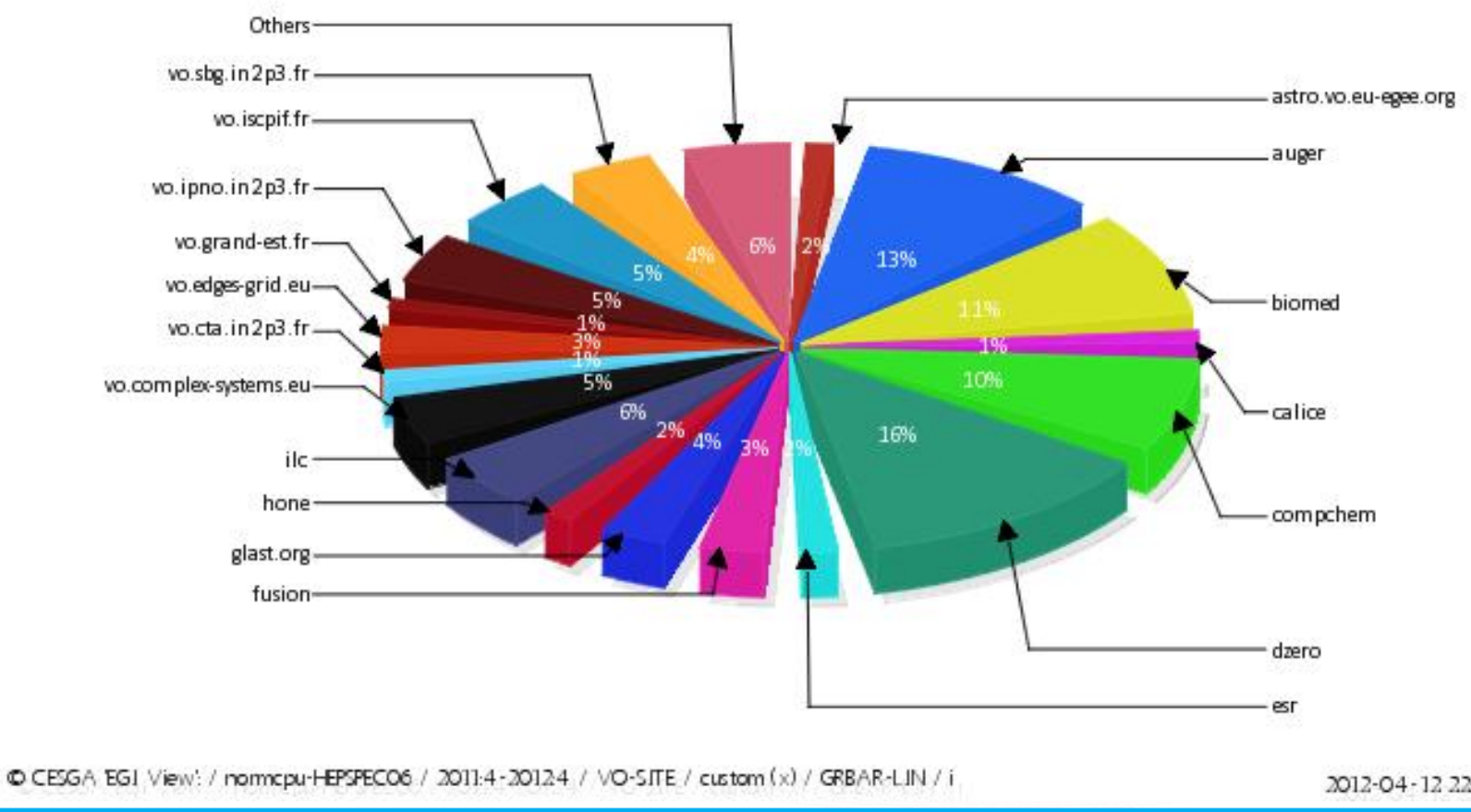
Normalised CPU time (HEPSPEC06) per VO without LHC

April 2010 – April 2011



Normalised CPU time (HEPSPEC06) per VO without LHC

April 2011 – April 2012



Training details

		nb days	nb trainees
Tutorial utilisateur grille	02/06/2010	1	15
Journée Outils pour le calcul Scientifique en Rhône-Alpes	29/06/2010	1	23
GRISBI - Calcul Scientifique sur Grille pour la Bioinformatique	27/09/2010	5	27
Formation utilisateurs - Présentation et initiation aux grilles de calcul	18/10/2010	2	16
Formation utilisateurs ligne de commande du middleware de grille de calcul Glite	22/10/2010	1	23
Formation sur l\'utilisation de OGC (service web) dans le contexte de la Grille EGI	26/10/2010	1	14
Formation sur l\'utilisation de DIRAC dans le contexte de la Grille de Calcul EGI	27/10/2010	1	20
Formation utilisateurs DIRAC	09/12/2010	1	11
Formation utilisateurs - Présentation et initiation aux grilles de calcul & Apprentissage de l\'outil DIRAC	01/02/2011	3	19
Open Nebula - Stratulab	30/03/2011	2	8
Open Nebula - Stratulab	04/04/2011	2	5
GateLab	08/04/2011	1	20
Formation utilisateurs - Présentation et initiation aux grilles de calcul & Apprentissage de l\'outil DIRAC	19/04/2011	3	19
Formation utilisateurs d\'EGI	18/05/2011	3	18
Journées GRISBI 2011	26/05/2011	2	30
ACGRID3 school	24/10/2011	9	52
Tutoriel sur l\'infrastructure de grille de calcul GRISBI	04/11/2011	1	30
StratusLab Cloud Training	17/11/2011	2	25
Formation utilisateurs - Présentation et initiation aux grilles de calcul	04/04/2012	2	16

Why not the EGI.eu AppsDB?

- Users' proposal in a previous operation workshop
- French is important for our target
- Plume is a multidisciplinary tool used by many scientific fields -> good mean to disseminate DCI!
- People are used to Plume and Plume federates people.
- Several cards are not directly related to EGI.eu
- Plume is a link to the computing research

questionnaire

- ask twice 10 people for 2-3 contacts -> 4 answers
- 6 final responses
- 3 researchers/3engineers
- 5 men/1woman
- High energy astrophysics
- 2 biology
- complex systems
- Earth Sciences Océanographie physique, domaine littoral
- astroparticules

- motivation:
  - number of CPUs, high number of jobs, jobs durations, big data
- training: 4 yes/2no (1 with specialist)
- all found close support
- all found benefit
- 4 found issues
  - 3 reliability
  - 1 big data transfer
- 2 HPC users
- 1 think to cloud