



## DELIVERABLE D4.4 APPENDIX A

### TESTBED EXTENSION AND SITE STATUS

#### WP4

---

Document Filename:	<b>CG-4.4-D4.4-v1.0-LIP013-CrossGridTestbed.doc</b>
Work package:	<b>WP4</b>
Partner(s):	<b>LIP</b>
Lead Partner:	<b>LIP</b>
Config ID:	<b>CG-4.4-D4.4-v1.0-LIP013-CrossGridTestbed</b>
Document classification:	<b>PUBLIC</b>

---

#### Abstract:

This document describes the status of the CrossGrid *production* and *test and validation* testbeds.



### Delivery Slip

	Name	Partner	Date	Signature
<b>From</b>				
<b>Verified by</b>				
<b>Approved by</b>				

### Document Log

Version	Date	Summary of changes	Author
1-0-DRAFT-A	26/12/2002	Draft version	Jorge Gomes
1-0-DRAFT-B	15/01/2003	Draft version	Jorge Gomes
1-0-FINAL	19/02/2003	Final version after review	Jorge Gomes

## CONTENTS

<b>1</b>	<b>INTRODUCTION</b> .....	<b>5</b>
<b>2</b>	<b>ACRONYMS AND ABBREVIATIONS</b> .....	<b>6</b>
<b>3</b>	<b>TESTBED STATUS</b> .....	<b>7</b>
3.1	PRODUCTION SITES .....	7
3.1.1	<i>Sites statistics</i> .....	11
3.2	PRODUCTION CENTRAL SERVICES .....	18
3.2.1	<i>Production Resource Broker</i> .....	19
3.2.2	<i>Production Replica Catalogue</i> .....	22
3.2.3	<i>Production VO server</i> .....	23
3.2.4	<i>Production MyProxy</i> .....	24
3.2.5	<i>Production Mapcenter</i> .....	24
3.3	TEST AND VALIDATION SITES.....	27
3.3.1	<i>Sites statistics</i> .....	28
3.4	TEST AND DEVELOPMENT CENTRAL SERVICES .....	30
3.4.1	<i>Test and validation Resource Broker</i> .....	31
3.4.2	<i>Test and Validation Replica Catalogue</i> .....	33
3.4.3	<i>Test and validation VO server</i> .....	34
3.4.4	<i>Test and validation MyProxy</i> .....	35
3.4.5	<i>Test and validation Mapcenter</i> .....	35
<b>4</b>	<b>TESTBED EXTENSION</b> .....	<b>37</b>
4.1	PRODUCTION TESTBED .....	37
4.2	TEST AND VALIDATION TESTBED .....	38
4.3	CENTRAL SERVICES .....	38
<b>5</b>	<b>TESTBED USAGE</b> .....	<b>40</b>
5.1	THE USER CERTIFICATE .....	40
5.2	RUNNING EXECUTABLES REMOTELY WITH GLOBUS .....	41
5.3	FILE TRANSFER .....	42
5.4	SUBMITTING JOBS WITH GLOBUS.....	44
5.5	SIMPLE JOB SUBMISSIONS WITH RSL .....	45
5.6	FILE COPY AND FILE ACCESS USING GASS .....	46
5.7	MPI PARALLEL JOB SUBMISSION.....	47
5.8	JOB SUBMISSION THROUGH THE RB.....	48
<b>6</b>	<b>CERTIFICATION AUTHORITIES</b> .....	<b>50</b>
<b>7</b>	<b>SUPPORT</b> .....	<b>51</b>
7.1	HELPDESK .....	51
7.2	VO SUPPORT.....	51
7.3	WEB PAGES .....	51
7.4	MAPCENTER .....	51
7.5	CVS .....	51
7.6	STATISTICS .....	52
7.7	HOST INSTALLATION AND CONFIGURATION VERIFICATION .....	52
<b>8</b>	<b>TESTBED DEPLOYMENT</b> .....	<b>53</b>
8.1	INSTALLING THE MIDDLEWARE.....	53
8.1.1	<i>LCFG</i> .....	53
8.2	CONFIGURING THE MIDDLEWARE .....	54
8.2.1	<i>CE/Gatekeeper</i> .....	54
8.2.2	<i>WN</i> .....	55
8.2.3	<i>SE</i> .....	55

---

8.2.4	UI .....	55
8.2.5	RB.....	55
8.2.6	MyProxy.....	55
8.2.7	RC.....	56
8.2.8	VO server .....	56
8.2.9	Mapcenter .....	56
<b>9</b>	<b>TESTBED MAINTENANCE .....</b>	<b>57</b>
9.1	VO MAINTENANCE .....	57
9.2	RB MAINTENANCE.....	57
9.3	RC MAINTENANCE.....	58
9.4	MAPCENTER MAINTENANCE .....	58
9.5	SITE MAINTENANCE.....	58
<b>10</b>	<b>SITE DEPLOYMENT PROBLEMS .....</b>	<b>59</b>
10.1	DNS AND REVERSE LOOKUPS .....	59
10.2	NUMBER OF PROCESSORS AND STORAGE SIZE .....	59
10.3	POLLED ACCOUNTS.....	59
10.4	CA CERTIFICATES .....	59
10.5	FIREWALLS .....	60
10.6	MDS CONFIGURATION.....	60
10.7	VO SETUP .....	60
10.8	GDMP SETUP .....	60
10.9	INPUT AND OUTPUT SANDBOX TRANSFER .....	60
10.10	LCFG AND SYSTEM BOOT.....	61
10.11	LCFG AND DISK PARTITIONS .....	61
10.12	RPM LISTS.....	61
10.13	SITE CONFIGURATION .....	61
10.14	MPI AND PRIVATE IP ADDRESSES .....	62
<b>11</b>	<b>DEVELOPMENTS AND CORRECTIONS.....</b>	<b>63</b>
11.1	CHECKING SITES AND SERVICES.....	63
11.2	STATISTICS .....	63
11.3	TEST LIBRARY FUNCTIONS.....	63
11.4	MAPCENTER .....	64
11.5	VO TOOLS.....	64
11.6	VOMS.....	64
11.7	LCFG .....	64
<b>12</b>	<b>TESTS.....</b>	<b>65</b>
12.1	BASIC TESTS .....	65
12.2	FIRST PROGRAM PROTOTYPES .....	65
12.3	FIRST INTEGRATION .....	66
12.4	IST DEMONSTRATION .....	66
12.4.1	<i>IFIC CE Usage statistics for the IST demonstration period.....</i>	<i>66</i>
12.4.2	<i>LIP CE Usage statistics for the IST demonstration period .....</i>	<i>66</i>
12.4.3	<i>FZK CE Usage statistics for the IST demonstration period .....</i>	<i>67</i>
<b>13</b>	<b>CONCLUSIONS .....</b>	<b>68</b>
<b>14</b>	<b>REFERENCE DOCUMENTS.....</b>	<b>69</b>

---

## 1 INTRODUCTION

CrossGrid will deploy separate testbeds for development, validation and production. However initially a single testbed prototype was built allowing the deployment efforts to be concentrated in a single infrastructure. The initial testbed was recently divided in two testbeds, a production testbed and a test and validation testbed.

The production testbed is the infrastructure where the first CrossGrid application prototypes are being tested and is based on EDG 1.2.2 and Globus middleware. The new test and validation testbed is based on EDG 1.4.3 and Globus. Both testbeds have been deployed and maintained in the context of the WP4 tasks 4.1 (International Testbed Organization) and 4.4 (Verification and Quality Control).

This document aims to describe the status of the two CrossGrid testbeds.

Many of the terms and concepts referenced in this document are explained more deeply in the “*CrossGrid Test and Validation Testbed Architecture*”, it is highly advisable to read this document prior to proceed.

---

## 2 ACRONYMS AND ABBREVIATIONS

ACL	Access Control List
API	Application programming interface
CA	Certification Authority
CASTOR	CERN Advanced Storage Manager
CE	Computing Element
CN	Common Name
CRL	Certificate Revocation List
CrossGrid	The EU CrossGrid Project IST-2001-32243
DataGrid	The EU DataGrid Project IST-2000-25182
DBMS	Database Management System
EDG	European DataGrid
FTP	File Transfer Protocol
HTTP	HyperText Transport Protocol
II	Information Index
JDL	Job Description Language
JSS	Job Submission Service
LB	Logging and Bookkeeping
LDAP	Lightweight Directory Access Protocol
LCFG	Local ConFiGuration system
LFN	Logical File Name
MDS	Meta Directory Services
NIC	Network Interface Card
NFS	Network File System
OU	Organizational Unit
PKI	Public Key Infrastructure
PFN	Physical File Name
QoS	Quality of Service
GDMP	Grid Data Mirroring Package
GID	Unix Group ID
GIIS	Grid Information Index Service
GRAM	Grid Resource Allocation Manager
GRIS	Grid Resource Information Service
GSI	Grid Security Infrastructure
RC	Replica Catalogue
RM	Replica Manager
RDBMS	Relational Database Management System
RB	Resource Broker
RFIO	Remote File I/O
SE	Storage Element
T&V	Test and Validation
UI	User Interface
UID	Unix User ID
VO	Virtual Organization
XML	Extensible Markup Language
WMS	Workload Management System
WN	Worker Node
WP	Work Package

### 3 TESTBED STATUS

The first CrossGrid testbed activities started at the project very beginning. At the time some partners were already involved in grid testbed deployment and research activities, namely through the involvement in the CERN LHC grid computing initiatives. They provided the seed for the initial CrossGrid testbed.

Since the CrossGrid middleware is still being developed and one of the CrossGrid goals is the compatibility with the DataGrid middleware, it was decided that initially the testbed middleware would have to be based entirely in DataGrid (EDG) and Globus middleware distributions. Delays in the release of a stable EDG testbed 1 middleware delayed the deployment of many CrossGrid sites. The efforts to establish an integrated CrossGrid testbed started with the release of EDG 1.2.0. However several problems were found in the first installed sites. The release of EDG 1.2.2 with improved stability has allowed more sites to join the initial CrossGrid testbed in spite of some middleware limitations.

Recently the initial testbed has given origin to two testbeds dedicated to production and test and validation activities. The CrossGrid production testbed uses EDG 1.2.2 and 1.2.3 while EDG version 1.4.3 has been deployed and tested in the test and validation testbed. It is expected that the EDG versions 1.4.x will overcome many of the major limitations of the previous releases with higher performance. Once found stable EDG 1.4.x will probably become the basis for the CrossGrid production middleware, allowing also the interconnection of both CrossGrid and DataGrid production testbeds.

#### 3.1 PRODUCTION SITES

At least 16 sites are foreseen to be included in the CrossGrid production testbed, this number may be increased with the addition of external sites interested in the technologies and applications being developed by CrossGrid, and with more sites owned by CrossGrid partners. Several organizations external to the project have already shown interest in joining the CrossGrid production testbed.

Of the 16 production sites foreseen 10 are now deployed and registered into the CrossGrid production Resource Broker (RB), 2 are deployed but have been temporarily moved to the test and validation testbed, 2 are deployed and almost ready to be added to the RB. Some of them are being tested and will be registered into the RB once the validation of their configuration is successfully completed. The 2 remaining sites are deployed but not yet available for testing.

The two sites temporarily in the test and validation testbed are involved in the test of the EDG 1.4.x middleware. EDG 1.4.x has some incompatibilities with the previous EDG 1.2.x middleware currently deployed in the production testbed. These incompatibilities prevent the registration of 1.4.x sites in the production RB. In spite of not being registered in the RB these two sites can still be used in the context of the production testbed to test applications that depend only on Globus middleware such as the first CrossGrid MPI application prototypes.

The next table shows the deployment status of the 16 CrossGrid production sites.

