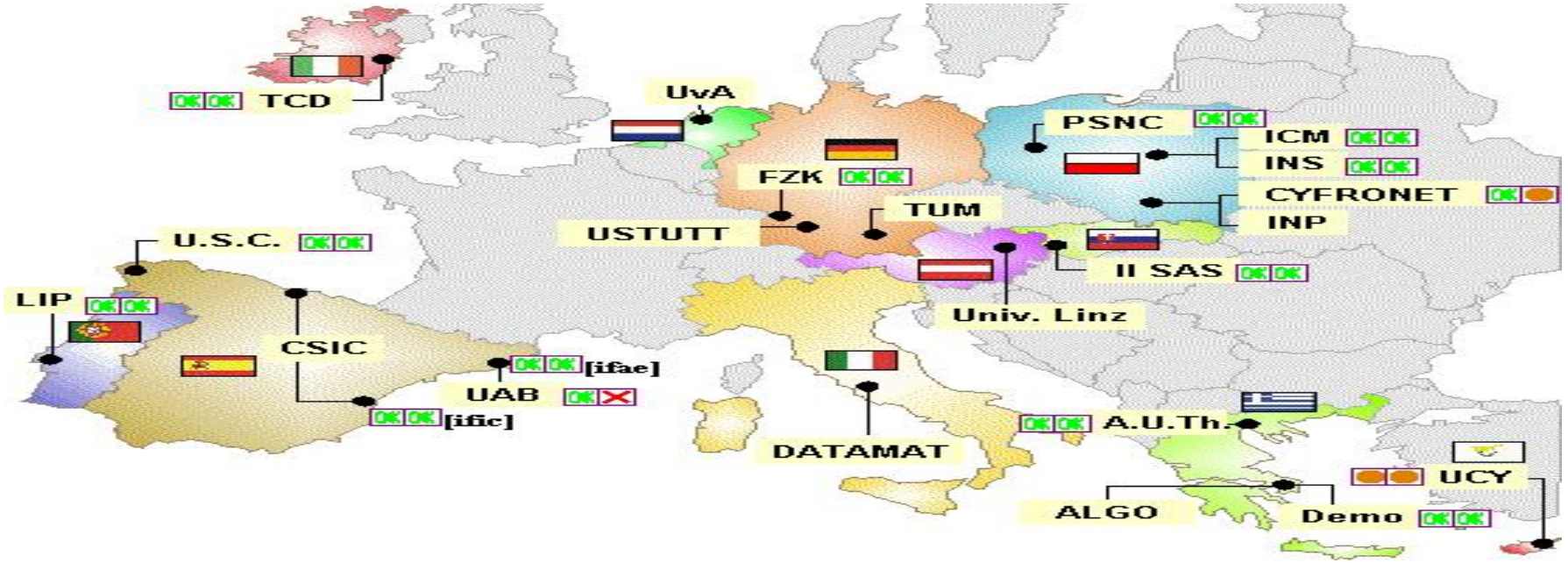
WP4

Status of the CrossGrid Testbed
EDG WP6 Meeting, Barcelona



Jesús Marco (marco@ifca.unican.es)
Instituto de Física de Cantabria, IFCA
Consejo Superior de Investigaciones Científicas, **CSIC,** Santander, SPAIN
grid.ifca.unican.es/crossgrid/wp4

First Prototype of the CrossGrid Testbed

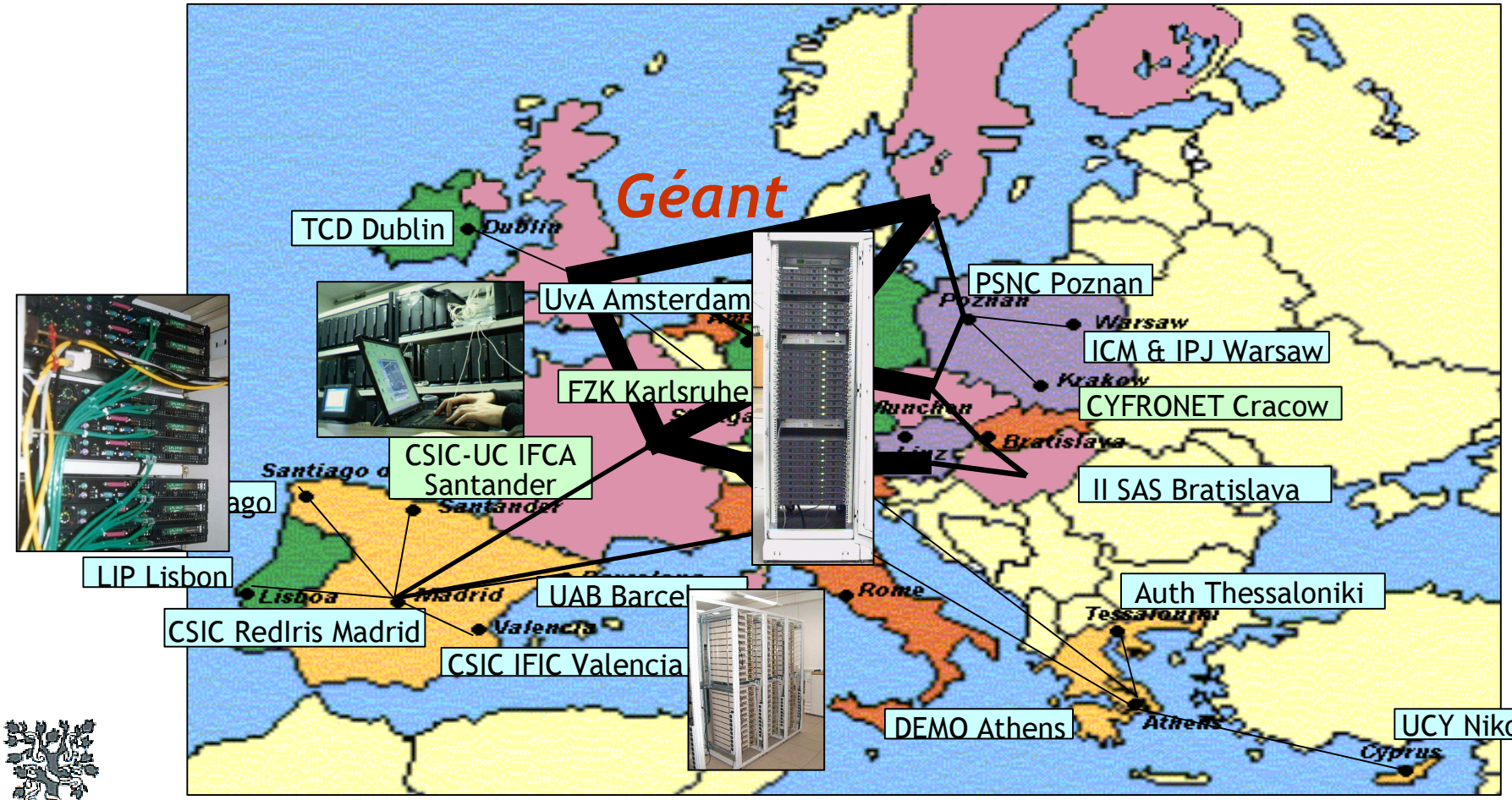
J.Gomes[1], M.David[1], J.Martins[1], L.Bernardo[1], J.Marco[2], R.Marco[2], D.Rodríguez[2], J.Salt[3], S.Gonzalez[3], J.Sánchez[3], A.Fuentes[4], M.Hardt[5], A.García[5], P.Nyczyk[6], A.Ozieblo[6], P.Wolniewicz[7], M.Bluj[8], K.Nawrocki[9], A.Padee[8,9,10], W.Wislicki[8,9], C.Fernández[11], J.Fontán[11], A.Gómez[11], I.López[11], Y.Cotronis[12], E.Floros[12], G.Tsouloupas[13], W.Xing[13], M.Dikaiakos[13], J.Astalos[14], B.Coghlan[15], E.Heymann[16], M.Senar[16], G.Merino[17], C.Kanellopoulos[18], and G.D.van Albada[19]

