NA4: User Community Support and Expansion

C. Loomis (CNRS/LAL)

NA4 Activity Manager

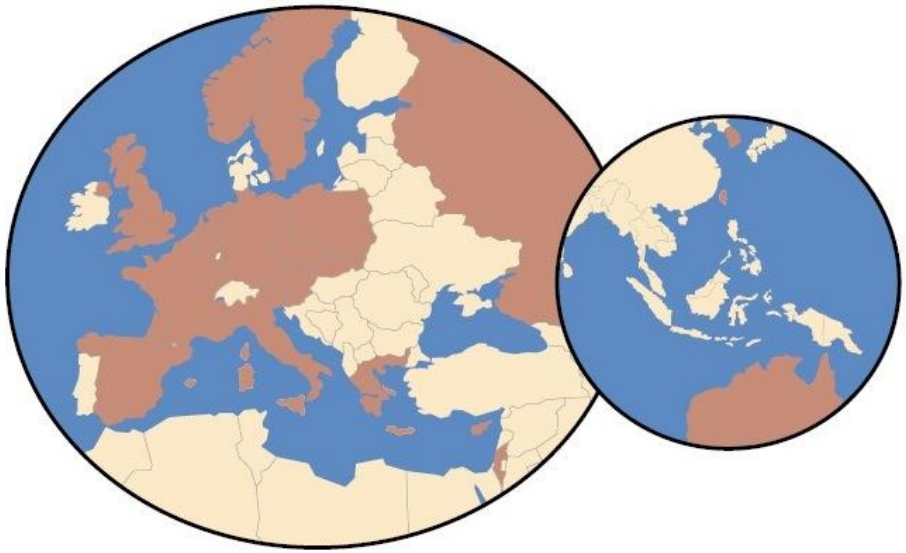
EGEE-III Final Review, 23-24 June 2010

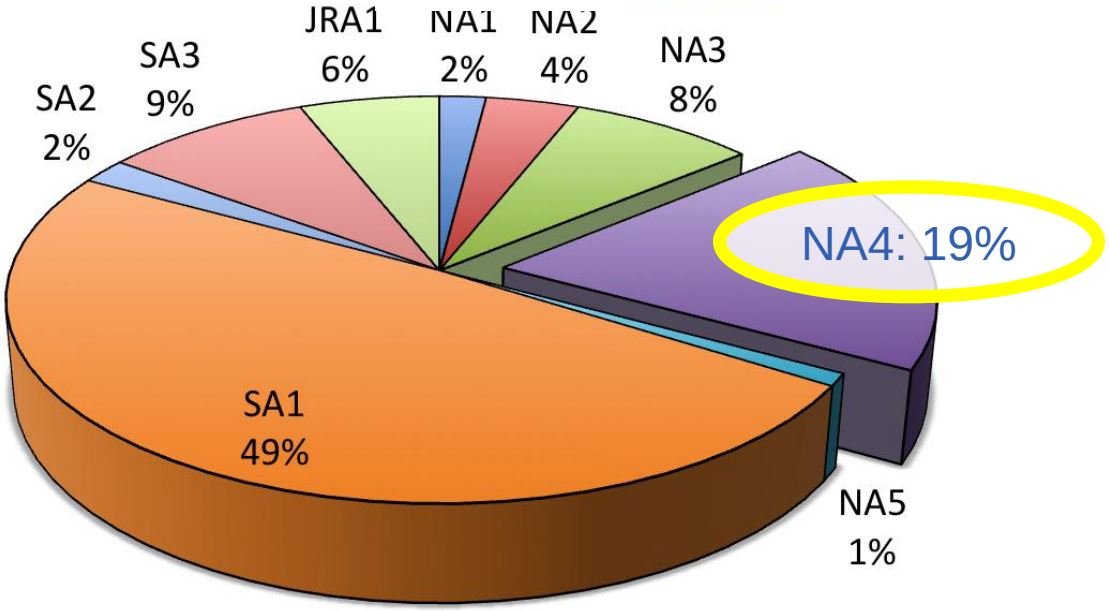
www.eu-egee.org

e-infrastructure

Activity Overview

44 Partners
287 People
19 Countries



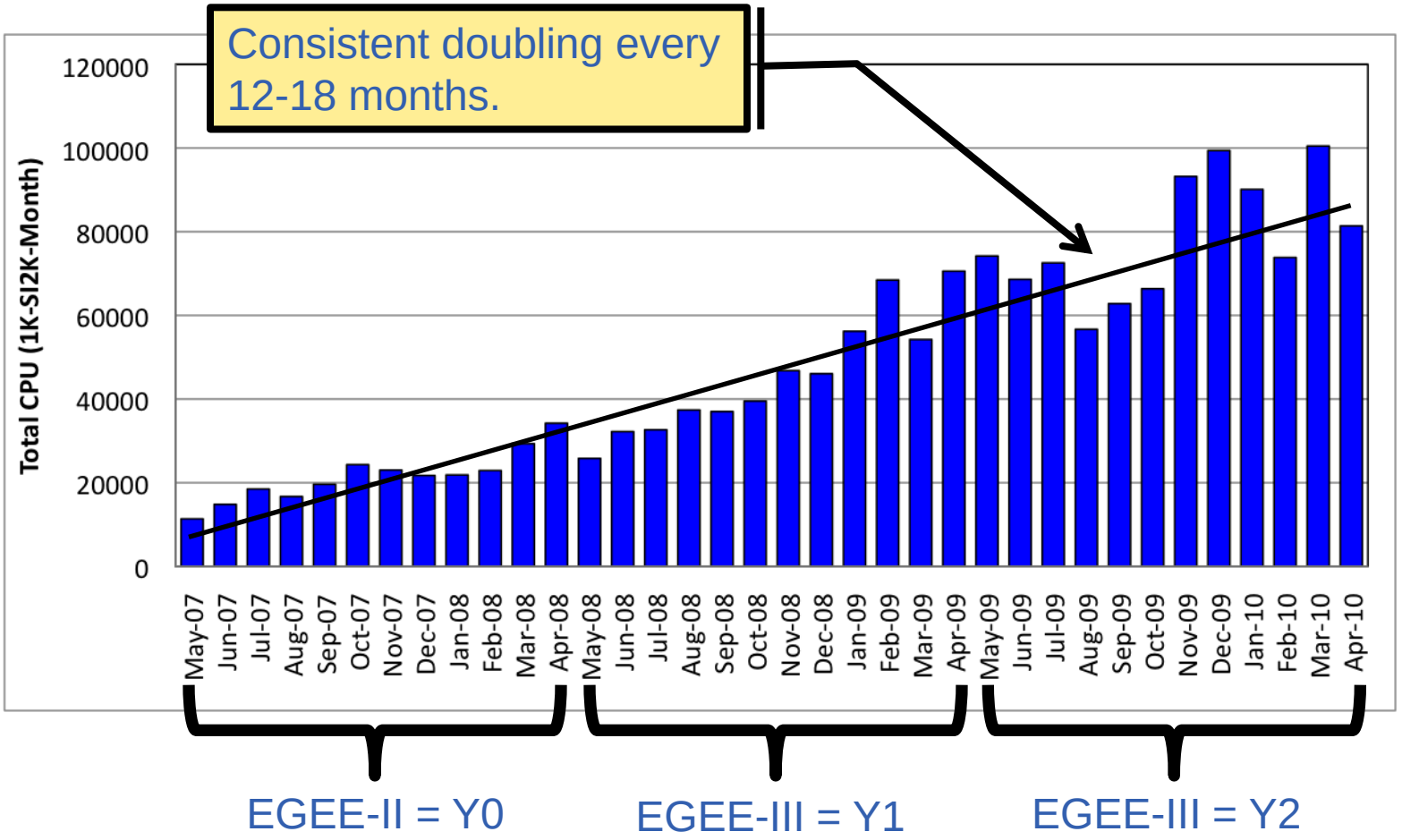


Country	PM	FTE
Austria	6	0.3
Belgium	12	0.5
CERN	162	6.8
Cyprus	12	0.5
Czech Republic	17	0.7
France	257	10.7
Germany	47	2.0
Greece	92	3.8
Hungary	126	5.3
Israel	12	0.5
Italy	209	8.7
Netherlands	29	1.2
Norway	30	1.3
Poland	42	1.8
Russia	66	2.8
Slovakia	24	1.0
Spain	211	8.8
Sweden	16	0.7
UK	41	1.7
TOTAL	**1421**	**59.2**

Tasks

- **TNA4.1: Support**
  - Virtual Organization Support (VOS)
  - Application Porting Support (APS)
  - Direct User Support (DUS)
- **TNA4.2: Strategic Discipline Clusters**
  - High Energy Physics (HEP)
  - Life Sciences (LS)
  - Earth Sciences (ES)
  - Grid Observatory (GO, CS)
  - Computational Chemistry (CC)
  - Astronomy & Astrophysics (AA)
  - Fusion (F)
- **TNA4.3: Activity Coordination**
  - Activity Management
  - Regional Coordination

CPU Utilization Over Time



CIC Portal: http://cic.gridops.org/

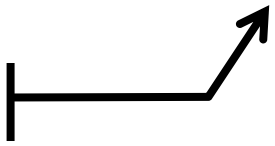
Accounting Portal: http://www3.egee.cesga.es/

CPU Utilization by Discipline

CPU (1k-SI2k-hours)

Domain	Y0	Y0 (%)	Y1	Y1 (%)	Y2	Y2 (%)	Y1/Y0	Y2/Y1
AA	1855	0.8%	15115	3.1%	12155	1.4%	8.1	0.8
CC	12654	5.8%	35112	7.2%	8239	1.0%	2.8	0.2
CS	0	0.0%	53	0.0%	2	0.0%		
ES	318	0.1%	2675	0.6%	2532	0.3%	8.4	0.9
F	2226	1.0%	2446	0.5%	4589	0.5%	1.1	1.9
HEP	177857	81.0%	376910	77.6%	784035	90.7%	2.1	2.1
INF	3334	1.5%	1337	0.3%	1858	0.2%	0.4	1.4
LS	10650	4.8%	20609	4.2%	15373	1.8%	1.9	0.7