Site	Location	Status
CYFRONET	Cracow	Running 1.2.2, registered into the RB
ICM	Warsaw	Running 1.2.2, registered into the RB
INS	Warsaw	Running 1.2.2, registered into the RB
UvA	Amsterdam	Deployed 1.2.2, not tested
FZK	Karlsruhe	Running 1.4.3, temporarily in the test and validation testbed. Site available for production for Globus only applications.
IISAS	Bratislava	Running 1.2.3, registered into the RB
PSNC	Poznan	Running 1.2.2, registered into the RB
UCY	Nikosia	Deployed 1.2.2, under test
TCD	Dublin	Running 1.2.3, registered into the RB
IFIC	Valencia	Running 1.2.3, registered into the RB
IFCA	Santander	Deployed 1.2.2, not tested
UAB	Barcelona	Deployed 1.2.2, under test
USC/CESGA	Santiago	Running 1.2.2, registered into the RB
Demokritos	Athens	Running 1.4.3, temporarily in the test and validation testbed. Site available for production for Globus only applications.
AUTH	Thessaloniki	Running 1.2.2, registered into the RB
LIP	Lisbon	Running 1.2.3, registered into the RB

Table 1. CrossGrid production sites status.

According with the CrossGrid testbed architecture and minimum hardware requirements each site must have at least five system components.

- **Gatekeeper:** is the system that provides the gateway through which Jobs are submitted to local farm nodes.
- **Worker Node (WN):** is a local farm-computing node where jobs are actually executed. Jobs arrive to the WN through the local batch scheduling system. Many worker nodes can exist behind a single Gatekeeper.
- **Storage Element (SE):** is the generic name for any storage resource that includes a Grid interface ranging from large Hierarchical Storage Management Systems to disk pools.
- **User Interface (UI):** is used by end-users to submit jobs to the grid testbed through the resource broker, or directly to CEs using the Globus GRAM service.
- **LCFG installation server:** is used to install, configure and maintain the above systems from a single management system. The installation server is required for the proper installation of the EDG middleware.

The experience gained so far has shown that the site configuration and installation even using LCFG is a complex procedure that takes a long time to master. This is the main reason why achieving a correct site installation and configuration is a non-trivial task that may take several weeks when done by the first time or by a non-expert. However once the installation

server is successfully deployed and the client profiles correctly defined adding new systems can be easily and quickly performed.

For the external observer each grid site has two main entry points, the gatekeeper that provides the interface with the local computing resources and the storage element that provides the interface with the local storage resources. Table 2 contains the list of the systems available at the 14 production sites currently available to CrossGrid users.

Site	Location	Status	
CYFRONET	Cracow	CE	Name: <a href="http://zeus.cyf-kr.edu.pl">zeus.cyf-kr.edu.pl</a>
			Type: dual PIII 1GHz, 512MB, 40GB
			WNs: 19 WN (dual P4 Xeon 2.4GHz, 1GB)
		SE	Name: <a href="http://zeus03.cyf-kr.edu.pl">zeus03.cyf-kr.edu.pl</a>
Type: dual PIII 1GHz, 512MB, 40GB			
Storage: 460GB			
FZK	Karlsruhe	CE	Name: <a href="http://ce010.fzk.de">ce010.fzk.de</a>
			Type: dual PIII 1.2GHz, 1GB, 38GB
			WNs: 1 WN (dual PIII 1.2GHz, 1GB)
		SE	Name: <a href="http://se010.fzk.de">se010.fzk.de</a>
Type: dual PIII 1.2GHz, 1GB, 38GB			
Storage: 1.2TB			
IISAS	Bratislava	CE	Name: <a href="http://cluster.ui.sav.sk">cluster.ui.sav.sk</a>
			Type: P4 2.4GHz, 1GB, 40GB, GETH
			WNs: 16 WN (Intel P4 1.8GHz, 256MB)
		SE	Name: <a href="http://Flood-vo.ui.sav.sk">Flood-vo.ui.sav.sk</a>
Type: Pentium 233MHz, 128MB, 40GB			
Storage: 30GB			
PSNC	Poznan	CE	Name: <a href="http://sequoia.crossgrid.man.poznan.pl">sequoia.crossgrid.man.poznan.pl</a>
			Type: P4 1.7GHz, 512MB, 40GB
			WNs: 2 WN (dual Intel PIII 1.26GHz, 1GB, 80GB)
		SE	Name: <a href="http://se.crossgrid.man.poznan.pl">se.crossgrid.man.poznan.pl</a>
Type: PIII 1.26GHz, 1GB, 60GB			
Storage: 320GB			
IFIC	Valencia	CE	Name: <a href="http://bee001.ific.uv.es">bee001.ific.uv.es</a>
			Type: AMD Athlon 1GHz, 768MB, 38GB
			WNs: 1 WN (AMD Athlon 1GHz, 768MB)
		SE	Name: <a href="http://bee003.ific.uv.es">bee003.ific.uv.es</a>
Type: AMD Athlon 1GHz, 768MB, 38GB			
Storage: 50GB			
Demokritos	Athens	CE	Name: <a href="http://cgnode00.di.uoa.gr">cgnode00.di.uoa.gr</a>
			Type: P4 2GHz, 256MB, 70GB
			WNs: 2 WN (Intel P4 2GHz)
		SE	Name: <a href="http://cgnode04.di.uoa.gr">cgnode04.di.uoa.gr</a>
Type: P4 2GHz, 256MB, 70GB			
Storage: 50GB			

Site	Location	Status	
CESGA	Santiago	CE	Name: <a href="http://ce.grid.cesga.es">ce.grid.cesga.es</a>
			Type: dual PIII 1.12GHz, 512MB, 17.5GB
			WNs: 5 WN (dual PIII 1.12GHz, 512MB)
		SE	Name: <a href="http://se.grid.cesga.es">se.grid.cesga.es</a>
			Type: PIII 1.12GHz, 1GB, 17.5GB
			Storage: 36GB
INS	Warsaw	CE	Name: <a href="http://cms.fuw.edu.pl">cms.fuw.edu.pl</a>
			Type: PII 233MHz, 64MB, 8GB
			WNs: 2 WN (dual PIII 1GHz, 1GB)
		SE	Name: <a href="http://cmsse.fuw.edu.pl">cmsse.fuw.edu.pl</a>
			Type: Pentium MMX 233MHz, 64MB, 9GB
			Storage: 160GB
ICM	Warsaw	CE	Name: <a href="http://xgrid.icm.edu.pl">xgrid.icm.edu.pl</a>
			Type: PIII 450MHz, 64MB, 15GB
			WNs: 4 WN (dual PIII 1.26GHz, 1GB)
		SE	Name: <a href="http://xgridse.icm.edu.pl">xgridse.icm.edu.pl</a>
			Type: Pentium MMX 225MHz, 48MB, 10GB
			Storage: 300GB
LIP	Lisbon	CE	Name: <a href="http://Ingrid02.lip.pt">Ingrid02.lip.pt</a>
			Type: P4 1.7GHz, 512MB, 40GB
			WNs: 1 WN (Intel P4 1.7GHz)
		SE	Name: <a href="http://Ingrid03.lip.pt">Ingrid03.lip.pt</a>
			Type: P4 1.7GHz, 512MB, 40GB
			Storage: 38GB
UAB	Barcelona	CE	Name: <a href="http://aocegrid.uab.es">aocegrid.uab.es</a>
			Type: P4 1.8GHz, 512MB, 38GB
			WNs: 4 WN (Intel P4 1.8GHz, 512MB)
		SE	Name: <a href="http://aosegrid.uab.es">aosegrid.uab.es</a>
			Type: P4 1.8GHz, 512MB, 40GB
			Storage: 40GB
AUTH	Thessaloniki	CE	Name: <a href="http://grid01.physics.auth.gr">grid01.physics.auth.gr</a>
			Type: P3 1.26GHz, 512MB, 17GB
			WNs: 1 WN (Intel P3 1.26GHz, 512MB)
		SE	Name: <a href="http://grid04.physics.auth.gr">grid04.physics.auth.gr</a>
			Type: P4 1.26GHz, 512MB, 17GB
			Storage: 38GB
UCY	Nikosia	CE	Name: <a href="http://apelatis.grid.ucy.ac.cy">apelatis.grid.ucy.ac.cy</a>
			Type: dual PIII 1.2GHz, 1GB, 18GB
			WNs: 4 WN (Intel P4 2GHz, 1GB)
		SE	Name: <a href="http://psellos.grid.ucy.ac.cy">psellos.grid.ucy.ac.cy</a>
			Type: P4 2GHz, 1GB, 40GB
			Storage: 30GB

Site	Location	Status	
TCD	Dublin	CE	Name: <b>cagnode45.cs.tcd.ie</b>
			Type: <b>PIII 1GHz, 1GB, 60GB</b>
			WNs: <b>8 WN (dual PIII 1GHz, 1GB)</b>
		SE	Name: <b>cagnode47.cs.tcd.ie</b>
			Type: <b>PIII 1GHz, 1GB, 60GB</b>
			Storage: <b>50GB</b>

**Table 2. Configuration of the CrossGrid production sites.**

### 3.1.1 Sites statistics

This section contains job submission statistics collected with software developed at LIP. They are based on the Globus gatekeeper log files. Since the data origin is a log file and not an accounting file it might contain a level of uncertainty introduced by missing information, problems parsing the log files, or in performing the correlations required to create the accounting records.

Also in certain circumstances events can be recorded as errors without corresponding to real software failures. These are examples of such events:

- The “GSS failed jobs” most of the time aren’t the result of job failures but the result of monitoring tools and hacker tools scanning the TCP port 2119 used by the gatekeeper to receive job submission requests.
- The CRL failures aren’t due to middleware problems but to less careful CA management, they happen when a user tries to submit a job to a CE but the CA that has issued the user certificate has let its CRL expire.
- Some of the CRL failures in the statistics tables aren’t even related with CrossGrid CA’s but with DataGrid CA’s and DataGrid users submitting jobs into some CrossGrid systems.
- The job manager errors reported happen mostly when the job-manager software components are being upgraded, reconfigured or restarted.

The statistics cover all job submissions since the deployment of the initial testbed. Therefore they also reflect much of the early installation and configuration testing and experimentation that was initially performed at all sites.

The next table explains the meaning of the CE statistics fields.

<b>Total user DNs</b>	Total different user certificates submitting jobs.
<b>Total different Ips</b>	Total different IP addresses submitting jobs.
<b>Total Jobs submitted</b>	Total jobs submitted.
<b>Total Ping Jobs submitted</b>	Total GRAM Ping jobs submitted (service checks).
<b>Total Jobs successfully submitted</b>	Total jobs that passed to the execution phase.
<b>Total failed jobs</b>	Total jobs that didn't reach the execution phase.
- LCAS failed jobs	Total jobs failed due to authorization issues.
- GSS failed jobs	Total jobs failed GSS authentication protocols.
- CRL failed jobs	Total jobs lost due to outdated CA CRLs.
- Jobman failed jobs	Total jobs lost due Jobmanager problems.