[1] Laboratório de Instrumentacão e Física de Partículas, Lisbon, Portugal
[2] Instituto de Física de Cantabria (CSIC),Santander, Spain
[3] Instituto de Física Corpuscular(CSIC),Valencia, Spain
[4] RedIris(CSIC),Madrid, Spain
[5] Forschungszentrum Karlsruhe GMBH, Germany
[6] Akademickie Centrum Komputerowe CYFRONET, Krakow, Poland
[7] Poznan Supercomputing and Networking Center, Poznan, Poland
[8] A.Soltan Institute for Nuclear Studies, Warsaw, Poland
[9] Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw,Poland
[10] Instytut Radioelektroniki PW, Warsaw, Poland
[11] CESGA, Centro de Supercomputacion de Galicia, Santiago de Compostela, Spain
[12] National Center for Scientific Research "Demokritos", National and Kapodistrian University of Athens, Dep. of Informatics and Telecommunications, Greece
[13] University of Cyprus, Cyprus
[14] Ustav Informatiky Slovenska Akademia Vied, Bratislava, Slovakia
[15] Trinity College Dublin, Ireland
[16] Universitat Autonoma de Barcelona, Spain
[17] Institut de Fisica d'Altes Energies, Barcelona, Spain
[18] Aristotle University of Thessaloniki, Greece
[19] Universiteit van Amsterdam, Netherlands

The CrossGrid Testbed

WP4

- A collection of distributed computing resources
  - 16 sites (small & large) in 9 countries, connected through ***Géant*** + NRN
  - + Grid Services: EDG middleware (based on Globus) RB, VO, RC…



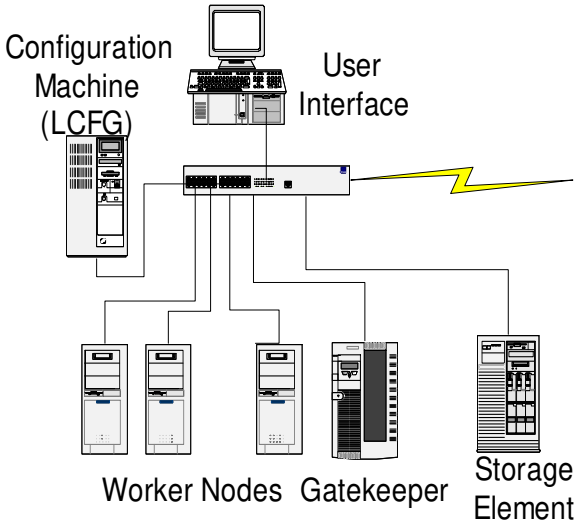
Human resources

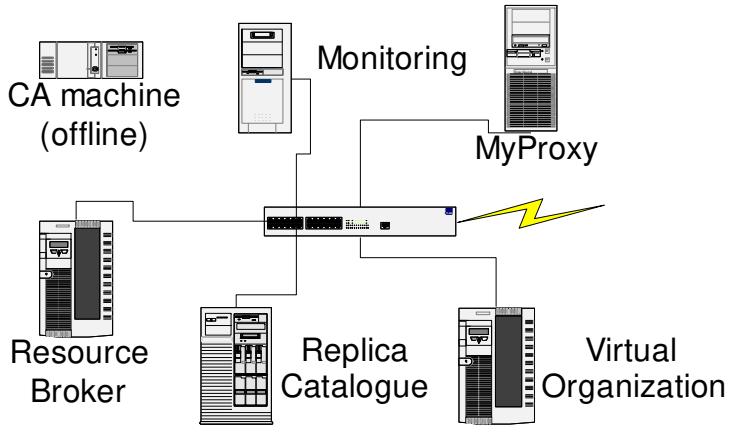
"An exceptional integration & site support team"

LIP Jorge, Mario
FZK Marcus, Ariel
IFIC Javier, Santi
IFCA Rafa
CESGA/USC Carlos, Javier
UAB Gonzalo
AuTH Christos
DEMO Vangelis
IISAS Jan
PSNC Pawel
CYFRONET Piotr
ICM/IPJ Adam, Michal
TCD Brian
UCY George

Computing resources

WP4





- Site testbed
  - LCFG configuration server
  - User Interface
  - Gatekeeper (Computing Element)
  - Worker Nodes
  - Storage Element
- 16 sites:
  - 115 CPUs (Worker Nodes)
  - 4 TB (Storage Elements)
- National Certification Authority machines
- Grid services (LIP)
  - Information Index
    - Top MDS Information Server, points to site Information Servers
  - Resource Broker
    - Matchmaking and load balancing scheduler
  - Replica Catalogue
    - Database for physical replica file location
  - Certificate Proxy Server
    - Short lived certificates for long lived processes, used by RB
  - Virtual Organization Server
    - Database for user authentication (CROSSGRID VO)
  - Monitoring
    - Mapcenter: network monitoring system