MV	5739	2.6%	11929	2.5%	8849	1.0%	2.1	0.7
OTH	2015	0.9%	13557	2.8%	19931	2.3%	6.7	1.5
UNK	3046	1.4%	5938	1.2%	6455	0.7%	1.9	1.1
TOTAL	**219694**	**100.0%**	**485680**	**100.0%**	**864018**	**100.0%**	**2.2**	**1.8**

Non-HEP: 2.6 (Y1/Y0), 0.7 (Y2/Y1), 1.9 (Y2/Y0)



Applications and Users

Discipline	Apps.
AA	55
CC	17
CS	30
ES	19
F	6
HEP	13
LS	55
OTH	53
TOTAL	**248**

Region	Apps.
Asia/US	3
Benelux	7
Central Europe	6
CERN/DE/CH	7
FR/UK/IRE	21
IT	154
Northern Europe	10
Southeast Europe	30
Southwest Europe	10
TOTAL	**248**

Domain	VOs	Users
AA	21	348
CC	4	439
CS	6	25
ES	10	267
F	2	80
HEP	43	4859*
LS	13	561
MV	29	1639
OTH	36	1782
TOTAL	**164**	**10000**

Total in Y1: 9540

Application Database: http://appdb.eu-egee.org

Alternate link: http://grid.ct.infn.it/egee_applications/

Active Virtual Organizations

	Y1			Y2		
Domain	VOs (>0%)	VOs (>10%)	VO Names (>10%)	VOs (>0%)	VOs (>10%)	VO Names (>10%)
AA	11	3	astro.vo.eu-egee.org, auger, virgo	13	5	argo, auger, glast.org, icecube, virgo
CC	4	2	compchem, trgrida	4	2	compchem, trgrida
CS	2	1	imath.cesga.es	2	1	vo.mosfet.es-ngi.eu
ES	5	1	esr	5	2	egeode, esr
F	1	1	fusion	1	1	fusion
HEP	29	4	alice, atlas, cms, lhcb	34	4	alice, atlas, cms, lhcb
INF	20	2	euindia, nw_ru	18	4	hungrid, nw_ru, ops, vo.rhone-alpes.idgrilles.fr
LS	8	1	biomed	8	3	biomed, emnr.eu, lsgrid
MV	16	3	balticgrid, seegrid, vo.gear.cern.ch	15	3	balticgrid, see, vo.northgrid.ac.uk
OTH	19	2	geant4, theophys	23	1	theophys
UNK	58	2	litgrid, vo.nanocmos.ac.uk	52	3	litgrid, vo.nanocmos.ac.uk, vo.ssp.ac.uk
TOTAL	**173**	**22**		**175**	**29**	

Additional Metrics

CPU usage provides an important, *but limited* view of users' activities on the grid infrastructure.