**Table 3 Description of the CE statistics fields**

**3.1.1.1 LIP CE Usage statistics (Ingrid02.lip.pt)**

<b>Total user DNs</b>	<b>51</b>
<b>Total different Ips</b>	<b>57</b>
<b>Total Jobs submitted</b>	<b>7154</b>
<b>Total Ping Jobs submitted</b>	<b>465</b>
<b>Total Jobs successfully submitted</b>	<b>3412</b>
<b>Total failed jobs</b>	<b>3277</b>
- LCAS failed jobs	<b>56</b>
- GSS failed jobs	<b>3103</b>
- CRL failed jobs	<b>25</b>
- Jobman failed jobs	<b>93</b>

**Table 4. LIP CE usage statistics.**

**3.1.1.2 TCD CE Usage statistics (cagnode45.cs.tcd.ie)**

<b>Total user DNs</b>	<b>6</b>
<b>Total different Ips</b>	<b>10</b>
<b>Total Jobs submitted</b>	<b>743</b>
<b>Total Ping Jobs submitted</b>	<b>1</b>
<b>Total Jobs successfully submitted</b>	<b>205</b>
<b>Total failed jobs</b>	<b>537</b>
- LCAS failed jobs	<b>0</b>
- GSS failed jobs	<b>267</b>
- CRL failed jobs	<b>245</b>
- Jobman failed jobs	<b>25</b>

**Table 5. TCD CE usage statistics.**

**3.1.1.3 IFIC CE Usage statistics (bee001.ific.uv.es)**

Total user DNs	32
Total different Ips	38
Total Jobs submitted	5726
Total Ping Jobs submitted	660
Total Jobs successfully submitted	3023
Total failed jobs	2043
- LCAS failed jobs	108
- GSS failed jobs	1787
- CRL failed jobs	101
- Jobman failed jobs	47

Table 6. IFIC CE usage statistics.

**3.1.1.4 UAB CE Usage statistics (aocegrid.uab.es)**

Total user DNs	4
Total different Ips	7
Total Jobs submitted	259
Total Ping Jobs submitted	0
Total Jobs successfully submitted	249
Total failed jobs	10
- LCAS failed jobs	0
- GSS failed jobs	6
- CRL failed jobs	4
- Jobman failed jobs	0

Table 7. UAB CE usage statistics.

**3.1.1.5 FZK CE Usage statistics (ce010.fzk.de)**

Total user DNs	83
Total different Ips	44
Total Jobs submitted	1799
Total Ping Jobs submitted	11
Total Jobs successfully submitted	1117
Total failed jobs	671
- LCAS failed jobs	120
- GSS failed jobs	421
- CRL failed jobs	7
- Jobman failed jobs	123

Table 8. FZK CE usage statistics.

**3.1.1.6 PSNC CE Usage statistics (sequoia.crossgrid.man.poznan.pl)**

Total user DNs	16
Total different IPs	12
Total Jobs submitted	625
Total Ping Jobs submitted	0
Total Jobs successfully submitted	535
Total failed jobs	90
- LCAS failed jobs	15
- GSS failed jobs	50
- CRL failed jobs	14
- Jobman failed jobs	11

Table 9. PSNC CE usage statistics.

**3.1.1.7 CESGA CE Usage statistics (ce.grid.cesga.es)**

Total user DNs	13
Total different Ips	20
Total Jobs submitted	7941
Total Ping Jobs submitted	1
Total Jobs successfully submitted	1021
Total failed jobs	6919
- LCAS failed jobs	82
- GSS failed jobs	6792
- CRL failed jobs	28
- Jobman failed jobs	17

Table 10. CESGA CE usage statistics.

**3.1.1.8 ICM CE Usage statistics (xgrid.icm.edu.pl)**

Total user DNs	17
Total different Ips	17
Total Jobs submitted	940
Total Ping Jobs submitted	35
Total Jobs successfully submitted	830
Total failed jobs	75
- LCAS failed jobs	8
- GSS failed jobs	41
- CRL failed jobs	24
- Jobman failed jobs	2

Table 11. ICM CE usage statistics.

**3.1.1.9 INS CE Usage statistics (cms.fuw.edu.pl)**

Total user DNs	6
Total different IPs	15
Total Jobs submitted	875
Total Ping Jobs submitted	18
Total Jobs successfully submitted	774
Total failed jobs	83
- LCAS failed jobs	6
- GSS failed jobs	37
- CRL failed jobs	20
- Jobman failed jobs	20

Table 12. INS CE usage statistics.

**3.1.1.10 Demokritos CE Usage statistics (cgnode00.di.uoa.gr)**

Total user DNs	13
Total different Ips	21
Total Jobs submitted	9481
Total Ping Jobs submitted	5
Total Jobs successfully submitted	1111
Total failed jobs	8365
- LCAS failed jobs	36
- GSS failed jobs	8278
- CRL failed jobs	0
- Jobman failed jobs	51

Table 13. Demokritos CE usage statistics.

**3.1.1.11 UCY CE Usage statistics (apelatis.grid.ucy.ac.cy)**

Total user DNs	2
Total different Ips	3
Total Jobs submitted	26
Total Ping Jobs submitted	0
Total Jobs successfully submitted	14
Total failed jobs	12
- LCAS failed jobs	10
- GSS failed jobs	2
- CRL failed jobs	0
- Jobman failed jobs	0

Table 14. UCY CE usage statistics.

**3.1.1.12 IISAS CE Usage statistics (cluster.ui.sav.sk)**

Total user DNs	18
Total different Ips	14
Total Jobs submitted	1774
Total Ping Jobs submitted	6
Total Jobs successfully submitted	1555
Total failed jobs	213
- LCAS failed jobs	0
- GSS failed jobs	99
- CRL failed jobs	15
- Jobman failed jobs	99

Table 15. IISAS CE usage statistics.

**3.1.1.13 IISAS CE Usage statistics (cluster.ui.sav.sk)**

Total user DNs	18
Total different Ips	14
Total Jobs submitted	1774
Total Ping Jobs submitted	6
Total Jobs successfully submitted	1555
Total failed jobs	213
- LCAS failed jobs	0
- GSS failed jobs	99
- CRL failed jobs	15
- Jobman failed jobs	99

Table 16. IISAS CE usage statistics.

**3.1.1.14 TCD CE Usage statistics (cagnode45.cs.tcd.ie)**

Total user DNs	6
Total different Ips	10
Total Jobs submitted	743
Total Ping Jobs submitted	1
Total Jobs successfully submitted	205
Total failed jobs	537
- LCAS failed jobs	0
- GSS failed jobs	267
- CRL failed jobs	245
- Jobman failed jobs	25

Table 17. TCD CE usage statistics.

**3.1.1.15 AUTH CE Usage statistics (grid01.physics.auth.gr)**

<b>Total user DNs</b>	<b>6</b>
<b>Total different Ips</b>	<b>13</b>
<b>Total Jobs submitted</b>	<b>451</b>
<b>Total Ping Jobs submitted</b>	<b>0</b>
<b>Total Jobs successfully submitted</b>	<b>236</b>
<b>Total failed jobs</b>	<b>215</b>
<b>- LCAS failed jobs</b>	<b>8</b>
<b>- GSS failed jobs</b>	<b>175</b>
<b>- CRL failed jobs</b>	<b>32</b>
<b>- Jobman failed jobs</b>	<b>0</b>

**Table 18. AUTH CE usage statistics.**

### 3.2 PRODUCTION CENTRAL SERVICES

A set of central services is required for proper testbed operation. These services are essential for user authentication, authorization, workload management, monitoring and location of file replicas. The following CrossGrid production central services are currently deployed:

- **Resource Broker (RB):** Is the heart of the testbed workload management system. The RB receives job requests sent by users through a UI and finds computing resources suitable to run the jobs. The matchmaking between jobs and resources is accomplished using the job description information and the site characteristics announced through the Globus MDS information system.
- **Replica Catalogue (RC):** Is a LDAP server that contains logical file names and the corresponding physical file replica location. The RC is used to obtain the location of physical files in the grid. The RB uses the RC to make scheduling decisions based on the location of the files required by the jobs.
- **MyProxy:** Is an authentication credential proxy used by the RB to renew the proxy certificates of long-lived jobs. MyProxy stores certificates on the behalf of other subjects and can issue or renew short-term proxy certificates based on the stored certificates.
- **VO server:** Is a repository of authorization information used by the testbed systems to rebuild the authorization databases. The CrossGrid VO server contains the list of all CrossGrid users authorized to use the testbed. Each grid node performs an authorization database update every 24 hours using information from the VO server.
- **Mapcenter:** Is a network-monitoring platform designed to monitor grid infrastructures.

In June of 2002 upon the deployment of the first sites it was clear that CrossGrid needed some central services at least for authorization and load management. A first possibility was to use the DataGrid central services. This would have the advantage of allowing CrossGrid to concentrate first on site deployment and would promote the integration of both testbeds from the beginning. However many CrossGrid partners were not involved in any of the communities of DataGrid and therefore couldn't easily achieve membership in a DataGrid VO, clearly CrossGrid due to its own characteristics and nature would need its own VO or VOs. Problems also appeared on the load management side, since the RB software available at the time was quite unstable suffering from scalability problems and could be brought down by sites with wrong configurations. DataGrid was not keen on adding new sites to the RB since it would increase the workload management instability. Another problem was the recognition by DataGrid of the new certification authorities being deployed within CrossGrid countries. These CA's were issuing certificates to CrossGrid users and systems but since they were still not recognized by DataGrid these systems and users could not use or be used by the DataGrid central services. At that point it was felt that CrossGrid needed its own central services even for a limited initial testbed based only on Globus and EDG middleware. These services were essential to integrate all CrossGrid sites into a testbed and allow CrossGrid users to use the project distributed computing resources.

At the time LIP was involved in DataGrid with a site already deployed and was testing the EDG central services middleware within the context of CrossGrid task 4.4. Therefore it was proposed that the existing central services deployed at LIP for test purposes could be reconfigured to provide central services for the whole CrossGrid project. Thanks to the experience gained in the test activities the conversion of the existing test services (RB and

VO server) and the deployment of some new required services (RC) was quickly accomplished. These services are now the basis of the CrossGrid production testbed.

The CrossGrid central services for the production testbed are located at the LIP Computer Center facilities in Lisbon. The systems are maintained in a cooled room and protected against power failures and surges by two UPS systems. The next diagram shows the configuration of the LIP production site including the central services.

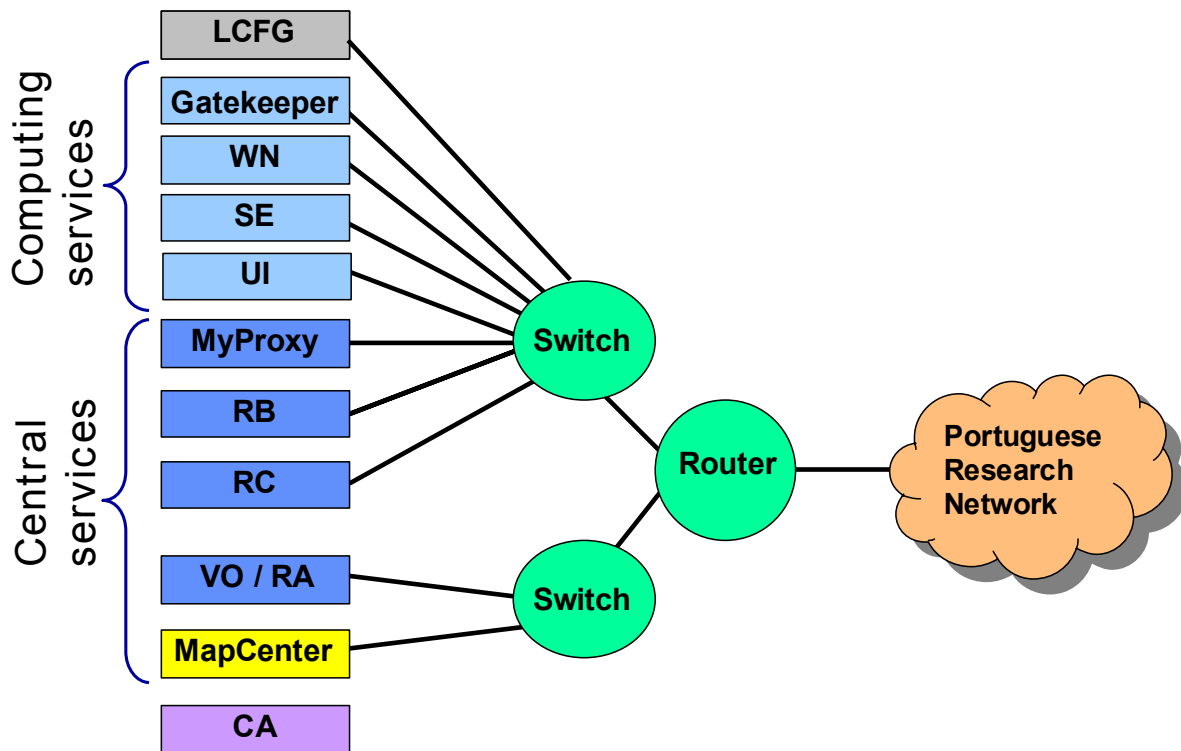


Figure 1. The CrossGrid production central services and the LIP production site.