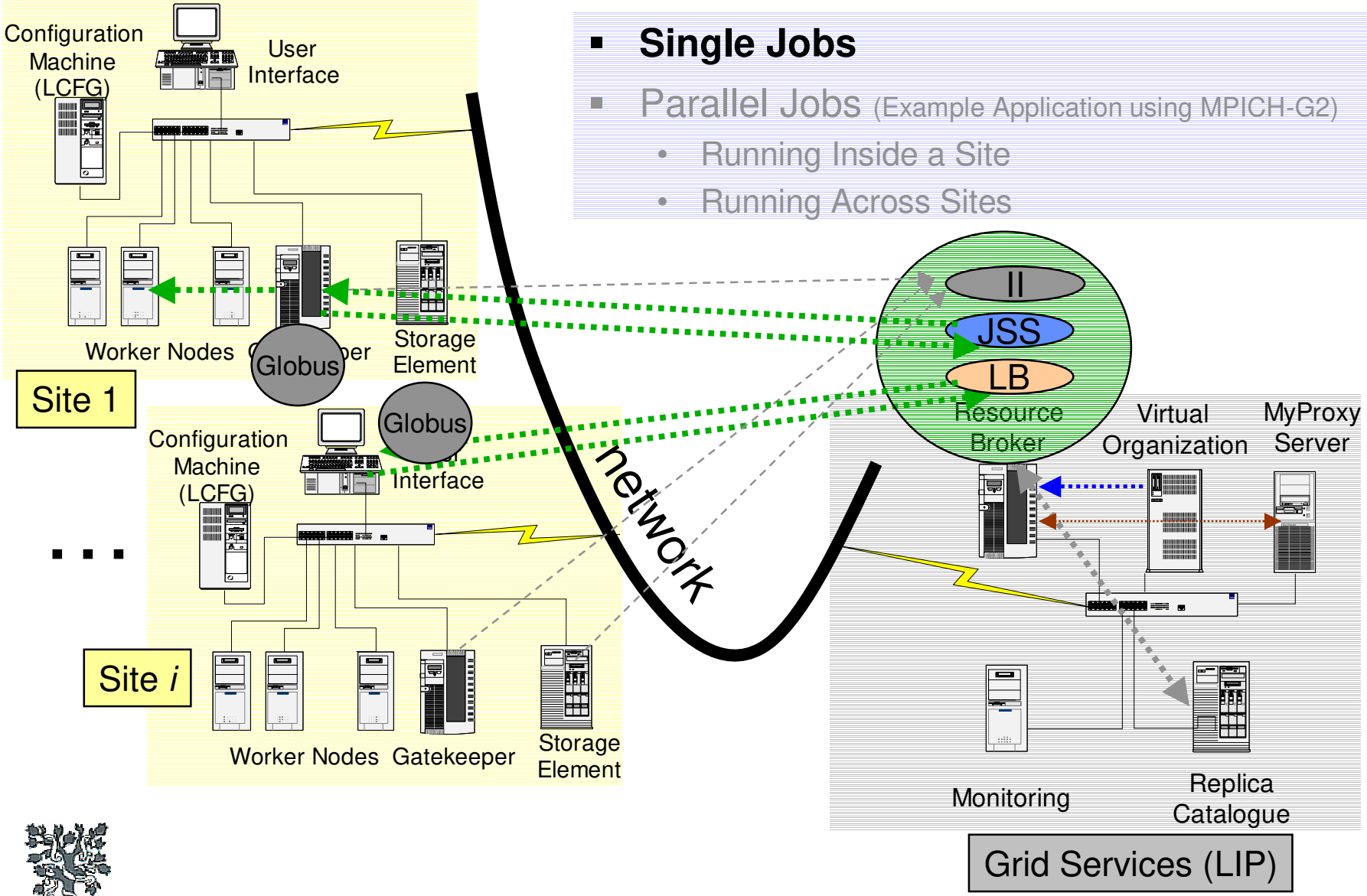
Testbed Status

Site	Location	Status
CYFRONET	Cracow	Running 1.4.4, Deploying 1.4.8
ICM	Warsaw	Running 1.2.2, Deploying 1.4.8
INS	Warsaw	Running 1.2.2, Deploying 1.4.8
UvA	Amsterdam	Deployed 1.2.2.
FZK	Karlsruhe	Running 1.4.8
IISAS	Bratislava	Running 1.2.3.
PSNC	Poznan	Running 1.2.2.
UCY	Nikosia	Deploying 1.4.8
TCD	Dublin	Running 1.2.3, Deploying 1.4.11
CSIC/IFIC	Valencia	Running 1.2.3, Deploying 1.4.8
CSIC/IFCA	Santander	Deploying 1.4.8
UAB	Barcelona	Deploying 1.4.8
USC/CESGA	Santiago	Running 1.2.2,Deploying 1.4.8
Demokritos	Athens	Running 1.4.8
AUTH	Thessaloniki	Running 1.2.2.
LIP	Lisbon	Running 1.4.11

Using the Testbed

WP4

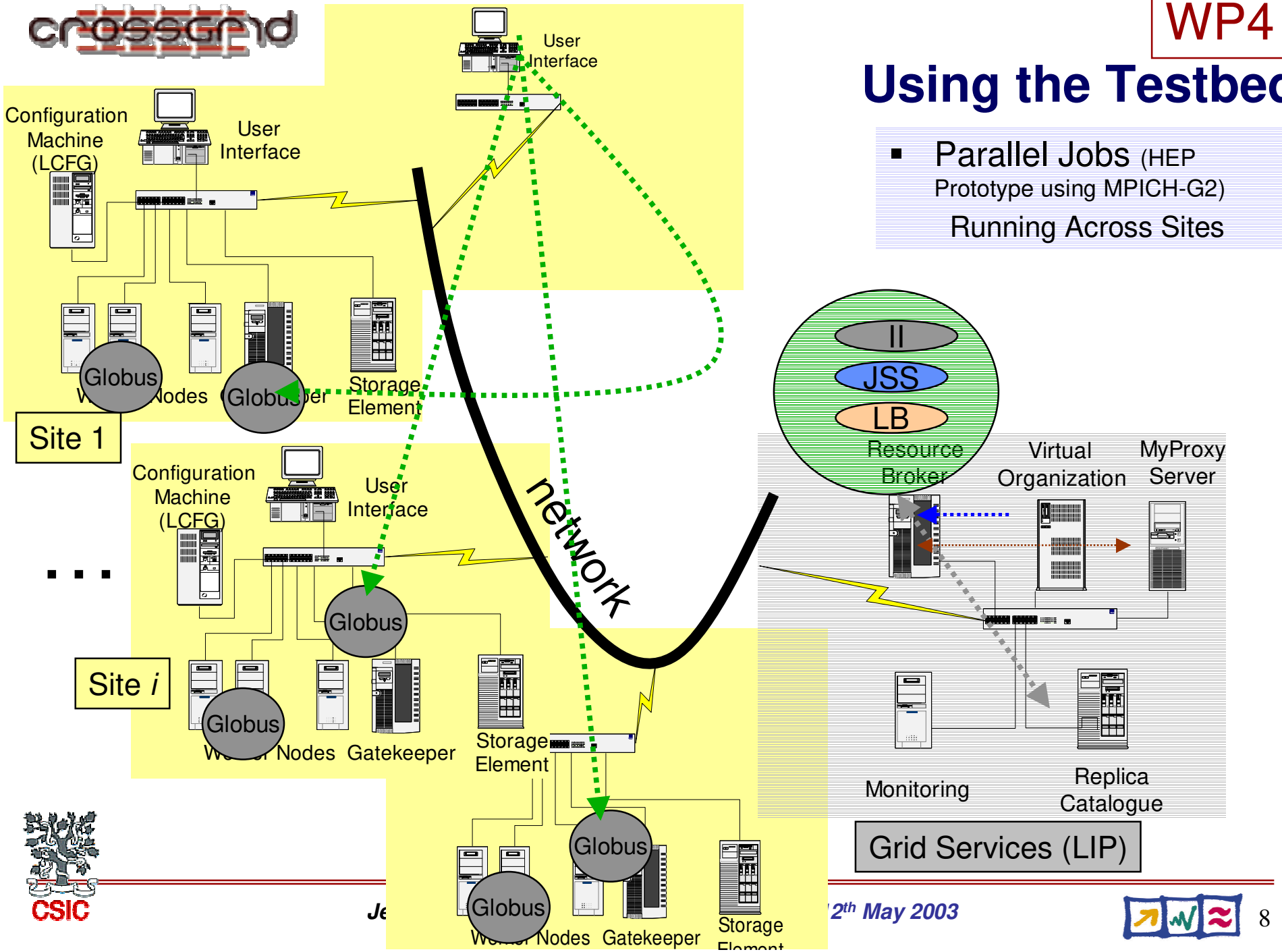
- **Single Jobs**
- Parallel Jobs (Example Application using MPICH-G2)
  - Running Inside a Site
  - Running Across Sites



WP4

Using the Testbed

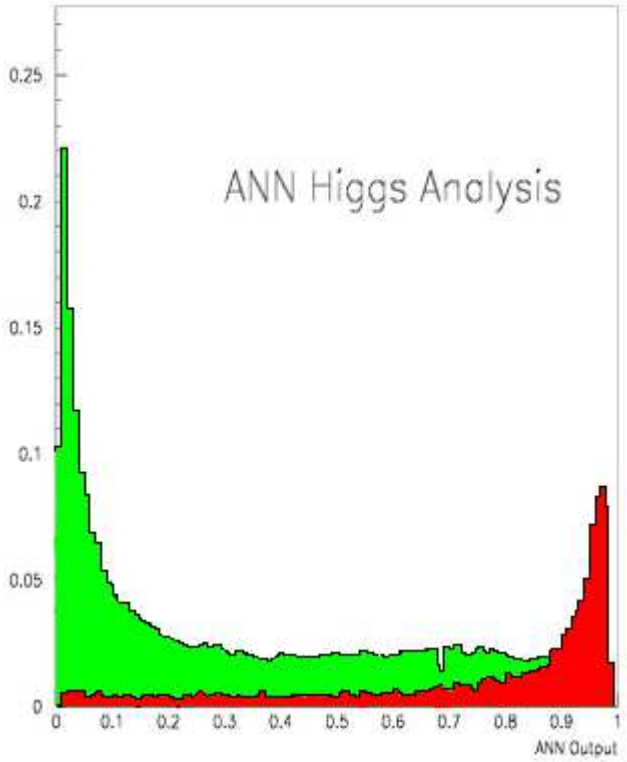
- Parallel Jobs (HEP Prototype using MPICH-G2) Running Across Sites