Number of valid certificates issued by each accredited CA
Number of users within a VO
Number of registered VOs (by scientific discipline and by VRC)
Number of sites supporting a VO
Number and types of services provided to a VO
Number and types of services provided to a VO and passing QA tests
Average reliability and availability of sites/services supporting a VO, domain, and VRC
Number of registered applications in a VO, domain, VRC
Number of registered applications in a VO, domain, VRC with associated publications
Amount of CPU used and number of (successful) jobs
Amount of disk storage used and number of files
Amount of network bandwidth used and number of transfers

Virtual Organization Support

- **VO monitoring services**
  - These are based on the standard monitoring tools of the EGEE infrastructure and will allow VOs to exclude sites that do not pass VO-specific tests, thereby reducing job failures. This is being prototyped with the biomed VO and will be extended to others after validation.
- **VOMS-enabled collaborative tools**
  - This system extracts membership and group information from the VO's VOMS server and uses that information to configure various collaborative tools like mailing lists, chat servers, and secured web sites.
- **VO software deployment tools**
  - Developed originally for the atlas VO, LJSFI could be used by many other VOs with few modifications. Instructions for using it have been written and are available on the team's wiki page.
- **CIC Portal**
  - This is a core service used in the VO registration and management. The group contributed to the evolution of this service in order to make it as pertinent as possible for all of the actors in EGEE.
- **Documentation and Support Provision**
  - The group has maintained on its wiki site a FAQ for typical VO manager questions and links to other documentation of interest to VO managers.
  - Links to VO documentation: https://twiki.cern.ch/twiki/bin/view/EGEE/VOSupport

Application Porting Support

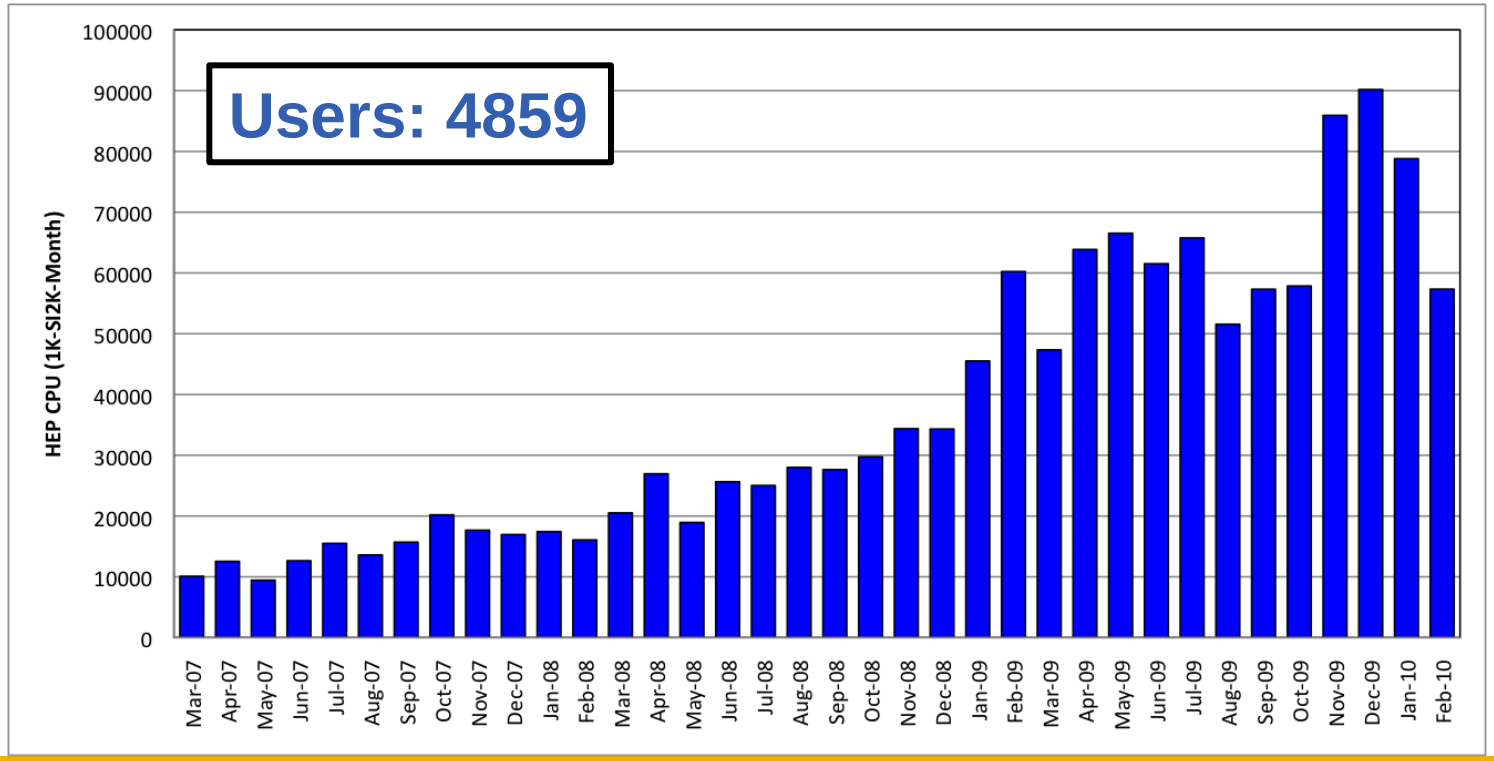
- **Consultancy and Porting**
  - Over the last year group has helped port around 20 applications to the EGEE grid infrastructure. These have resulted in 7 papers with more expected as those applications mature.
- **Use of RESPECT Tools**
  - The team is well versed in the utilization of the RESPECT packages that complement the functionality of gLite. Because of this the team members have participated in many of the training events organized by the NA3 activity.
- **Application Database**
  - This database keeps track of applications running on the EGEE infrastructure and provides a good view of the impact that the EGEE infrastructure has had on the user communities. This database was a cooperative effort of the Application Porting Support Team and the Greek Regional Coordination personnel.
- **MPI Improvements**
  - MPI is important for a large number of scientific domains. The group has participated strongly in the MPI Working Group to improve the support and efficiency of MPI on the EGEE grid infrastructure.
- **Case Studies**
  - For each ported application, a case study is written. This both documents the work of the group and provides information for porting similar applications. All of the case studies are available on the team's web pages.

Direct User Support

- **Ticket Handling**
  - The group created a support group within the GGUS system and operated 2-person, 2-week shifts to handle tickets. Tickets were generally handled quickly, although few tickets were assigned to this support group.
- **Use Cases**
  - General use cases were updated and maintained, although the main focus was on creating specific use cases for each scientific area. A number of new use cases were generated, with those from chemistry and astrophysics being the most popular.
  - http://www.eu-egee.org/fileadmin/documents/UseCases/Index.html
- **Screencasts**
  - In order to make the documentation more relevant and more appealing to users a set of 12 screencasts were created based on the existing generic use cases. Those are now available and will hopefully improve the user support experience.
  - http://www.vimeo.com/album/206141
  - http://www.youtube.com/view_play_list?p=761908FFB0570C2E