### 3.2.1 Production Resource Broker

The CrossGrid production resource broker performs load sharing and resource matching between submitted jobs and the CrossGrid computing resources. The RB accepts job requests from CrossGrid users submitted through the User Interface systems. Access control is performed using a local authorization file rebuilt every 24 hours from the CrossGrid VO database hosted at the central VO server. Once a user is authenticated and authorized to access the RB the job description is transferred from the UI to the RB and a matchmaking operation is started to find computing resources matching the job requirements. In the matchmaking the RB uses information about the existing computing resources in the testbed obtained from the Globus MDS information system, for this purpose the production RB contains an MDS Information Index (II) server with pointers to the MDS/GRIS information servers in the testbed computing elements and storage elements. The information index server must be updated “manually” each time a new site is added or removed from the testbed. When the job description contains references to logical files the RB must also query the central Replica Catalogue to obtain the location of the corresponding physical replicas so

that the selected CE has a near SE containing replicas of the required files. Finally the RB uses the central MyProxy server to renew the job credentials when needed.

The system hosting the RB also hosts the Logging and Bookkeeping (LB) database. The LB is used to record information about job status. The LB receives information from several components both internal and external to the RB host system. Internal sources are the components of RB such as the RB itself and the Job Submission System (JSS). External sources are the User Interfaces, and the Gatekeepers. The LB can be queried to obtain the status of a job or historic data about previously submitted jobs. The logging information is stored in a MySQL relational database.

The following diagram shows the CrossGrid RB and the sites registered into the RB Information Index and therefore candidates to execute CrossGrid jobs. The two sites marked as not available have been moved temporarily to the test and validation testbed and cannot be used through the production RB.

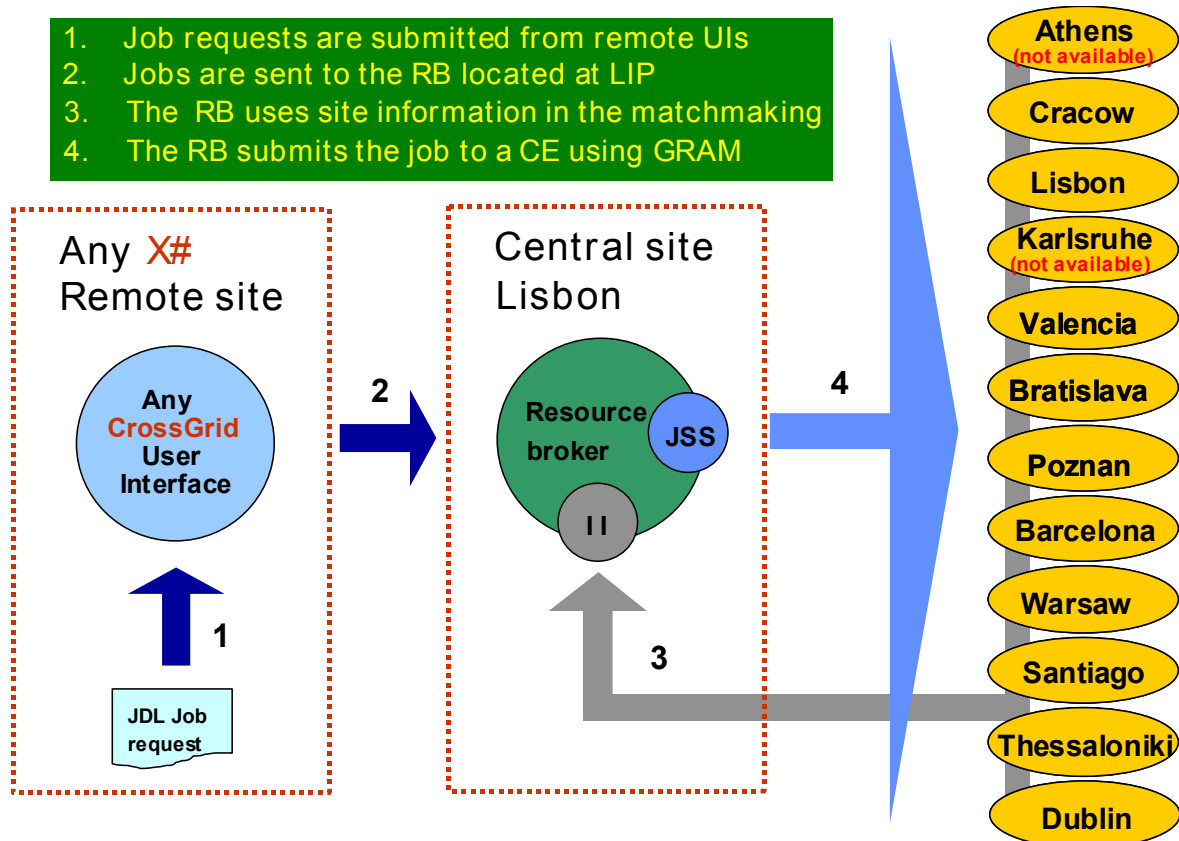


Figure 2. The Job submission through the CrossGrid production RB.

The gray arrow shows the flow of the site information provided by MDS and used by the RB in the matchmaking process through the Information Index.

**3.2.1.1 Production RB system characteristics:**

<b>NAME</b>	Ingrid06.lip.pt
<b>CPU</b>	Pentium 4 1.8GHz
<b>RAM</b>	512 MB
<b>STORAGE</b>	IDE DISK 40GB
<b>NIC</b>	Fast Ethernet
<b>OS</b>	Linux Red Hat 6.2

**Table 19. RB system characteristics.**

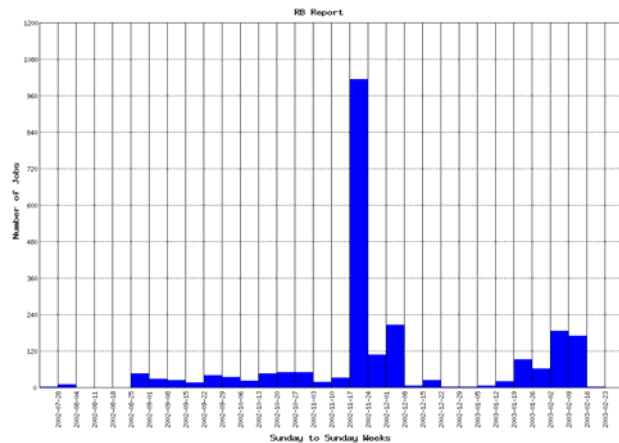
**3.2.1.2 Production RB TCP ports:**

<b>8881</b>	RB port to listen for JSS
<b>7771</b>	RB port to listen for UI
<b>9991</b>	JSS
<b>7846</b> <b>7890</b> <b>15830</b>	Logging and bookkeeping
<b>2170</b>	Information Index

**Table 20. RB ports.**

**3.2.1.3 Production RB Usage statistics**

<b>Total users</b>	<b>35</b>
<b>Total jobs sent to the RB</b>	<b>2313</b>
<b>Total jobs accepted by RB</b>	<b>2168</b>
<b>Total jobs with good match</b>	<b>2010</b>
<b>Total jobs submitted by JSS</b>	<b>1988</b>
<b>Total jobs run</b>	<b>1775</b>
<b>Total jobs done</b>	<b>1207</b>



**Table 21. RB statistics.**

The statistics show a peak of usage corresponding to a period of intense RB testing. However since the RB doesn't have currently support parallel job submissions most of the application tests are being performed directly with Globus thus bypassing the RB. CrossGrid is developing software that will allow the usage of the RB matchmaking capabilities for parallel job submissions. A second peak of activity can be observed in February of 2003. The peak corresponds to the software integration activities that culminated with the Santiago workshop where CrossGrid applications, tools and middleware were demonstrated.

### 3.2.2 Production Replica Catalogue

The Replica Catalogue is basically an LDAP server containing logical file names. For each logical file name the server contains the location of the physical file replicas. The physical file replicas can be stored in multiple testbed SEs. The replica catalogue is a key piece in the data access optimization since combined with the RB allows jobs to be executed at sites where the required data files are already present in the local SE, hence minimizing data transfers and reducing data access time and network bandwidth usage. The file replication software such as the Replica Manager and GDMP uses the Replica Catalogue to register the location of the replicated files.

The CrossGrid production RC server contains logical collections for CrossGrid, Atlas, CMS and DataGrid wpsix VOs.

#### 3.2.2.1 Production RC system characteristics:

<b>NAME</b>	Ingrid07.lip.pt
<b>CPU</b>	AMD K6 200MHz
<b>RAM</b>	128 MB
<b>STORAGE</b>	IDE DISK 4GB
<b>NIC</b>	Fast Ethernet
<b>OS</b>	Linux Red Hat 6.2

Table 22. RC system characteristics.

#### 3.2.2.2 Production RC TCP ports:

<b>9980</b>	CrossGrid RC
-------------	--------------

Table 23. RC ports.

#### 3.2.2.3 Production RC Logical collections:

<b>Crossgrid</b>	ldap://Ingrid08.lip.pt:9980/lc=cgtst0,rc=CrossGridReplicaCatalogue,dc=Ingrid08,dc=lip,dc=pt
<b>Wpsix</b>	ldap://Ingrid08.lip.pt:9980/lc=wpsixtst0,rc=CrossGridReplicaCatalogue,dc=Ingrid08,dc=lip,dc=pt
<b>Atlas</b>	ldap://Ingrid08.lip.pt:9980/lc=atlastst0,rc=CrossGridReplicaCatalogue,dc=Ingrid08,dc=lip,dc=pt
<b>Cms</b>	ldap://Ingrid08.lip.pt:9980/lc=cmstst0,rc=CrossGridReplicaCatalogue,dc=Ingrid08,dc=lip,dc=pt

Table 24. RC logical collections.

### 3.2.3 Production VO server

The CrossGrid VO server contains an authorization database used to rebuild the authorization files of all CrossGrid systems. All CE, SE and RB systems perform a rebuild of their local authorization files at least once a day.

The CrossGrid VO server is a LDAP server that contains two production virtual organizations; they are “CrossGrid” and “gdmpservers”. The “CrossGrid” VO contains the distinguish names of all CrossGrid users. The “gdmpservers” VO contains the distinguish names of the GDMP servers (SEs) in the testbed.

Each VO can have groups; the groups contain subsets of the VO members for authentication purposes. Access control is performed most times by authorizing or denying a VO group. For CrossGrid the group that contains the users authorized to use the production testbed is called “testbed1”. The “gdmpservers” VO has two groups called “apptb” and “devtb”. The “apptb” group contains the servers of the production testbed while the “devtb” group contains the development testbed servers.