CrossGrid WP1 – Task 1.3 Distributed Data Analysis in HEP

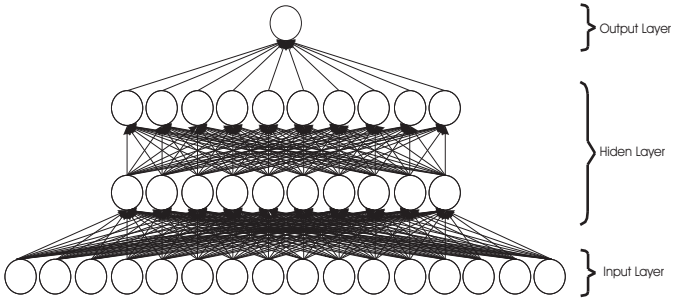
Subtask 1.3.2: Data-mining techniques on GRID

- ANN: example of architecture 16-10-10-1
  - 16 input variables
  - 2 hidden layers with 10 nodes each
  - 1 output layer, 1=signal, 0=background
- Trained on MC sample
  - Higgs generated at a given mass value
  - All types of Background
  - 10x real data statistics
- Applied on real collected data to order in S/B the candidates to Higgs boson
- Training process:
  - Minimize classification "error"
  - Iterative process
  - No clear "best strategy"
  - Computing Intensive: hours to days for each "try"



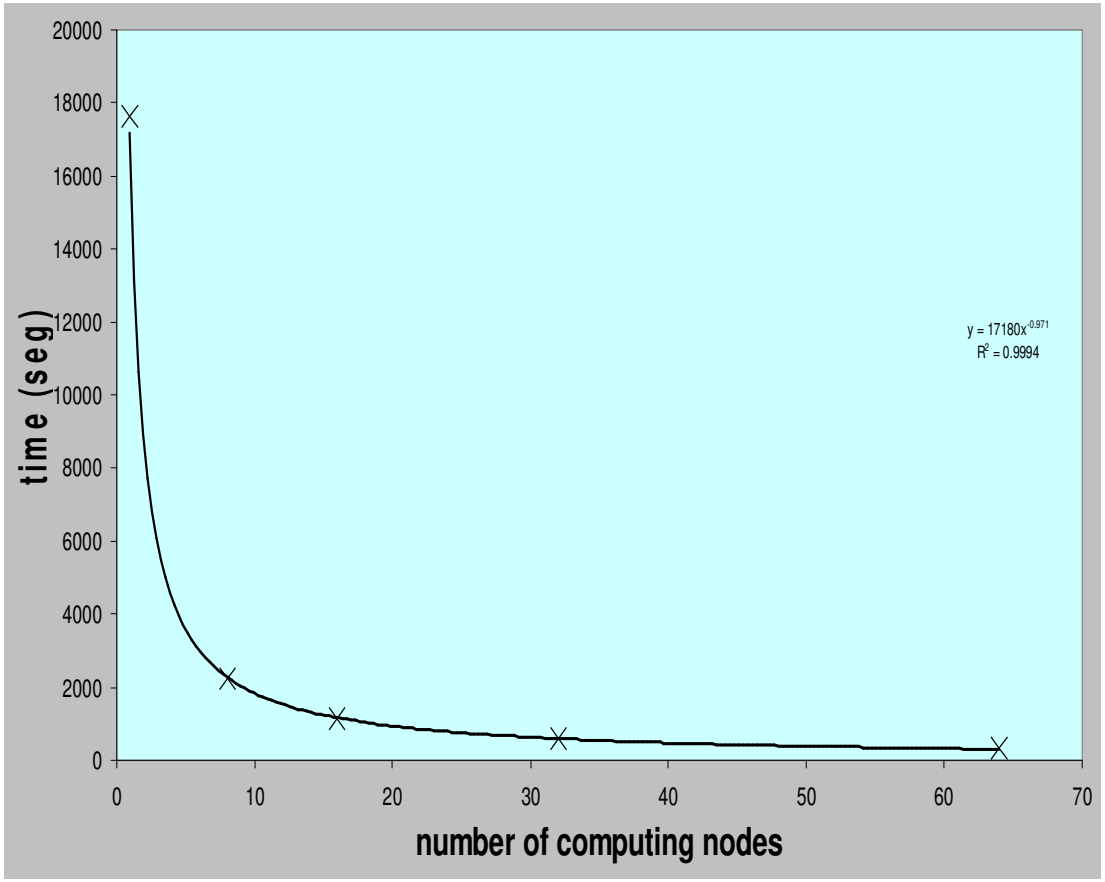
Distributed Training Prototype

- Distributed Configuration:
  - Master node and N slave nodes.
    - Scan to filter events & select variables
    - ResultSet in XML, split according to N (number of slave nodes)
- Training procedure:
  - Master reads input parameters and sets the initial weights to random values.
  - The training data is distributed to the slaves.
  - At each step:
    - The master sends the weights to the slaves.
    - The slaves compute the error and the gradient and return them to the master.
- This training procedure has been implemented using MPI and adapting the MLP-fit package.
- Conditions:
  - train an ANN with 644577 simulated realistic LEP events, 20000 of them corresponding to signal events.
  - Use a 16-10-10-1 architecture (270 weights)
  - Need 1000 epochs training.
  - Similar sized samples for the test.
  - BFGS training method.



Execution and results on a local cluster

First prototype on local cluster with MPI-P4



Scales as 1/N

644577 events,
16 variables
16-10-10-1 architecture
1000 epochs for training

Time reduction:
from **5 hours**
down to **5 min**
using 60 nodes!

Modelling including latency <300 ms needed !

Running with MPICH-G2 in a local cluster

- Migration to MPICH-G2 required
  - Installation of certificates for users and machines
  - Globus 2 installation in cluster
  - Program rebuilt, statically linked
  - Installation of software in the CVS repository at FZK
  - Use of globus-job-run, resource allocation through `.rsl` file
- **DEMO (shown in Santiago, CrossGrid Workshop)**
  - Running in local cluster, comparing two configurations with
    - 1 node (master+slave)
    - 20 nodes (1 master+20 slaves)
  - Certificates for authentication
  - Graphics shown:
    - Basic: ERROR EVOLUTION WITH TRAINING PROGRESS (number of iterations or epochs)
    - Signal-Background separation: NN output (classification) vs Discriminating variables
      - AQCD = Event shape
      - BTAG = Particles Lifetime
      - PWW, PZZ = Mass reconstruction

ann_compara.exe

M12 objective reached !!!