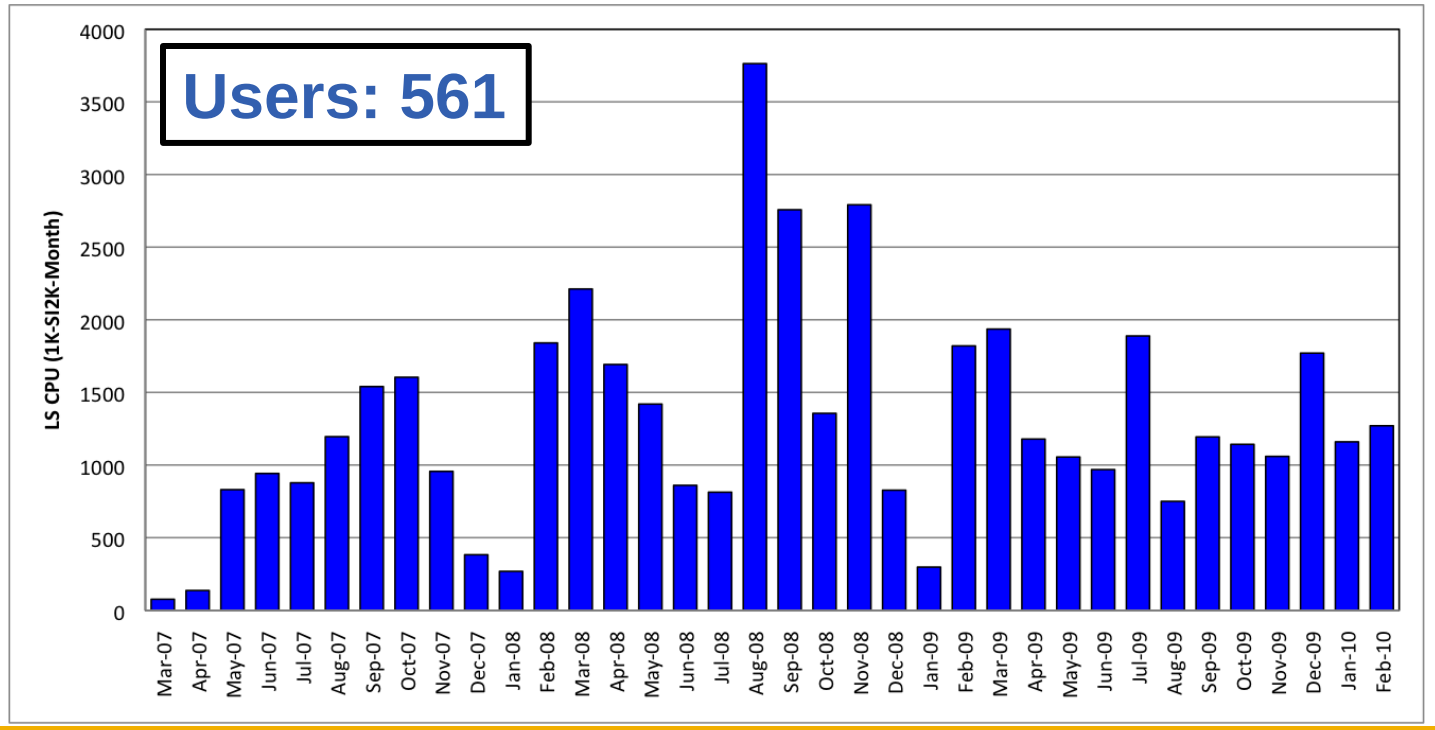
High Energy Physics

- **Large investment in testing and development of high-level tools validated by the smooth, successful start of LHC data taking and analysis.**
- **Developed tools (AMGA, Dashboard, Ganga/Diane) are heavily used within other scientific communities.**



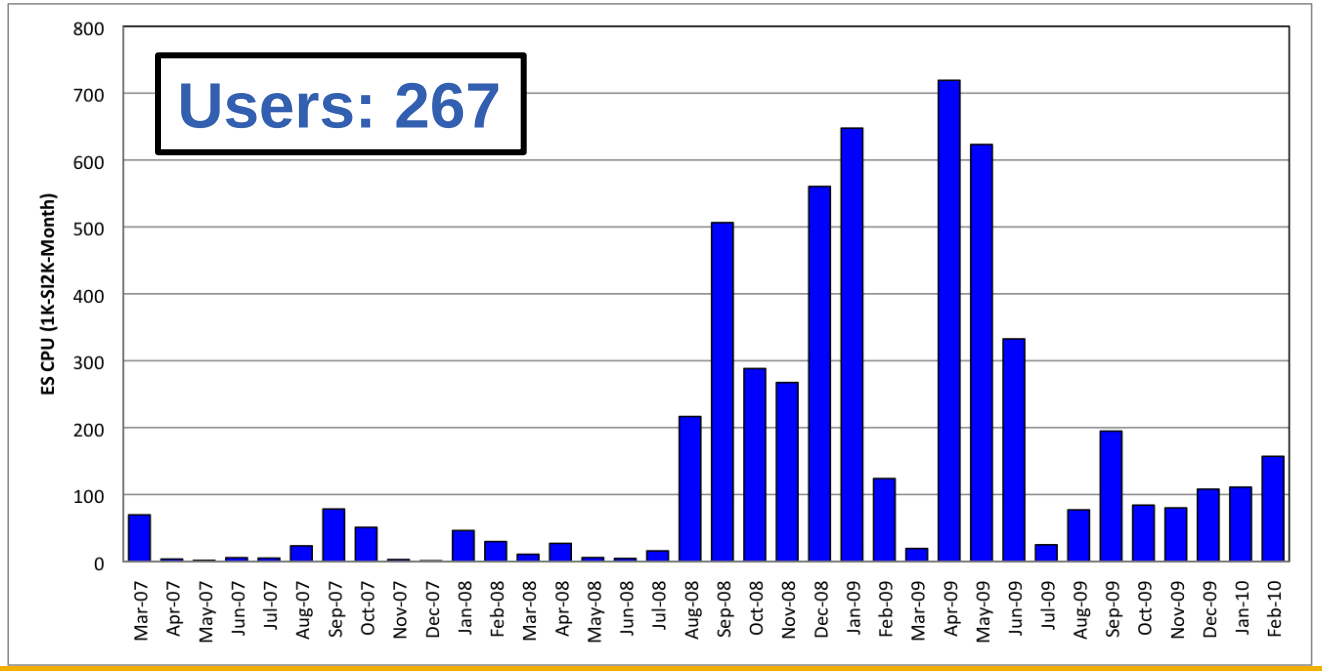
Life Sciences

- Goal: maximize grid utility while simplifying access
- Concentrated on adapting and improving tools for the life science communities
- Tools include AMGA and encrypted data management
- Organized shifts for dealing with operational problems