The names of VOs and groups like many other configuration options have been chosen to be compatible with the ones used in the DataGrid testbed in order to make both testbeds as compatible as possible and to make easier the participation of CrossGrid testbed sites in DataGrid.

#### 3.2.3.1 Production VOs and groups:

VO	GROUP	DESCRIPTION
crossgrid	testbed1	CrossGrid users
gdmpservers	apptb	Application GDMP servers
gdmpservers	devtb	Development GDMP servers

Table 25. VOs and groups.

#### 3.2.3.2 VO server system characteristics:

NAME	grid-vo.lip.pt
CPU	Pentium II 300MHz
RAM	128 MB
STORAGE	IDE DISK 30GB
NIC	Fast Ethernet
OS	Linux Red Hat 7.2

Table 26. VO system characteristics.

#### 3.2.3.3 VO server TCP ports:

9990	CrossGrid VO server
------	---------------------

Table 27. VO ports.

### 3.2.4 Production MyProxy

For long-lived jobs there is the possibility that some jobs could outlive the validity of its proxy certificate causing the job to fail. To avoid this, the workload management tools allow a proxy to be automatically renewed via a MyProxy server. The server manages long-lived proxies generated by the users and gives updated short-lived proxies to properly authenticated processes on behalf of the certificate holder.

Since valid user proxy certificates are stored in the MyProxy server, the security of the system hosting this service is central to the security of the whole testbed. The CrossGrid MyProxy server is hosted in a fully dedicated system.

#### 3.2.4.1 MyProxy System characteristics:

<b>NAME</b>	Ingrid08.lip.pt
<b>CPU</b>	AMD K6 200MHz
<b>RAM</b>	64 MB
<b>STORAGE</b>	IDE DISK 3GB
<b>NIC</b>	Fast Ethernet
<b>OS</b>	Linux Red Hat 6.2

**Table 28. MyProxy system characteristics.**

#### 3.2.4.2 MyProxy TCP ports:

<b>7512</b>	CrossGrid MyProxy
-------------	-------------------

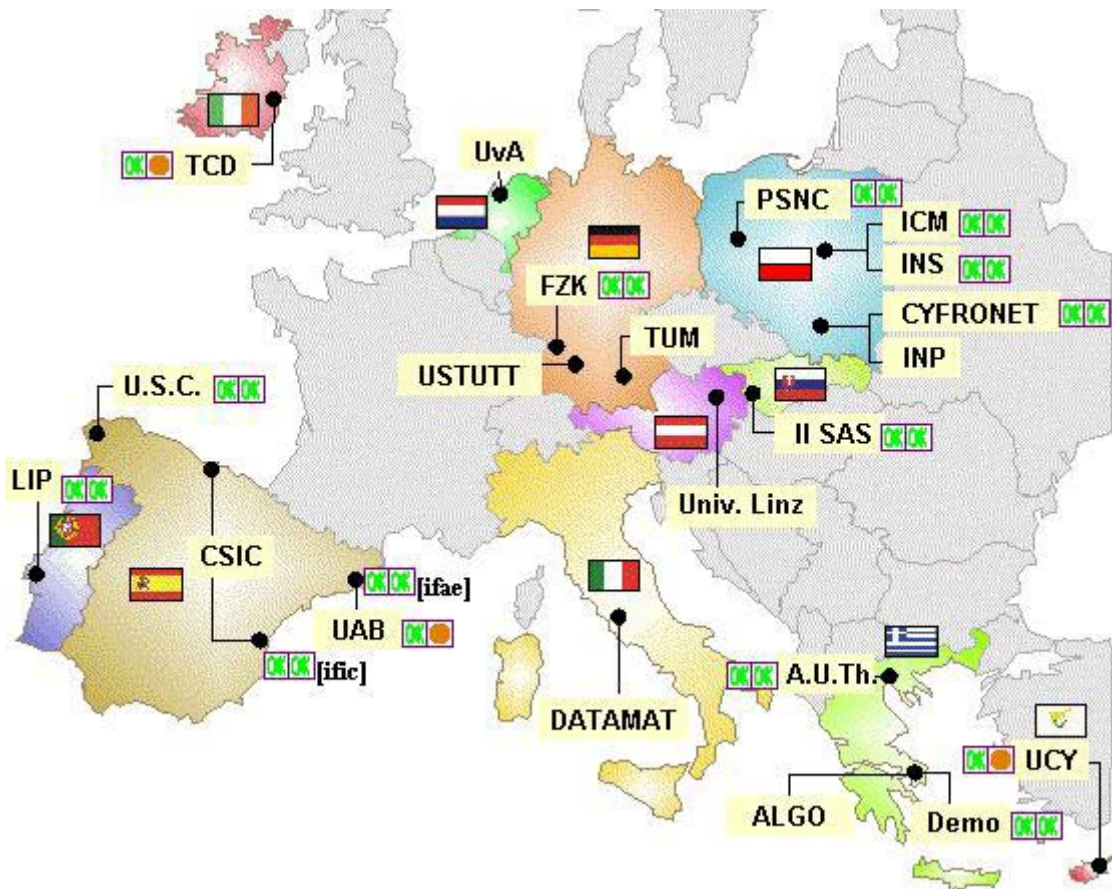
**Table 29. MyProxy ports.**

### 3.2.5 Production Mapcenter

Mapcenter is basically a network-monitoring tool for grid infrastructures. It is used to monitor the network reach ability of the grid services running in all testbed sites. Mapcenter also allows navigating and querying the MDS information servers. The monitoring results are made available through an Apache web server and the information can be displayed in several views namely geographically, by country or by site. The CrossGrid Mapcenter server is located at the LIP computer centre the URL for the main web page is <http://mapcenter.lip.pt>.

In the next figure the CrossGrid Mapcenter geographical view is shown, it should be noticed that some sites are being monitored but are not yet registered into the RB since they are still being validated.

The map shows for each site the minimum and maximum status. The minimum status appears first and corresponds to the less severe error detected in the resources registered for the site. The maximum status appears in second place and corresponds to the most severe error detected in the resources registered for the site. Each pair of status appearing in the map can represent many resources or even sites at the same location or belonging to the same organization. This is the case of LIP that is hosting in Lisbon two clusters and two sets of central services for the CrossGrid production and validation testbeds. The Mapcenter server at LIP is thus monitoring both the production and validation testbeds.



**Figure 3. Mapcenter geographical view of the CrossGrid sites.**

The above view captured in February shows three sites partially down. The circles in the UAB, TCD and UCY were caused by SE services (MDS or GRAM) that were down thus causing port connection failures. These are transient problems related with ongoing site configuration tuning. It can also be seen that the maximum status for the same sites is OK, meaning that all other services are available.

The CrossGrid Mapcenter has the following views:

- **CrossGrid Europe:** Geographical view showing the status and location of each site in the map of Europe.
- **Countries:** A view of the sites per country.
- **Organization:** A view of the sites per organization.
- **Main servers:** Central services.
- **Full tree view:** A view of all Countries sites and resources.
- **Objects:** A table of all resources with detailed information.
- **History:** Historical information about past occurring events.
- **Alerts:** Last alarms.

For each system registered in Mapcenter the connectivity status of each service is shown, some systems such as the RB's and CE's also include links to web pages containing usage statistics developed at LIP.

### 3.2.5.1 Mapcenter system characteristics:

<b>NAME</b>	mapcenter.lip.pt
<b>CPU</b>	Pentium III 866MHz
<b>RAM</b>	256 MB
<b>STORAGE</b>	IDE DISK 30GB
<b>NIC</b>	Fast Ethernet
<b>OS</b>	Linux Red Hat 6.2

Table 30. Mapcenter system characteristics.

### 3.2.5.2 Mapcenter TCP ports:

<b>80</b>	Mapcenter web server
-----------	----------------------

Table 31. Mapcenter ports.

### 3.3 TEST AND VALIDATION SITES

The decision to start deploying the test and validation infrastructure was taken due to the recent necessity of testing the EDG 1.4.3 middleware release. These tests must be performed without disturbing the production testbed in a tightly controlled infrastructure with a reduced set of sites. Incompatibilities between the 1.4.x and 1.2.x EDG releases due to changes in both EDG and Globus also contributed to the decision of not mixing testbed sites running different middleware versions.

The test and validation testbed has currently a main site at LIP in Lisbon and two other sites at AUTH in Demokritos and FZK in Karlsruhe. LIP has deployed a second infrastructure fully dedicated to test and validation. The FZK and Demokritos sites are temporarily being used for test and validation and they will be probably moved back to the production testbed once the tests are finished and the middleware certified. However Demokritos plans to dedicate its infrastructure to test and validation activities once a second local cluster dedicated to production is deployed.

The next table shows the deployment status of the 3 CrossGrid test and validation sites.

Site	Location	Status
FZK	Karlsruhe	Running 1.4.3, temporarily in the test and validation testbed.
Demokritos	Athens	Running 1.4.3, temporarily in the test and validation testbed.
LIP	Lisbon	Running 1.4.3, registered into the test and validation RB.

**Table 32. CrossGrid test and validation sites status.**

The test and validation testbed architecture obeys to the CrossGrid minimum hardware requirements therefore each test site must have the same five system components described for the production testbed. They are:

- **One Gatekeeper**
- **One or more Worker Nodes**
- **One Storage Element**
- **One User Interface**
- **One LCFG Installation Server**

Table 33 contains the list of the systems used as entry points for data access and job submission at the 3 sites currently available in the test and validation testbed.

Site	Location	Status	
FZK	Karlsruhe	CE	Name: ce001.crossgrid.fzk.de
			Type: dual PIII 1.2GHz, 1GB, 38GB
			WNs: 1 WN (dual PIII 1.2GHz)
		SE	Name: se001.crossgrid.fzk.de
			Type: dual PIII 1.2GHz, 1GB, 38GB
Storage: 1.2TB			
Demokritos	Athens	CE	Name: cgnode00.di.uoa.gr
			Type: P4 2GHz, 256MB, 70GB
			WNs: 2WN (Intel P4 2GHz)
		SE	Name: cgnode04.di.uoa.gr
			Type: P4 2GHz, 256MB, 70GB
Storage: 50GB			
LIP	Lisbon	CE	Name: ce01.lip.pt
			Type: P4 1.8GHz, 512MB, 40GB
			WNs: 1 WN (Intel P4 1.8GHz)
		SE	Name: se01.lip.pt
			Type: P4 1.8GHz, 512MB, 40GB
Storage: 38GB			

Table 33. Configuration of the CrossGrid test and validation sites.

### 3.3.1 Sites statistics

The statistics software developed to collect and publish information about the usage of the production CE's is also being applied the test and validation testbed. No changes to the statistics software were required. The statistics here presented for FZK and Demokritos reflect the sites usage since they joined the test and validation testbed.

#### 3.3.1.1 LIP CE Usage statistics (ce01.lip.pt)

Total user DNs	11
Total different Ips	15
Total Jobs submitted	75860
Total Ping Jobs submitted	5333
Total Jobs successfully submitted	70419
Total failed jobs	108
- LCAS failed jobs	36
- GSS failed jobs	68
- CRL failed jobs	0
- Jobman failed jobs	4

Table 34. LIP test and validation CE usage statistics.

**3.3.1.2 FZK CE Usage statistics (ce010.fzk.de)**

Total user DNs	16
Total different Ips	32
Total Jobs submitted	16406
Total Ping Jobs submitted	64
Total Jobs successfully submitted	16202
Total failed jobs	140
- LCAS failed jobs	38
- GSS failed jobs	39
- CRL failed jobs	13
- Jobman failed jobs	50

Table 35. FZK test and validation CE usage statistics.