Running in the CrossGrid Testbed

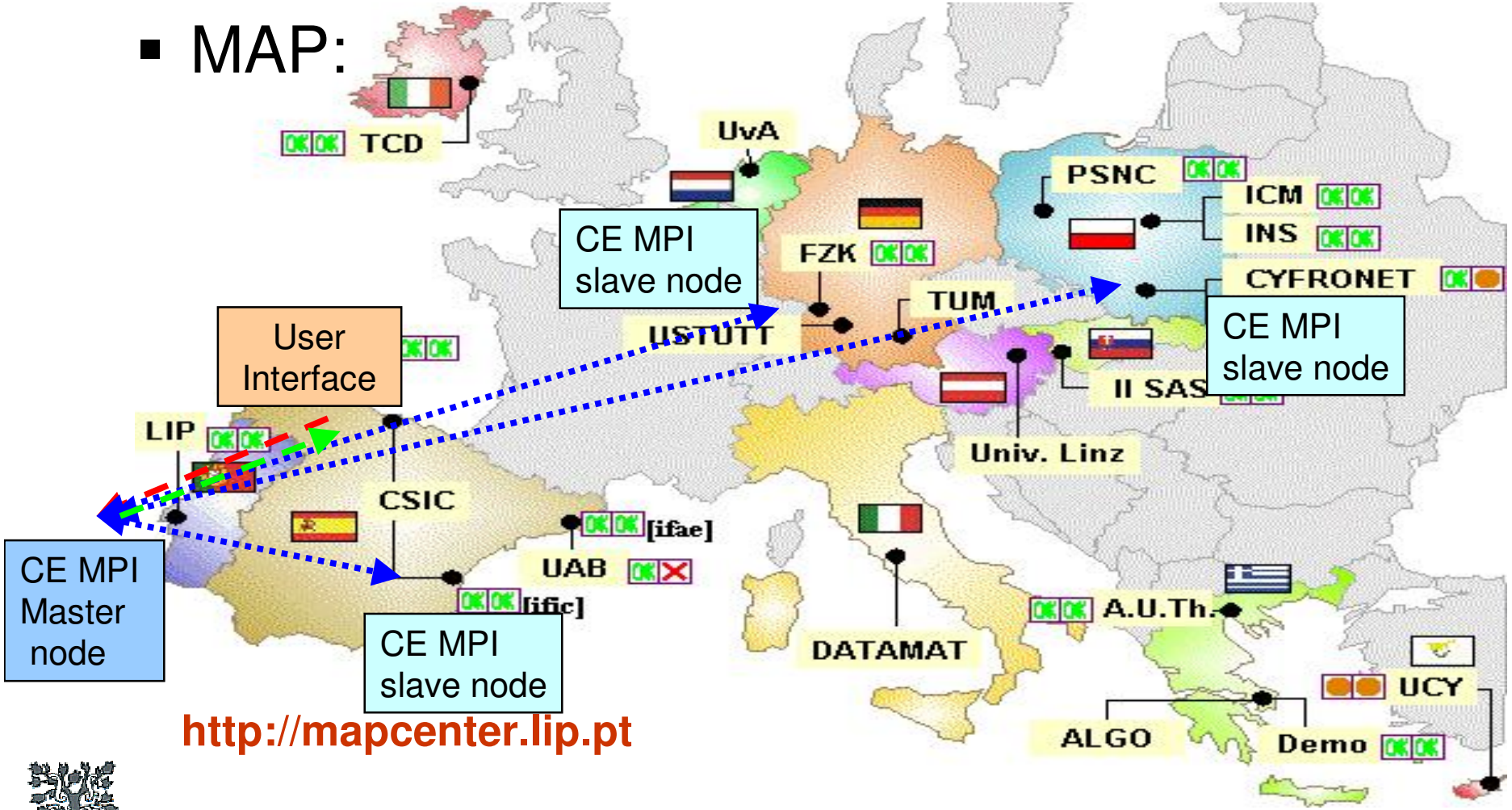
INTEGRATION AND DEPLOYMENT: Objective for these months!

- Steps:
  - User (with certificate) included in CrossGrid VO, logs in User Interface machine
  - Resource Allocation:
    - Direct .rsl file
    - Need public IP
  - Job Submission:
    - Copy executables + input files to each node via an script with Globus tools
    - Submit as before (globus-job-run)
  - Output:
    - Graphical output via X11
    - NN (weights) in XML format
- **DEMO (also shown in Santiago, CrossGrid Workshop):**
  - Running in testbed:
    - User Interface in Santander, Master node in CE at LIP
    - Slaves at Valencia (IFIC), Karlsruhe (FZK), Krakow (CYFRONET)

ann_in_testbed.exe

DEMO IN TESTBED

- MAP:



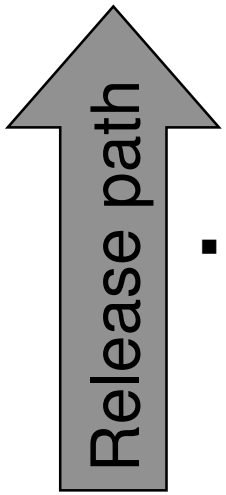
More related integration work …

- TOOLS (**DEMO**):
  - MPI Verification using MARMOT
    - Compilation with MARMOT
    - Running in the testbed
  - GPM & OCM-G
    - Monitoring

- GRID SERVICES:
  - ACCESS TO TESTBED RESOURCES **(DEMO)**
    - Use ROAMING ACCESS SERVER Via Portal or via Migrating Desktop
      - File transfer possibilities
  - JOB SUBMISSION (**DEMO**):
    - Job “parameters”: build XML form and translate into JDL
    - Submission for single node using JSS
      - Migrating Desktop
      - Portal
    - Output:
      - Graphical output via X11 (tunnelled), or using SVG
  - Network tracing with SANTA-G
      - Dump and analysis of packets

Testbed Evolution

- Initial testbed
  - Four initial sites, EDG 1.2, in July
  - Deployment of Grid Services at LIP in September
    - CrossGrid Virtual Organization server
    - Resource Broker, Replica Catalogue
- Current status:
  - **Production (Stable) testbed**
    - Objective: support applications
    - All sites, now with EDG 1.2.2/3, migrating to 1.4.8+ RH6.2+LCFG
  - **Validation testbed**
    - Objective: validation of new production middleware
    - LIP+FZK+DEMO, EDG 1.4.8+ RH6.2+LCFG
- Coming:
  - **Development testbed**
    - Objective: support development of new middleware and applications
    - Used in the integration process
    - Plan was EDG 2.0, now should be 1.4.8 but prefer RH7.3+LCFG-ng
    - Modified Resource Broker (MPI resources “finder”)



Production Testbed Use

WP4

Site Statistics

Sites	Connect	Pings	Jobs OK	Failed Jobs			
				LCAS	CRL exp	Jobman	GSS
LIP	6556	462	2836	50	17	92	3099
IFIC	5326	655	2649	100	97	45	1780
Cyfronet	4516	306	2522	0	20	111	1557
II SAS	1404	6	1185	0	15	99	99
FZK	1799	11	1112	118	7	123	428
Demo	9481	5	1111	36	0	51	8278
ICM	705	34	604	8	24	2	33
CESGA	7321	1	544	78	28	13	6657
UAB	600	14	519	0	9	14	44
INS	592	2	517	6	20	20	27
PSNC	582	0	496	15	14	11	46
TCD	145	0	131	0	0	2	12
AUTH	141	0	127	0	3	0	11
TOTAL	**39168**	**1496**	**14353**	**411**	**254**	**583**	**22071**

Resource Broker Statistics

Total users	33
Jobs submitted	**1943**
Jobs accepted	1904
Jobs with good match	1799
Jobs submitted by JSS	1781
Jobs run	1620
Jobs done	**1070**

Since the RB doesn't support parallel jobs, most job submissions pass unnoticed to the RB.

Validation Testbed Use

WP4

Site Statistics

Sites	Connections	Pings	Jobs OK	Failed Jobs			
				LCAS	CRL exp	Jobman	GSS
LIP	67365	2319	64995	21	0	4	26
FZK	8883	64	8671	38	12	50	48
Demo	10665	0	6170	4	6	2	4483
TOTAL	**86913**	2383	**79836**	63	18	56	4557

Resource Broker Statistics

Total users	8
Jobs submitted	**4173**
Jobs accepted	4173
Jobs with good match	4010
Jobs submitted by JSS	4007
Jobs run	3964
Jobs done	**3954**

95% success !!