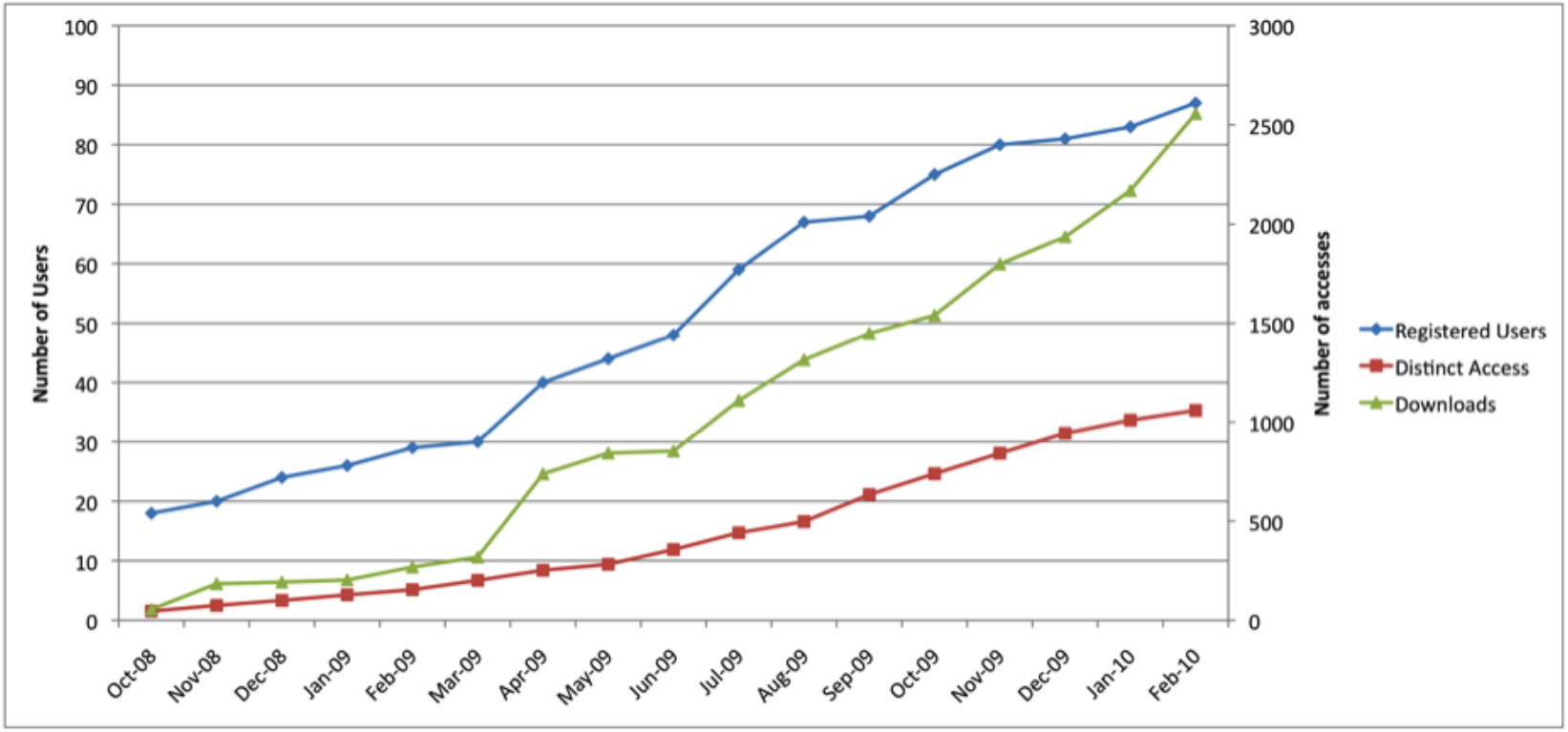
Earth Sciences

- Core strategy: community building
- Propagated grid technology to all ES disciplines, set up interactive collaboration among the members of the virtual ES grid community, and stimulated the interest of stakeholders on the political level
- Roughly doubled users within this cluster

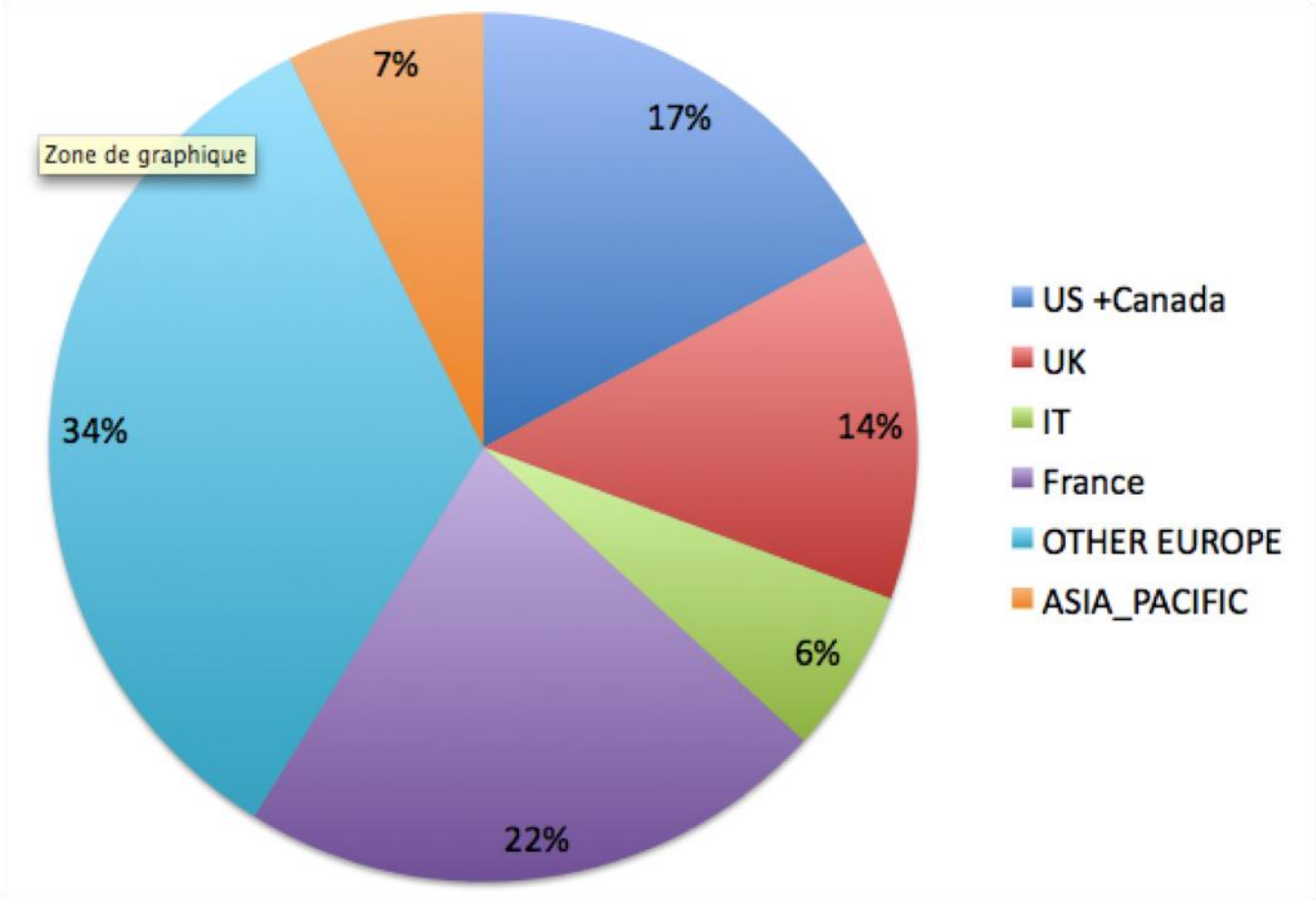


Grid Observatory

- Number of users doubled this last year: broad geographic distribution and steady increase of portal usage
- The GO has started to “live its own life”, with usage that has extended well beyond the cluster partners.