**3.3.1.3 Demokritos CE Usage statistics (cgnode00.di.uoa.gr)**

Total user DNs	10
Total different Ips	14
Total Jobs submitted	12602
Total Ping Jobs submitted	407
Total Jobs successfully submitted	7694
Total failed jobs	4501
- LCAS failed jobs	4
- GSS failed jobs	4483
- CRL failed jobs	9
- Jobman failed jobs	5

Table 36. Demokritos test and validation CE usage statistics.

### 3.4 TEST AND DEVELOPMENT CENTRAL SERVICES

As mentioned in the production central services section several central services are required for proper testbed operation, including user authentication, authorization, workload management, monitoring and location of file replicas. Most of the central services provided for the test and validation testbed have the same basic functionalities already explained for the central production services. The following CrossGrid test and validation central services are currently deployed:

- **Resource Broker (RB):** Is the heart of the testbed workload management system.
- **Replica Catalogue (RC):** Contains logical file names and the corresponding physical file replica location.
- **MyProxy:** Is an authentication credential proxy.
- **VO server:** Is an authorization database system.
- **Mapcenter:** Is a network-monitoring system.
- **Information Index (II):** This is a new service that provides the root entry point for the MDS information tree. All sites publish information about their resources in MDS servers contained in the CE and SE systems. For this information to be found by the RB these MDS servers must now register themselves into the central Information Index.

The EDG 1.4.x release contains many improvements and changes not only at the site services level but also at the central services level. These changes do not allow mixing most of the 1.2.x central services with the 1.4.x systems. Also the requirement of having independence between the production and validation testbeds has imposed the installation of new servers to support the test and validation central services.

The CrossGrid central services for the test and validation testbed are currently located at the LIP Computer Center facilities in Lisbon. LIP was already providing the central services for the CrossGrid initial testbed, and using it to conduct test and validation activities. It was also foreseen that in the context of the task 4.4 LIP would host and manage the central services for the test and validation testbed. Currently the central services include dedicated servers to host the RB, II and RC. The Mapcenter, MyProxy and VO services are being shared with the production testbed. Sharing these three services doesn't impose any deterioration of performance or functionality in both testbeds, and contributes to the optimization of resources and human effort required for their maintenance.

The next diagram shows the current configuration of the LIP site including the local and central services available for both production and test and validation.

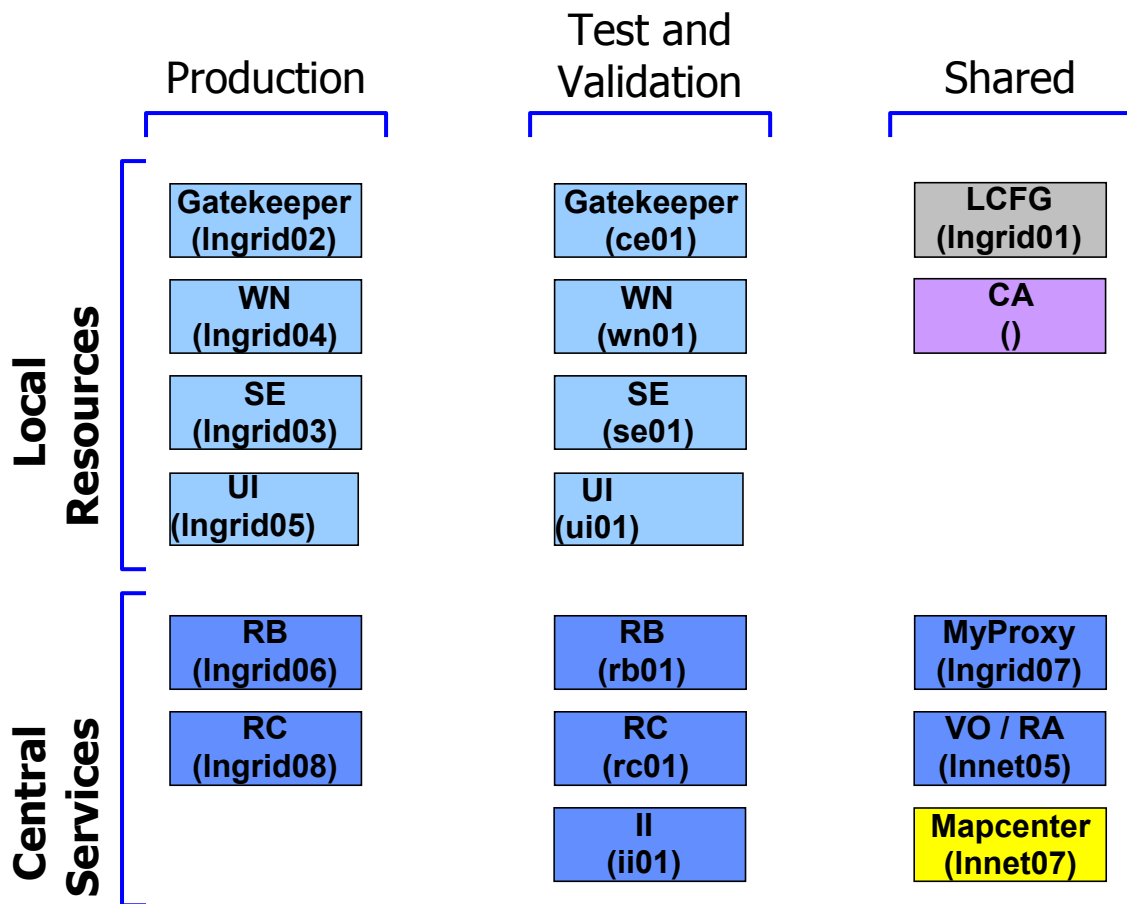


Figure 4. The CrossGrid services and resources at the LIP site.

### 3.4.1 Test and validation Resource Broker

The CrossGrid test and validation RB has the same basic functionalities as the production RB. Therefore the description already provided for the production RB is also valid for the test and validation. However there are some major differences not in the RB itself but in the way the RB accesses the MDS information. Previously computing resources were registered manually in a LDAP Information Index service running in the RB system. Now the Information Index service runs in a dedicated system, and MDS software running in the testbed Gatekeepers performs the site resources registration automatically. However the registration process is still controlled by the II administrator. Permission for registering resources is controlled and granted through policy configuration files that limit which DNS hosts or domains are allowed to register themselves. The registration request must be renewed every ten minutes to keep the corresponding resources in the MDS tree. In this way sites are automatically removed when unavailable and added when available contributing for the information accuracy and higher reliability of the MDS tree.

The test and validation RB still requires the presence of a RC, which is used when a job description contains references to logical files. In these cases the RB queries the RC to

obtain the location of the corresponding physical replicas. A dedicated RC has been deployed for the test and validation testbed.

The RB also uses a MyProxy server to renew job credentials and a VO server to obtain the list of the DN's authorized to submit jobs through the RB. These services are shared between the production and the test and validation testbeds.

The following diagram shows the CrossGrid test and validation RB and the sites that can be used to execute test jobs. The gray arrow shows the flow of the site information provided by MDS and used by the RB in the matchmaking process through the Information Index.

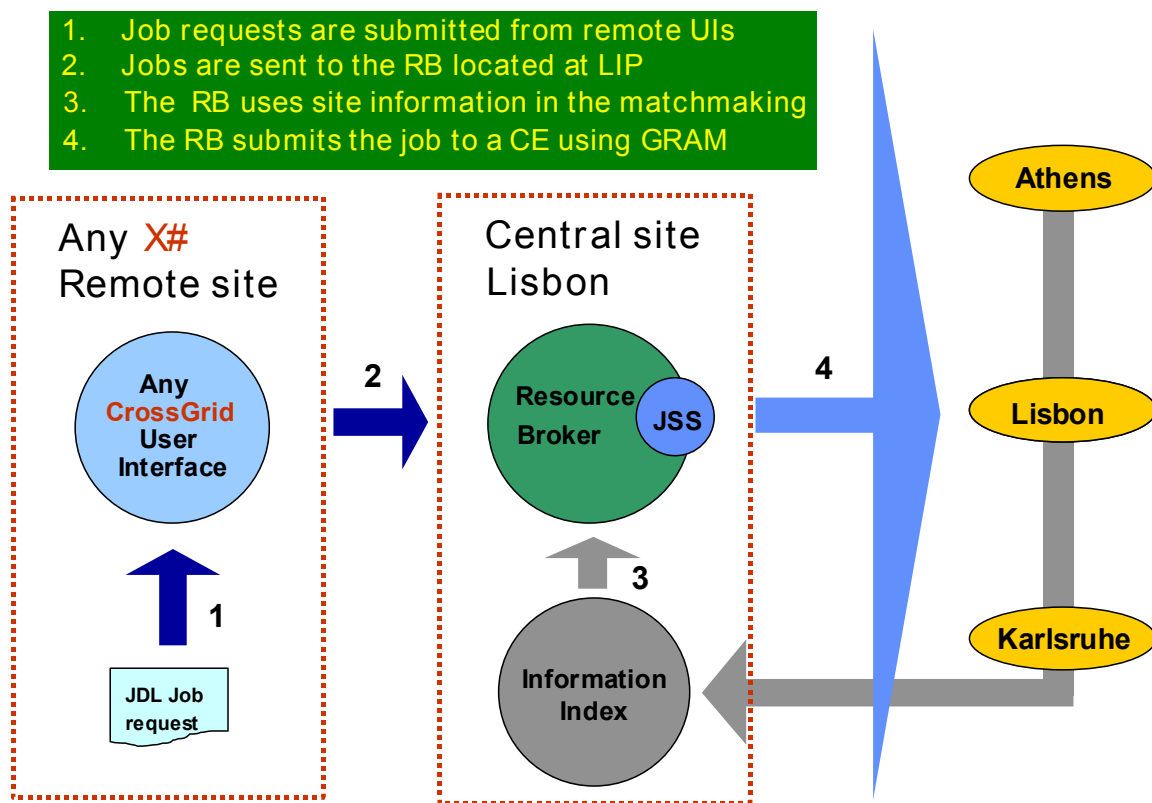


Figure 5. The Job submission through the CrossGrid test and validation RB.

**3.4.1.1 Test and Validation RB system characteristics:**

<b>NAME</b>	rb01.lip.pt
<b>CPU</b>	Pentium 4 1.8GHz
<b>RAM</b>	512 MB
<b>STORAGE</b>	IDE DISK 40GB
<b>NIC</b>	Fast Ethernet
<b>OS</b>	Linux Red Hat 6.2

Table 37. RB system characteristics.

**3.4.1.2 Test and Validation RB TCP ports:**

8881	RB port to listen for JSS
7771	RB port to listen for UI
9991	JSS
7846 7890 15830	Logging and bookkeeping

Table 38. RB ports.

**3.4.1.3 Test and Validation RB Usage statistics**

Total users	9
Total jobs sent to the RB	5407
Total jobs accepted by RB	5396
Total jobs with good match	5178
Total jobs submitted by JSS	5167
Total jobs run	5022
Total jobs done	4976

Table 39. RB statistics.

**3.4.2 Test and Validation Replica Catalogue**

As previously explained in the production section the Replica Catalogue is basically an LDAP server containing sets of logical file names. For each logical file name the server contains the location of the physical file replicas. The RB uses the RC in the matchmaking process to find the location of the file replicas, and the file replication software uses the Replica Catalogue to register the location of the replicated files.

The CrossGrid RC server contains logical collections for CrossGrid, Atlas, CMS and DataGrid wpsix VO's.

**3.4.2.1 Test and Validation RC system characteristics:**

NAME	rc01.lip.pt
CPU	Celeron 1.4GHz
RAM	512 MB
STORAGE	IDE DISKs 4GB + 2GB
NIC	Fast Ethernet
OS	Linux Red Hat 6.2

Table 40. RC system characteristics.