IST Demonstration

- CrossGrid has participated in the World Grid demonstration involving European and US sites from CrossGrid, DataGrid, GriPhyN and PPDG.
- It took place in November 2002.
- It was the largest grid testbed in the world.
- Applications from the CERN/LHC experiments CMS and Atlas were used.
- CrossGrid participated with 3 sites:
  - LIP - Lisbon
  - FZK - Karlsruhe
  - IFIC - Valencia

Coordination

- Collaborative Work:
  - 17 VRVS meetings
    - keep regular contact, share info
    - quite an effort!
  - Several presential meetings
    - Cracow
    - CERN (april, july, october)
    - Linz
    - Santiago
- Mailing lists: *crossgrid-wp4@lists.cesga.es*
- WP4 Web pages
  - *http://grid.ifca.unican.es/crossgrid/wp4*
- First publication (ACROSSGRID Workshop)
- **Integration Team**

Deliverables used to trigger and organize work

Testbed setup

- **Installation of testbed sites and middleware deployment**
- Certification Authorities:

Country / Institution	CrossGrid	DataGrid
Poland / PSNC	Deployed and accepted	Accepted in October
Netherlands / NIKHEF	Accepted	Already in DataGrid
Germany / FZK	Deployed and accepted	Accepted in June
Slovakia / IISAS	Deployed and accepted	Accepted in October
Ireland /TCD	Accepted	Already in DataGrid
Spain / CSIC	Accepted	Already in DataGrid
Greece / AUTH	Deployed and accepted	Accepted in October
Portugal / LIP	Accepted	Already in DataGrid
Cyprus / UCY	Deployed and accepted	Accepted in January

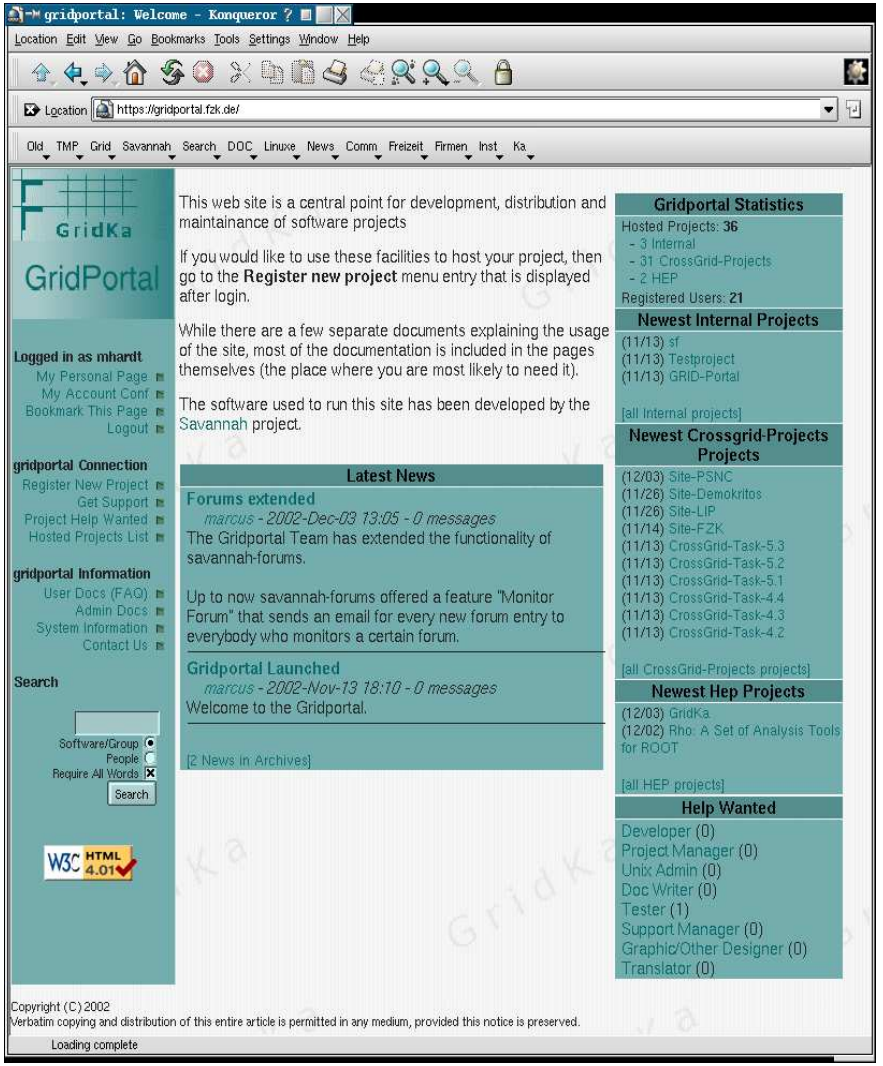
http://grid.ifca.unican.es/crossgrid/wp4/ca

Infrastructure Support

- **Software repository**

http://gridportal.fzk.de

- Customized GNU Savannah (based on SourceForge)
- CVS browsable repository
- Main current usage:
  - ca. 35 web-visits per day (=1000 hits)
  - 7000 files, 356MB, 850.000 code-lines, 15.000 doc-lines + 174 doc/pdf-files
  - 44 X#-RPMs in the download repository



Infrastructure Support

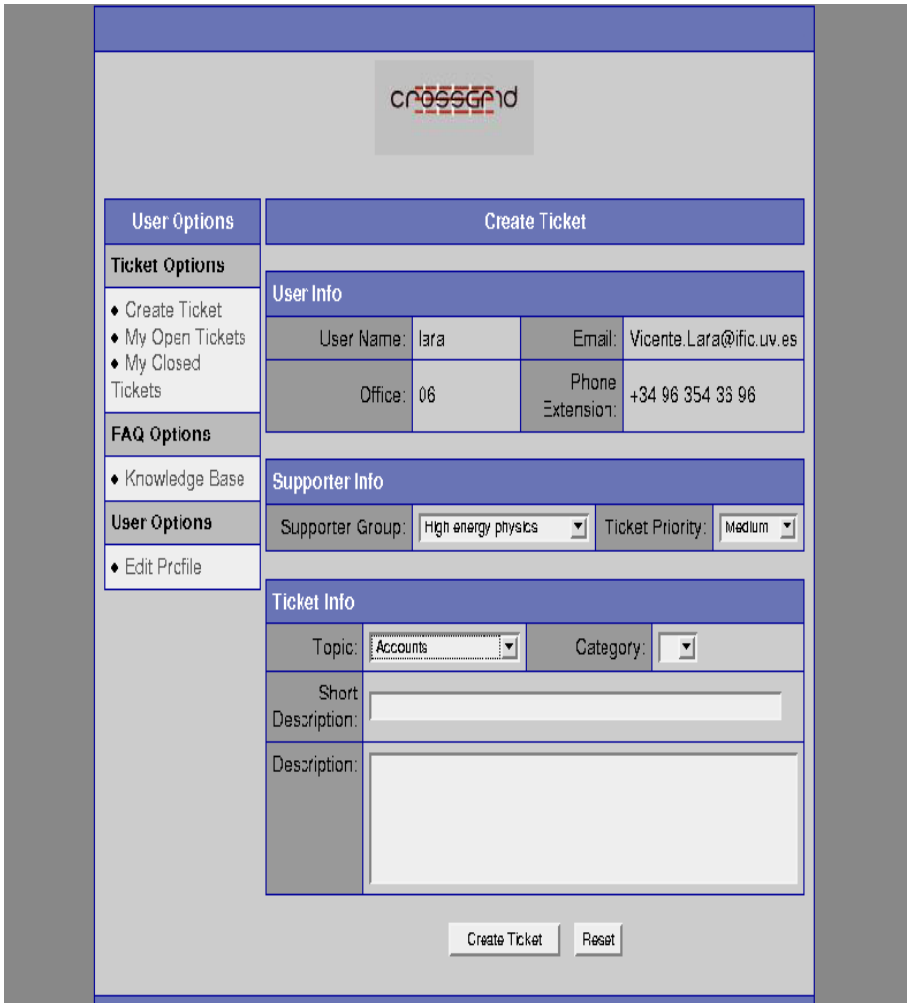
- **HelpDesk**

  http://cg1.ific.uv.es/hlpdesk/

- Question-Answer Mechanism

(follow the evolution of question via tickets)