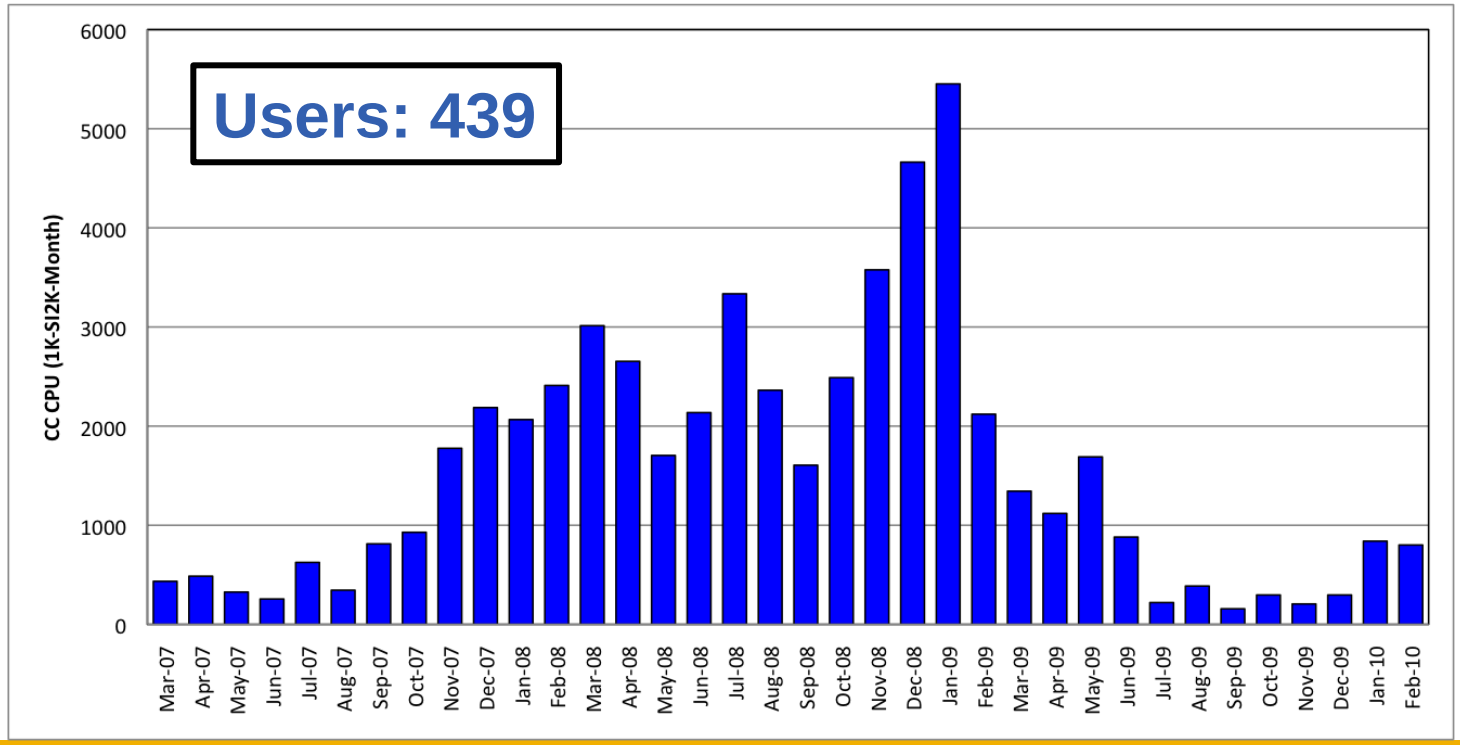
Regional Distribution



Grid Observatory: http://www.grid-observatory.org/

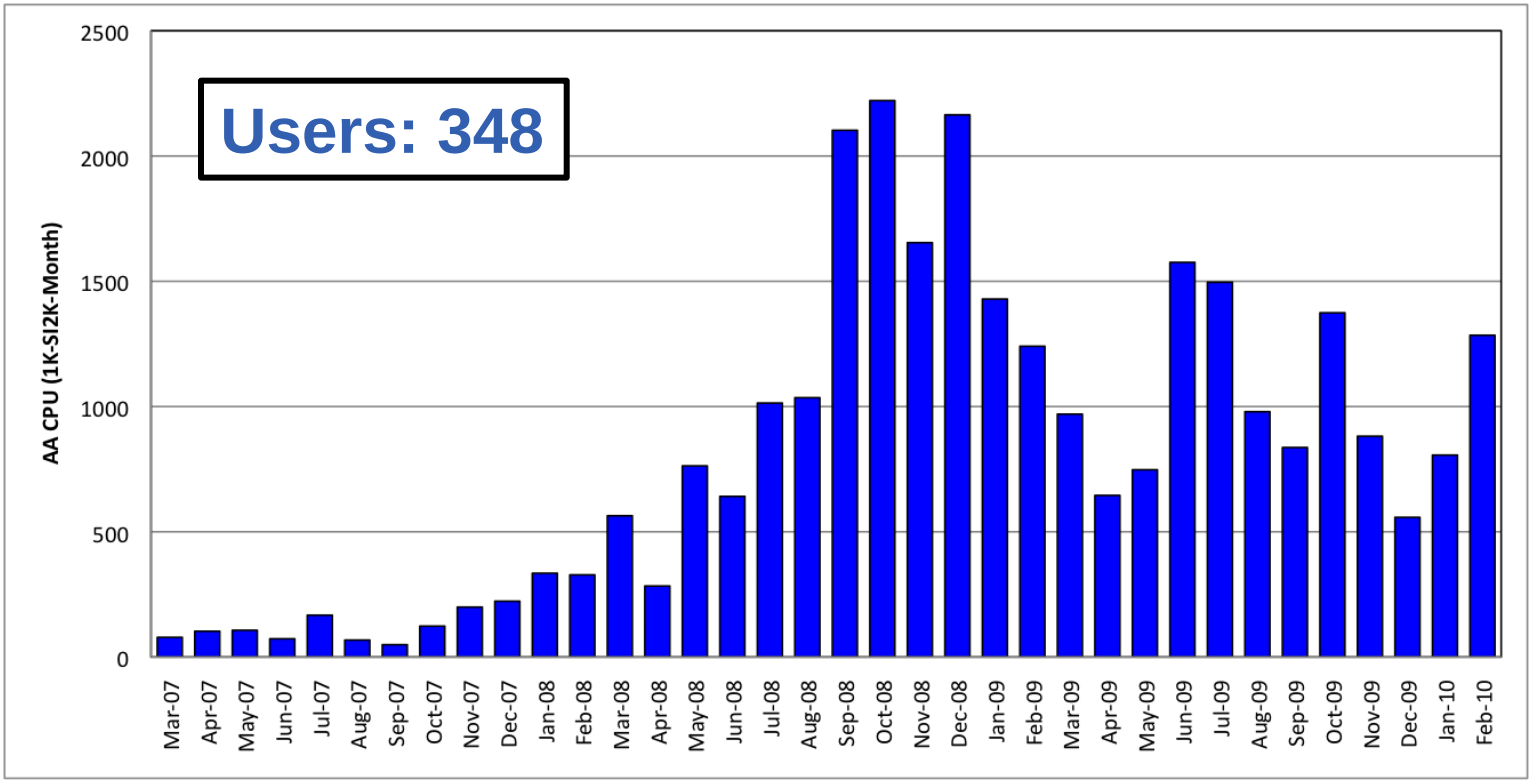
Computational Chemistry

- Create customized portal for simplified grid access
- Adapt applications to treat chemical interactions in new regimes that are difficult, impossible without grid
- Expand and parallelize standard chemical software packages for use on the grid



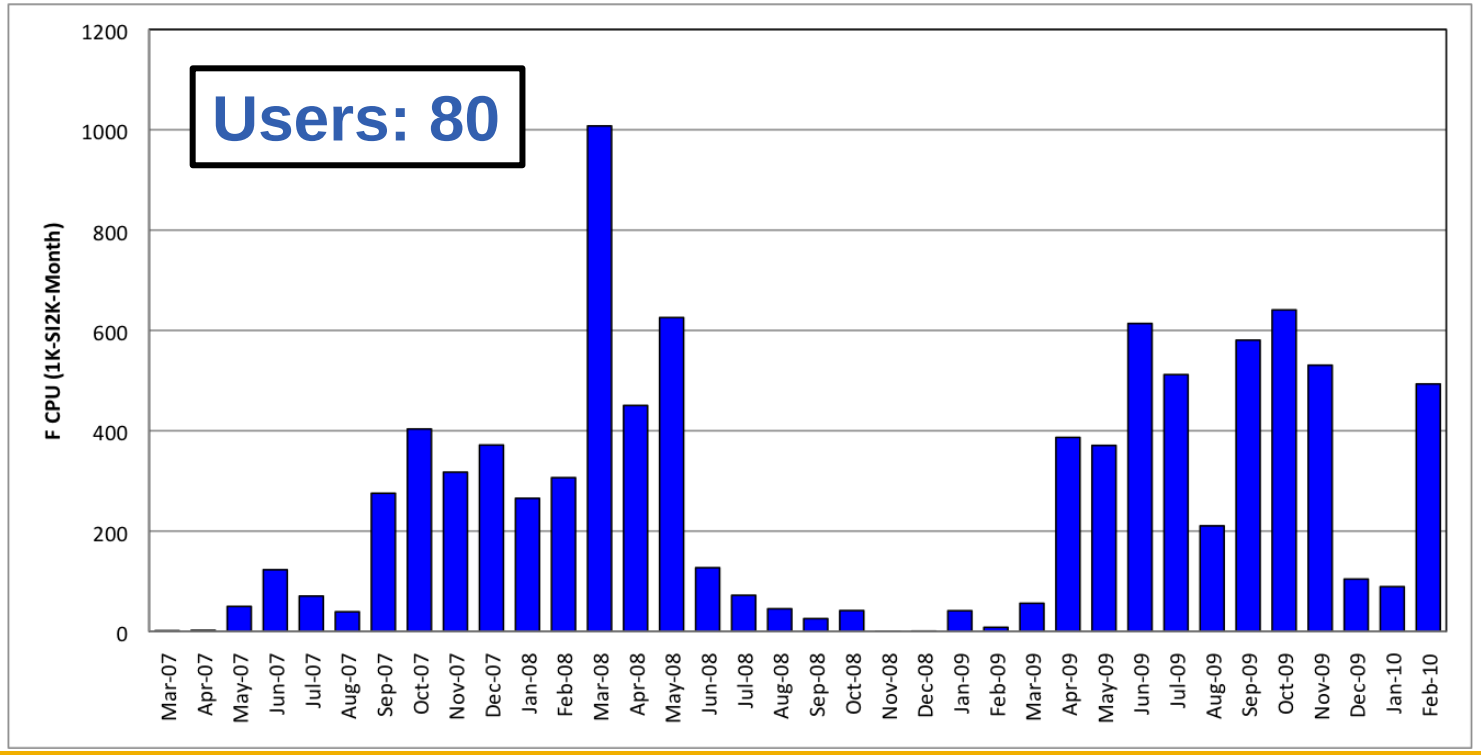
Astronomy & Astrophysics

- Ported of new applications to the grid infrastructure
- Improved grid utility for astronomers via tool provision
- Interfaced to important data stores outside of the grid
- Trained new users



Fusion

- Improved collaboration between fusion institutes, projects, and infrastructures.
- Connected grid and HPC resources: A complex workflow between EGEE-III, EUFORIA, and DEISA infrastructures has been carried out for the first time (demo).



Scientific Impact

- **DNA4.5 summarizes scientific and technical impact**
- **JOGC: EGEE applications and supporting technologies**
  - 11 papers selected from 43 submissions using ~100 reviewers
  - NA4 guest editors: V. Floros, C. Germain, F. Harris, P. Mendez
- **213 scientific papers identified**
  - 18 highlighted in DNA1.7.2

AA	Monthly notices of the Royal Astronomical Society, J. Phys.: Conf. Ser. 219, 072057
CC	ICTP Lecture Notes, Lecture Notes in Computer Science
ES	Earth Science Informatics, AGU transactions
F	Ph.D thesis, Cluster Computing
GO	KDD 2009, Lecture Notes in Computer Science
HEP	J. Phys.: Conf. Ser. 219, 062029
LS	Nature Genetics, Nature Methods

Activity Coordination

- **Activity Management**
  - All milestones and deliverables have been achieved
  - Maintained and expanded the RESPECT program
  - Encouraged community interaction via the User Forum
  - Contributed to EGEE → EGI migration (see following)