### 3.4.2.2 Test and Validation RC TCP ports:

9980	CrossGrid RC
------	--------------

**Table 41. RC ports.**

### 3.4.2.3 Test and Validation RC Logical collections:

<b>Crossgrid</b>	ldap://rc01.lip.pt:9980/lc= <b>cgtst0</b> ,rc=CrossGridReplicaCatalogue,dc=rc01,dc=lip,dc=pt
<b>Wpsix</b>	ldap://rc01.lip.pt:9980/lc= <b>wpsixtst0</b> ,rc=CrossGridReplicaCatalogue,dc=rc01,dc=lip,dc=pt
<b>Atlas</b>	ldap://rc01.lip.pt:9980/lc= <b>atlastst0</b> ,rc=CrossGridReplicaCatalogue,dc=rc01,dc=lip,dc=pt
<b>Cms</b>	ldap://rc01.lip.pt:9980/lc= <b>cmstst0</b> ,rc=CrossGridReplicaCatalogue,dc=rc01,dc=lip,dc=pt

**Table 42. RC logical collections.**

### 3.4.3 Test and validation VO server

In the current setup the VO server is shared between the production and the test and validation testbeds. However the VO server contains VO's and groups dedicated to either the production or the test and validation testbeds.

Besides the two VO's already described in the production section there is a third VO called "cgTV" that contains the distinguish names of the CrossGrid test and validation user certificates. The "cgTV" VO also has two groups they are called "alpha" and "beta" following the names given to the two test phases. In test phase "alpha" the middleware is tested by the test and validation group while in phase "beta" the test and validation testbed is opened to users that will validate the testbed by running their applications over the deployed middleware.

The "gdmpservers" VO has also a group called "tvb" that contains the distinguish names of the GDMP servers in the test and validation testbed.

**3.4.3.1 Test and validation VOs and groups:**

VO	GROUP	DESCRIPTION
gdmpservers	tvtb	Validation GDMP servers
cgTV	alpha	Test and validation group
cgTV	beta	Test and validation end users

**Table 43. VOs and groups.****3.4.3.2 VO server system characteristics:**

There is only one VO server that is shared by the production and test and validation testbeds. The characteristics of the VO server are described in the production section.

**3.4.3.3 VO server TCP ports:**

There is only one TCP port used by the VO server to serve both the production and the test and validation testbeds. This port is described in the production section.

**3.4.4 Test and validation MyProxy**

As described in the production section the MyProxy server is a certificate repository that can issue short-lived proxy certificates to properly authenticated processes on behalf of the certificate holder.

Currently there is only one MyProxy server to support both the production and the test and validation testbeds.

**3.4.4.1 MyProxy System characteristics:**

There is only one MyProxy server that is shared by the production and test and validation testbeds. The characteristics of the MyProxy server are described in the production section.

**3.4.4.2 MyProxy TCP ports:**

There is only one TCP port used by the MyProxy server to serve both the production and the test and validation testbeds. This port is described in the production section.

**3.4.5 Test and validation Mapcenter**

Mapcenter is basically a network-monitoring tool for grids with a web interface. There is only one Mapcenter server for both the production and the test and validation testbeds. The Mapcenter interface monitors sites in both testbeds. The full tree view clearly identifies the resources dedicated to test and validation activities.

#### **3.4.5.1 *Mapcenter system characteristics:***

There is only one Mapcenter server that is shared by the production and test and validation testbeds. The characteristics of the Mapcenter server are described in the production section.

#### **3.4.5.2 *Mapcenter TCP ports:***

There is only one TCP port used by the Mapcenter server to serve both the production and the test and validation testbeds. This port is described in the production section.

## 4 TESTBED EXTENSION

It is expected that after the full integration of the first 16 official sites more sites will be added to the testbed. In fact several request have already been received from project partners and external organizations. Adding new sites to the CrossGrid testbed will contribute to the expansion of the grid technologies in Europe and will bring more human and computational resources into the project.

### 4.1 PRODUCTION TESTBED

The next table shows the list of possible new sites to be added to the production testbed in a near future.

Site	Country	Location	Project
University Complutense of Madrid	Spain	Madrid	External
University of Linz	Austria	Linz	CrossGrid
LIP Coimbra	Portugal	Coimbra	CrossGrid
CSIC additional sites	Spain	Madrid	National Grid

**Table 44. Possible future production sites.**

The addition of new sites will be postponed until the middleware being tested in the test and development testbed (currently EDG 1.4.3) is released into the CrossGrid production testbed. This measure is intended to maintain a high availability level in the current production testbed that is being used to test the application prototypes. Since many of the current production testbed issues can be improved or resolved with the newer middleware, only then the extension of the testbed will be considered.

The addition of future sites may impose also the deployment of new grid Certification Authorities. In the table above it can be seen the presence of Austria, a country where those services aren't yet deployed. Similarly the distinguish names of the new user certificates will need to be registered into a VO.

The registration of external user certificates into a VO is still being studied. The issue here is whether external users should be registered into the CrossGrid VO as any other project user, or whether these users should be registered in a different VO or in a dedicated group under the CrossGrid VO. Probably the last option will be chosen since it has the advantage of showing all users under the CrossGrid domain and simultaneously allowing site administrators to choose between authorizing project users only or project users and external users. The registration of external users into the VO will also be subjected to the signature of a testbed usage policy.

All sites to be added will follow the normal CrossGrid site approval procedure where sites to be added are subjected to a validation process performed by members of the testbed quality assurance (task 4.4). The validation process verifies the site configuration and functionality through the submission of jobs first through Globus and later through the resource broker, the site information published through MDS and the GSI file transfer service are also verified. During the whole validation process the network connectivity is monitored. When problems are found they are reported back to the site administrator with information about the possible

problem origin. This process implies a strong cooperation between the test and validation team and the site administrator. The test and validation team reports problems and the site administrator takes actions to correct them. Once the reported problems are corrected a new iteration is performed to validate the site. If problems aren't found the site is declared correctly deployed and is finally added to the resource broker. In the validation process a tool developed at LIP to detect most of the common configuration and installation issues (named Host Check) is used. The tool also produces daily web pages with reports that can be consulted by the site administrator.

#### 4.2 TEST AND VALIDATION TESTBED

Has explained in the test and validation description there are currently three sites in the test and validation testbed. However it is expected that at least two more sites will join the test and validation testbed. The complete list of test and validation sites is described in the next table.

Site	Location	Status
FZK	Karlsruhe	Non dedicated infrastructure deployed
IFIC	Valencia	Not currently in the test and validation testbed
USC/CESGA	Santiago	Not currently in the test and validation testbed
Demokritos	Athens	Non dedicated infrastructure deployed
LIP	Lisbon	Dedicated infrastructure deployed

Table 45. The complete list of test and validation sites.

Since the test and validation testbed must be kept small and well maintained the inclusion of more sites is not expected. However collaboration on testing activities with other projects may contribute to extend the testbed to include a limited set of external sites. This was the case of the IFAE DataGrid site in Barcelona with which several replication tests have already been conducted.

#### 4.3 CENTRAL SERVICES

A viability study to connect the systems hosting the CrossGrid central services directly to the backbone of the Portuguese Academic Network (FCCN) is also being conducted. The systems would be hosted at the Network Operations Center of the Portuguese Academic Network in Lisbon enjoying a privileged connection to the Géant infrastructure and a more reliable environment including a diesel powered electric generator. However the systems would still be completely controlled and managed by the LIP personnel. For that purpose a remote power switch and a remote manageable KVM switch would be provided to support the server's remote management. Some tests will be carried out at the Portuguese Academic Network premises in February of 2003 aiming to validate the concept.

In a first phase only the production systems that would benefit most from the bandwidth increase and lower latency will be moved to the network backbone. These systems are the RB, RC and II. However the deployment of an installation server to support locally the maintenance and upgrade of these services will be also required.

This is the result of a close collaboration between LIP and the Portuguese Academic Network in the field of grid computing and networking.

If successful the migration of the production central systems to the new location will probably take place after the new EDG 1.4.x middleware is deployed in the testbed.

## 5 TESTBED USAGE

This section contains simple usage examples based on the current production testbed running EDG 1.2.x. The examples are intended for new users wishing to use the testbed to run their applications.

Grid operations should be carried out from a properly configured user interface system. This is especially important when performing job submissions through the RB. These examples assume that all commands are being issued from a user interface.

### 5.1 THE USER CERTIFICATE

The user certificate and the corresponding private key must first be copied to a directory called *.globus* in the user home directory. The certificate must be stored in a file name *usercert.pem* and the private key in a file called *userkey.pem*. The file permission for *usercert.pem* must be 444 and for *userkey.pem* 400.

Once the files are available in the *.globus* directory the user can invoke the command *grid-proxy-init* to create a proxy certificate that will be used to authenticate the user in the testbed. The first command example creates a proxy valid for 12 hours. The command will ask for the certificate private key password.

```
$ grid-proxy-init -hours 12
```

The next command example can be used to create a proxy from a certificate and private key stored in a floppy disk or in a non-standard directory. However the floppy file system must first be mounted. The command will ask for the certificate private key password.

```
$ grid-proxy-init -cert /mnt/floppy/usercert.pem \  
-key /mnt/floppy/userkey.pem
```

For higher security the proxy certificate should be destroyed before logging out. This can be accomplished with the command *grid-proxy-destroy*.

```
$ grid-proxy-destroy
```

The command above removes the user proxy certificate. This is especially important when the proxy has a long lifetime. Users must understand that the private keys of proxy certificates aren't encrypted and therefore are only protected through the file system security. If a system where a valid proxy exists is hacked the proxy certificate can be used to impersonate the user until it expires.

To obtain information about an already created proxy use the command *grid-proxy-info*. The next example shows how to obtain all the relevant information from an existing proxy.

```
$ grid-proxy-info -all
```

## 5.2 RUNNING EXECUTABLES REMOTELY WITH GLOBUS

The job submission under Globus can be performed in different ways. This section presents some job submission examples.

The first example shows how to obtain the list of files contained in a remote home directory. The command runs in foreground and sends the output to the user terminal.

```
$ globus-job-run zeus.cyf-kr.edu.pl /bin/ls -l
```

The command above performs a `ls -l` command in the remote system `zeus.cyf-kr.edu.pl`. When specifying system commands or any other executable the full pathname must be provided therefore the full pathname for `ls` is specified. Arguments following the executable pathname are considered as arguments for the executable.

More than one program can be run in a single job submission. For that purpose the shell must be invoked.

```
$ globus-job-run zeus.cyf-kr.edu.pl /bin/sh -c "cd /tmp; pwd"
```

Each site can have more than one job manager. Usually a fork job manager and a PBS job manager are available. The fork job manager runs the job in the remote Gatekeeper system while the PBS job manager sends the job to a PBS queue, therefore the job will run in a farm worker node. In most cases the jobs should be submitted to the PBS job managers since this is the only way of using the remote worker nodes. The next example shows how to make sure that the submitted job is sent to a PBS scheduler at a remote site.

```
$ globus-job-run zeus.cyf-kr.edu.pl/jobmanager-pbs /bin/hostname
```

The next example shows how to make sure that the submitted job is sent to a fork job manager in the remote Gatekeeper.

```
$ globus-job-run zeus.cyf-kr.edu.pl/jobmanager-fork /bin/hostname
```

The hostname obtained through the two commands above should be different. The first should display the name of a worker node while the second should display the name of the Gatekeeper itself.

Since the user home directories are shared between all worker nodes and the Gatekeeper any command submitted to either the PBS or the fork job managers will see the same user files.