- Unified knowledge database with EDG
- Interacion levels:
  - User
  - Supporter
  - Administrator
- Based on **OneOrZero v1.4 RC2**
  - distributed under GPL
  - a web based helpdesk system incorporating PHP, MySQL Javascript

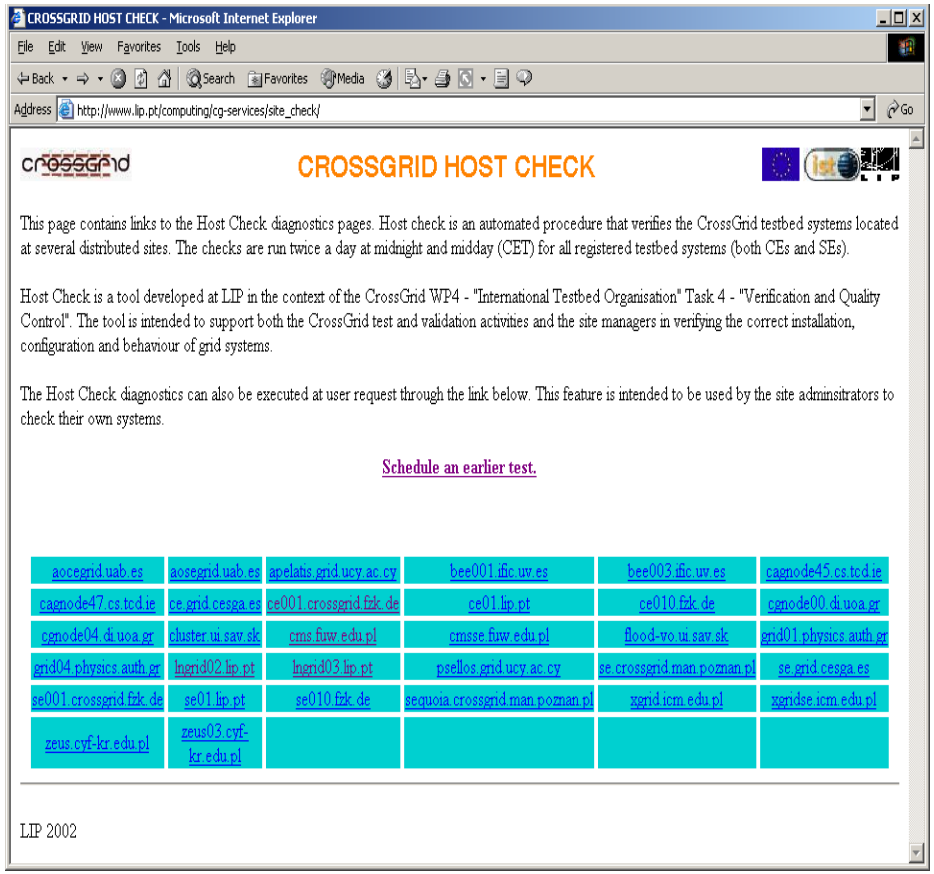


Verification and Quality Control

- **Test and Validation Testbed**
- First Tests of Applications
  - HEP prototype using MPICH-G2
- Support:

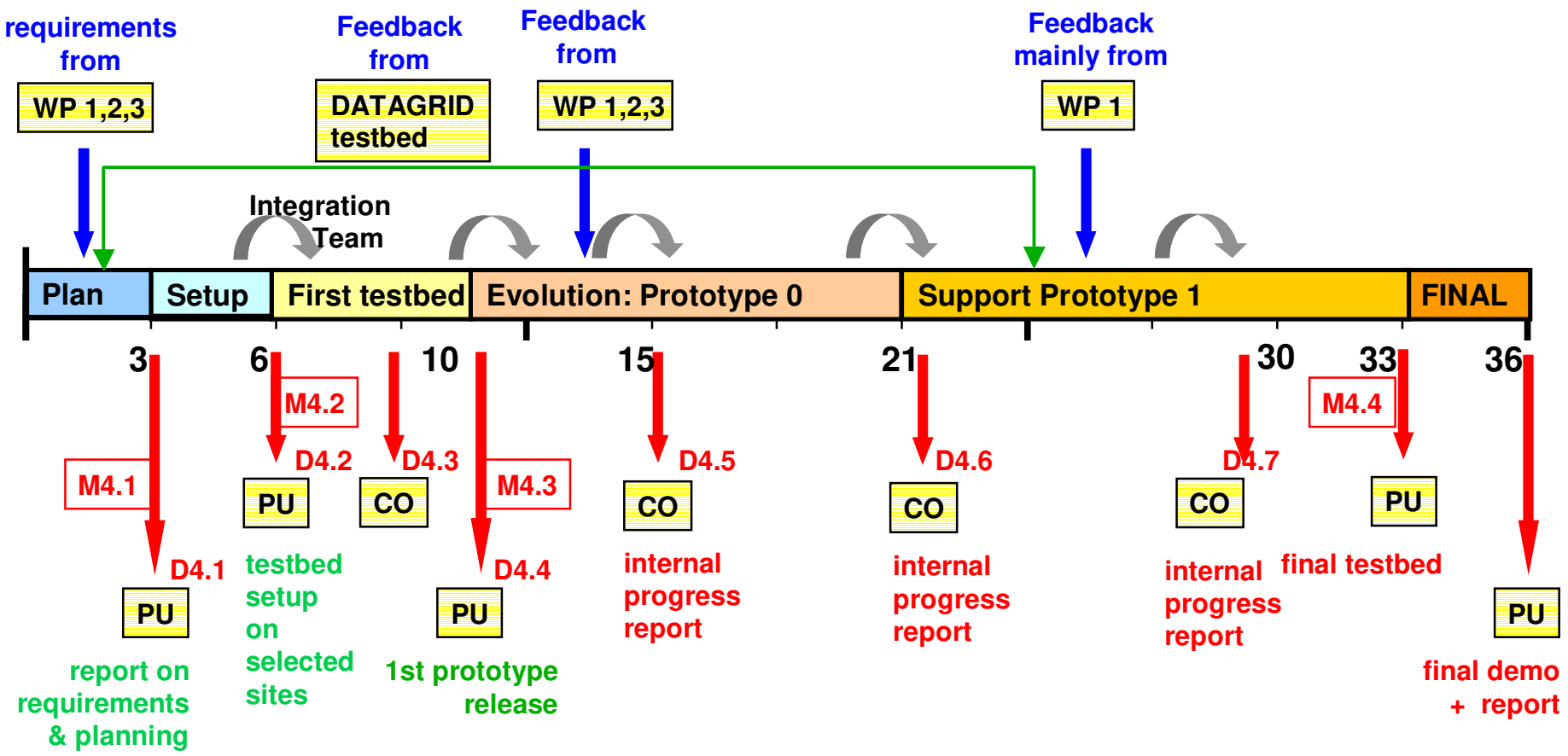
  grid.support@lip.pt

  http://www.lip.pt/computing/cg-tv-services
- Usage Statistics
- Tools:
  - Monitoring (mapcenter)
  - CrossGrid Host Check Tool