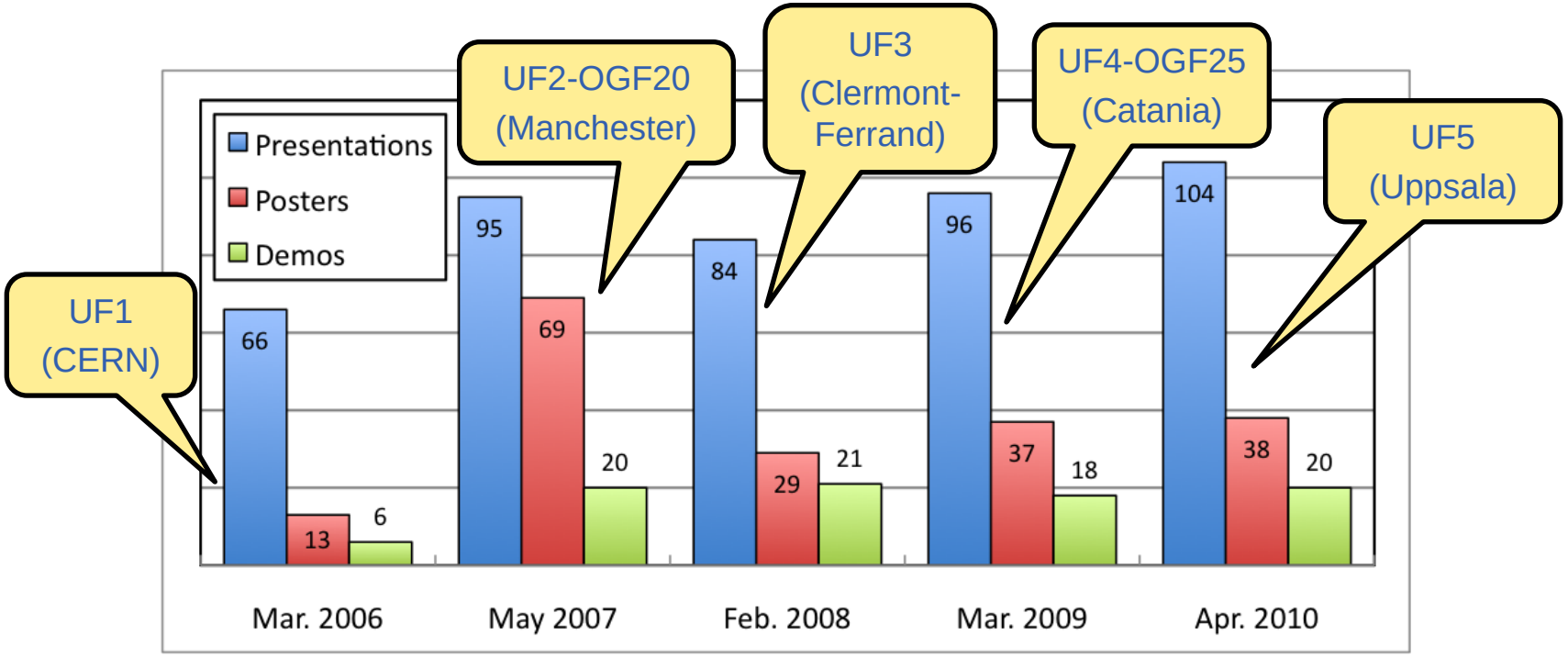
- **Regional Coordination**
  - Design, implementation, and filling of Application Database
  - Developed knowledge of user communities within regions

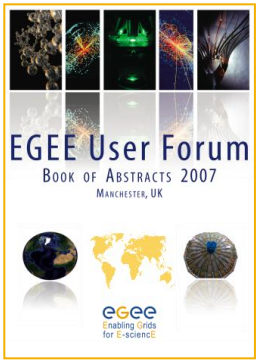
RESPECT

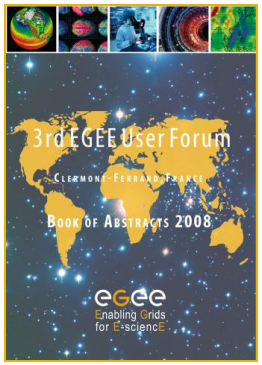
- **Identify third-party software that works well with gLite.**
  - http://technical.eu-egee.org/index.php?id=290
  - Integrated with application database in EGI

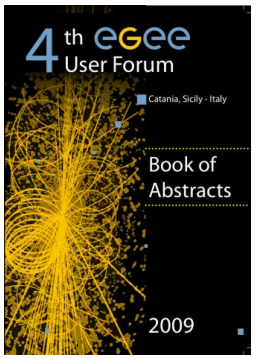
- **Y1 Tools (11)**
  - P-GRADE, Ganga, Migrating Desktop, g-Eclipse, i2glogin, Virtual Control Room, GridWay Metascheduler, DIANE, GRelC, Instrument Element, StoRM
- **Y2 Tools (7)**
  - ROOT/PROOF, VBrowser, Dashboard, DiGS, Work Binder Application Service, ToPoS, ESIP

User Forum









electronic

http://technical.eu-egee.org/index.php?id=148

JOGC (V8, No.2, June 2010): http://www.springerlink.com/content/g6234w046001/

Transition EGEE → EGI

- **Series of "SSC" workshops:**
  - Used to define user community model in EGI
  - Used to define strategy for VRC organization and support
- **Virtual Research Communities**
  - Defined needs to move towards sustainability
  - Broadened and strengthened ties within existing communities
  - New communities: Humanities, Photon, and Complexity Science
- **Result:**
  - EC project proposals: ROSCOE, SAFE, CUE, SGI, TAPAS, …
  - None passed initial selection thresholds
  - Limited funding of existing communities through EGI-InSPIRE

Support Activities After EGEE

- **VO Support**
  - Existing documentation available: https://twiki.cern.ch/twiki/bin/view/EGEE/VOSupport
  - Recommendations for VO management: https://edms.cern.ch/document/1070320/1
  - CIC Portal will continue as part of EGI operations tools
- **Application Porting Support**
  - Existing team and portal will continue in EGI-era
  - Expertise conserved within the EGI community
- **Direct User Support**
  - Use cases and screencasts available for use/evolution
  - Current ticket handling team will not continue
  - Difficult to transfer expertise as responsibility is very distributed

Virtual Research Communities

- **Failure of EC-funded proposals:**
  - Raised serious doubts about VRC model itself
  - Left (most) existing communities without funded effort for European-level coordination of grid-related activities
- **Beginning of EGI:**
  - Need to review/validate user community organization
  - Need to review/validate interaction model with EGI
  - Communities will continue on best-effort basis in short term with existing resources
- **Concerns:**
  - Lack of European-level coordination may lead to fragmentation of communities
  - Reliance on ESFRI projects alone will ignore large segments of scientific communities

Summary

- **User Community**
  - 10000 users, 248 applications, 175 VOs
  - Majority of use from 29 core VOs
  - Overall CPU 2x greater, driven by HEP usage
- **Impact**
  - DNA4.5 contains details on technical and scientific impact
  - Books of Abstracts and special JOGC issue show breadth
- **EGEE → EGI Transition**
  - Communities worked hard for realistic, viable transition
  - Current situation concerns most communities