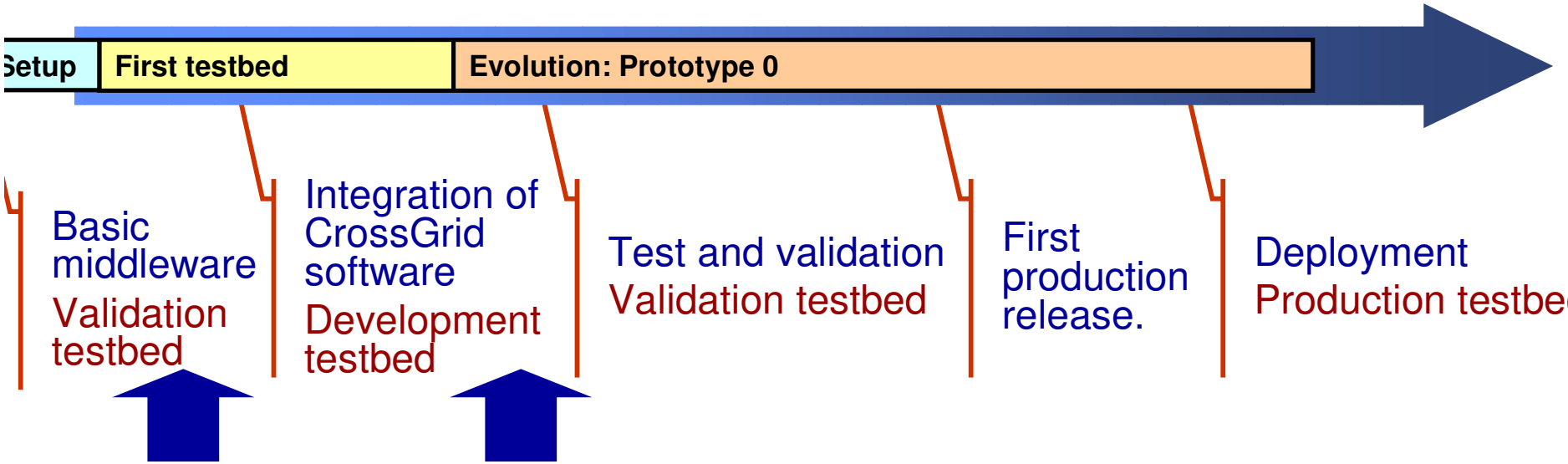


http://www.lip.pt/computing/cg-services/site_check

Planning


requirements from
WP 1,2,3
Feedback from
DATAGRID testbed
Feedback from
WP 1,2,3
Feedback mainly from
WP 1
Integration Team
Plan
Setup
First testbed
Evolution: Prototype 0
Support Prototype 1
FINAL
3
6
10
15
21
30
33
36
M4.1
D4.1
PU
report on requirements & planning
M4.2
D4.2
PU
testbed setup on selected sites
D4.3
CO
M4.3
D4.4
PU
1st prototype release
D4.5
CO
internal progress report
D4.6
CO
internal progress report
D4.7
CO
internal progress report
M4.4
PU
final testbed
PU
final demo + report

Software Integration and Testbed Release



Basic middleware
EDG 1.4.3 Globus 2

CORE WP4 Integration Team

- Development tools
  - MPI verification
  - Benchmarking
  - Performance prediction
  - Monitoring
- Middleware:
  - RAS, Portal, Migrating Desktop
  - Parallel Interactive Scheduler
  - Monitoring
  - Data Access Optimization
- Application Prototypes

CrossGrid Integration Team

- WP2 + WP4 people
- WP3 + WP4 people
- WP1 + WP4 people

Integration Work in Santiago

WP2 + WP4

Developer Workstation (On top of UI, RH6.2), first at CESGA

- MARMOT (Bettina, Rafa) RPM? test? Difficulties?
- GRIDBENCH (George, Christos)
- PERFORMANCE PREDICTION (Fran, Carlos+Javier)
- GPM (Roland+Wlodek,Piotr) installed at Munich
- Processing of monitoring data (Adam)

WP3 + WP4

- RAS+Migrating desktop (Pawel, Marcin) server(s) at PSNC
- Portal Application Server (Yannis,Angelos,Jesus,Dani,Javier,Javier)
- JSS (Stephano+support from Jorge) JSS server at PSNC (in another machine)
- Modified RB (Alvaro,Santi,Javier) RB server at IFIC, and at CESGA
- SANTA-G (David,Mario) RGMA server at LIP, tcpdump where?
- OCM-G (Roland, Bartosz, Ariel) machine at CYFRONET?
- Integrating reports for Jiro (Slawek, Marcus), Data Access Optimization (Lukasz,Piotr)

Integration Work in Santiago

WP1 + WP4

HEP prototype work (Jorge, David, Celso):

Three types of execution modes: (a) “EDG-like”, (b) MPI inside cluster, (c) MPI across sites

- (a) Full chain: portal/migrating desktop+ JSS, GRAPHICS?
- (b) Submission using modified procedure, could we send from Portal? Graphics?
- (c) Same question

FLOODING CONTROL APPLICATION

- Installation at external cluster started (Jan, Viet, Rafa)

METEO APPLICATION

- 1.4.3 Air pollution: Installation at testbed started (Jose Carlos , Carlos+Javier)
- 1.4.2 Mesoscale + Datamining: Installation at testbed started (Antonio, Rafael,Daniel)
- 1.4.1 Meteo (Bogumil , Michal+ Adam)

BIOMEDICAL APPLICATION

- Installation at cluster started (Alfredo , Dick)

Work for next months

- Maintain testbed stability:
  - Evolution of production testbed to 1.4.4 these weeks
- **Work on integration, establish the development testbed**
- Installation: test LCFGng and *support clusters*
- Prepare migration to EDG 2 and RH 7.3, *and CERN LCG-1*
- Support the extension to new sites.
  - More sites internal to the project (Linz)
  - Possible external sites and users (policy needed).
- **Study the usage of QoS in CrossGrid**
  - Create a QoS test infrastructure on network and computing resources
- **Start the security group activities**
  - Policies, guidelines, tracking of problems, patches.
- Stress testing of the infrastructure.

DataGrid / CrossGrid Spanish CA

working on improvements

- Web Access , Unix command access
  - Users: For making requests, getting certificates
  - RAs: To approve/reject requests
- Security:
  - CA is kept offline
  - RAs online: https (use certificates)+ password + code card
- Based on Apache + PHP + MySQL
- Layered structure programming to allow easy modifications.
- To be presented at next CA managers meeting

Collaboration with DataGrid

- Tracking EDG releases
- Collaboration on:
  - **Helpdesk and support**
  - **Software repository and autobuild**
  - ***Installation and large cluster integration***
  - **Test and validation with the EDG testgroup**
  - ***Network QoS issues***
  - **Participation in Security group (CA)**
- Advice on migration?

Conclusions

- The *CrossGrid testbed is operative*
  - Regular tests and use, including direct MPI execution
- Evolution
  - Testing and validation for:
    - Applications (WP1)
    - Programming environment (WP2)
    - New services & tools (WP3)
  - Emphasis on interoperability with EU-DataGrid (EDG)
  - Extension of GRID across Europe
  - Advice to follow GT3 while need to follow LCG…
  - to be analyzed!