Another common problem faced by the testbed users is how to run in another grid site a program that was compiled locally and that remains in the User Interface machine. There are two solutions. The first is to copy the file to the remote site, an example of such copy procedure is provided in the next section. The second is to stage the executable. Staging is a feature provided by the Globus job submission software that copies the executable to the remote system prior to execution. The next example shows an example of how to stage and execute a program using *globus-job-run*.

```
$ globus-job-run ce.grid.cesga.es/jobmanager-pbs -s \  
/home/user/myprogram.exe
```

Prior to submit a job one might want to know whether the target site has CPUs available in the PBS farm. This information can be obtained easily from the MDS information system running in the site Gatekeeper. The next example shows how to obtain this information.

```
$ grid-info-search -x -h ce.grid.cesga.es -p 2135 -b 'o=Grid' \  
FreeCPUs | grep FreeCPUs
```

In the same way it is possible to know whether the user certificate is authorized to access the remote site.

```
$ DN=`grid-proxy-info -subject | rev | cut -f2- -d/ | rev`\  
$ grid-info-search -x -h ce.grid.cesga.es -p 2135 -b 'o=Grid' \  
AuthorizedUser | grep "$DN"
```

All the information published by the Gatekeeper MDS can be obtained through the following simpler command.

```
$ grid-info-search -x -h ce.grid.cesga.es -p 2135 -b 'o=Grid'
```

### 5.3 FILE TRANSFER

File transfer between grid systems can be accomplished through the *globus-url-copy* command.

```
$ globus-url-copy file:/// $PWD/test.dat \  
gsiftp://lngrid02.lip.pt/tmp/test.dat
```

The example above shows how to copy a file named *test.dat* from the current directory to the */tmp* directory in a remote CE called *Ingrid02.lip.pt*. As it can be seen both the source and destination full pathnames must be specified. This limitation raises the issue of obtaining the home directory pathname. In fact the user home directory pathname is different from site to site. Therefore before copying a file to a remote site one must obtain the home directory pathname at the site. The pathname can be obtained through a simple *globus-job-run*.

```
$ globus-job-run Ingrid02.lip.pt /bin/pwd
```

The above command will print the user home directory at *Ingrid02.lip.pt*. The pathname can then be used to perform the copy as it can be seen below.

```
$ REMOTE_PATH=`globus-job-run Ingrid02.lip.pt /bin/pwd`  
$ globus-url-copy file:/// $PWD/test.dat \  
  gsiftp://Ingrid02.lip.pt/$REMOTE_PATH/test.dat
```

The next example shows how to copy an executable using *gsiftp*. The executable is copied as any other file. However after the copy the file execution bit must be set.

```
$ REMOTE_PATH=`globus-job-run Ingrid02.lip.pt /bin/pwd`  
$ globus-url-copy file:/// $PWD/prog.exe \  
  gsiftp://Ingrid02.lip.pt/$REMOTE_PATH/prog.exe  
$ globus-job-run Ingrid02.lip.pt /bin/chmod a+rx prog.exe
```

Although *gsiftp* is the recommended copy method there are incompatibilities between different implementations of the *gsiftp* client and server. Namely the *gsiftp* client provided with EDG 1.2.x is incompatible with the *gsiftp* server bundled with EDG 1.4.x.

A simpler method to copy files between grid nodes is to publish the files in a web server. A job can then be submitted to the remote site (using *globus-job-run*) to perform the copy with *wget*. The next example illustrates this approach. It assumes that the file *test.dat* was previously published in *http://www.lip.pt/~john/test.dat*.

```
$ globus-job-run Ingrid02.lip.pt /usr/bin/wget \  
  http://www.lip.pt/~john/test.dat
```

## 5.4 SUBMITTING JOBS WITH GLOBUS

Globus also allows batch job submissions where the job is submitted detached from the user terminal. The user can submit the job and obtain the output later. Five commands are used for this purpose:

- **globus-job-submit:** is used to submit the batch job.
- **globus-job-get-output:** is used to obtain the job output.
- **globus-job-clean:** is used to remove the job files.
- **globus-job-cancel:** is used to cancel a job execution.
- **globus-job-status:** shows the status of a job.

The next example shows how to submit a batch job with Globus.

```
$ globus-job-submit Ingrid02.lip.pt/jobmanager-pbs /bin/hostname
```

The command above will submit a batch job to the PBS job manager in *Ingrid02.lip.pt*. The *globus-job-submit* will print an URL that is used as a job identifier. The job identifier is required for obtaining the job output, clean the job files, obtain the job status and cancel the job.

The next examples assume that the job identifier returned by the *globus-job-submit* command was *https://Ingrid02.lip.pt:2012/1787/1042764379/*.

Prior to obtaining the job output one might want to check whether the job has finished. This can be accomplished with the command *globus-job-status*. When a job finishes the command will report the status as DONE. Other possible status are PENDING, ACTIVE, SUSPENDED and FAILED.

```
$ globus-job-status \  
https://Ingrid02.lip.pt:2012/1787/1042764379/
```

To obtain the job output use the next command that will print the job output in the terminal screen.

```
$ globus-job-get-output \  
https://Ingrid02.lip.pt:2012/1787/1042764379/
```

Cleaning the job output files can be performed with the command *globus-job-clean*.

```
$ globus-job-clean \  
https://Ingrid02.lip.pt:2012/1787/1042764379/
```

Finally the next commands can be used to cancel a previously submitted job.

```
$ globus-job-cancel \  
  https://Ingrid02.lip.pt:2012/1787/1042764379/  
$ globus-job-clean \  
  https://Ingrid02.lip.pt:2012/1787/1042764379/
```

## 5.5 SIMPLE JOB SUBMISSIONS WITH RSL

Besides the simple job execution commands already explained (*globus-job -run* and *globus-job-submit*) there is a third lower level command called *globusrun*. The job submissions are actually performed by the command *globusrun*. In fact the commands previously explained are just shell scripts that invoke *globusrun* after creating an RSL job description. RSL is the Globus Resource Specification Language and is used to describe the job requirements.

The following two commands perform exactly the same operation. However the first uses *globus-job-run* while the second uses *globusrun*.

```
$ globus-job-run bee001.ific.uv.es /bin/ls -l -a  
$ globusrun -r bee001.ific.uv.es \  
  '&(executable=/bin/ls)(arguments="-l -a")'
```

The *-r* option in the *globusrun* command indicates that the next option is a resource (Gatekeeper name). The last command option is a RSL specification containing two attributes (*executable* and *arguments*).

RSL has many attributes and can be quite complex. Usually the RSL specification is written into a file and then passed to *globusrun* as in the next example.

```
$ cat > test1.rsl <<EOF  
&  
(executable=/bin/ls)  
(arguments=-l -a)  
(maxCpuTime=1)  
(stdout=/tmp/xxx.lis)  
(directory=/etc)  
EOF  
$ globusrun -r bee001.ific.uv.es -f test1.rsl
```

The example above creates a RSL job description that changes the working directory to */etc* And performs the command *ls -l -a* storing the execution output in */tmp/xxx.lis*. The maximum execution time is specified as 1 minute.

## 5.6 FILE COPY AND FILE ACCESS USING GASS

Besides the file copy methods already presented another way of performing file transfers and file access is by using the Globus GASS. The Globus GASS server is basically a web server that allows the put and get of files between a GASS server and a client such as *globus-url-copy*. This is the mechanism used by the executable staging feature previously explained. The GASS server must be started from the user account. Unfortunately this means that firewalls configured to prevent connections from external systems to user daemons running in local systems may block GASS. The classic usage of GASS is with *globusrun* or *globus-job-run* by specifying the option `-s`. This option starts a GASS server in the UI machine that can then be used by the submitted process to copy files to or from the UI. The next example shows how to stage and execute a program using RSL and *globusrun*.

```
$ globusrun -s -r ce001.crossgrid.fzk.de \
  '&(executable=$(GLOBUSRUN_GASS_URL) # "/bin/lis")'
```

The following example performs a *grep* remotely (searching for the word *quark*) over a local file. Both the *grep* command and the input file are transferred using GASS.

```
$ globusrun -s -r ce001.crossgrid.fzk.de \
  '&(executable=$(GLOBUSRUN_GASS_URL) # "/bin/grep") \
  (arguments=quark) \
  (stdin=$(GLOBUSRUN_GASS_URL) # "/home/john/mydb.txt")'
```

The following example starts the GASS server on the Gatekeeper side. This approach has more changes to succeed.

```
$globus-job-run zeus.cyf-kr.edu.pl \
  /bin/sh -c "export GLOBUS_INSTALL_PATH=/opt/globus; \
  /opt/globus/bin/globus-gass-server -c" | \
  ( read URL; \
  globus-url-copy file:/// $PWD/test.txt \
  ${URL}/home/wpsix001/test.txt; \
  globus-url-copy file:/dev/null \
  ${URL}/dev/globus_gass_client_shutdown )
```

The example above starts the GASS server remotely with the option `-c` that enables the clients to shutdown the server. The URL of the GASS server is written to the standard output by the *globus-job-run* and sent through a pipe to a shell process running locally in the UI. The shell process reads the URL from the pipe and uses it to perform the file transfer of *test.txt* from the UI to the remote Gatekeeper using *globus-url-copy*. A second *globus-url-copy* is then performed to shutdown the remote GASS server.

## 5.7 MPI PARALLEL JOB SUBMISSION

Parallel jobs using MPICH-G2 can be submitted using several methods. The first method uses *globus-job-run* to submit and MPI job to three CrossGrid sites.

```
$ globus-job-run \
  -: Ingrid02.lip.pt/jobmanager-pbs      -np 1      -env \
    GLOBUS_DUROC_SUBJOB_INDEX=0        -s mpi1.exe \
  -: bee001.ific.uv.es/jobmanager-pbs   -np 1      -env \
    GLOBUS_DUROC_SUBJOB_INDEX=1        -s mpi1.exe \
  -: zeus.cyf-kr.edu.pl/jobmanager-pbs  -np 2      -env \
    GLOBUS_DUROC_SUBJOB_INDEX=2        -s mpi1.exe
```

The example above uses 4 processors one at LIP, one at IFIC and 2 at Cyfronet. The environment variable `GLOBUS_DUROC_SUBJOB_INDEX` is required to identify each site/instance and therefore must be different for each resource. The next example shows how to perform the same job submission with *globusrun* and RSL.

```
$ cat > mpi1.rsl <<EOF
+
( &(resourceManagerContact="Ingrid02.lip.pt/jobmanager-pbs")
  (subjobStartType=strict-barrier)
  (label="subjob 0")
  (executable=$(GLOBUSRUN_GASS_URL) #
    "/home/john/ExampleApplication.exe")
  (environment= ("GLOBUS_DUROC_SUBJOB_INDEX" "0"))
  (count=1)
)
( &(resourceManagerContact="bee001.ific.uv.es/jobmanager-pbs")
  (subjobStartType=strict-barrier)
  (label="subjob 1")
  (executable=$(GLOBUSRUN_GASS_URL) #
    "/home/john/ExampleApplication.exe")
  (environment= ("GLOBUS_DUROC_SUBJOB_INDEX" "1"))
  (count=1)
)
( &(resourceManagerContact="zeus.cyf-kr.edu.pl/jobmanager-pbs")
  (subjobStartType=strict-barrier)
  (label="subjob 2")
  (executable=$(GLOBUSRUN_GASS_URL) #
    "/home/john/ExampleApplication.exe")
  (environment= ("GLOBUS_DUROC_SUBJOB_INDEX" "2"))
  (count=2)
)
)
EOF

$ globusrun -q -s -f mpi1.rsl
```

## 5.8 JOB SUBMISSION THROUGH THE RB

As shown in the previous sections submitting a job can be done entirely with Globus. However Globus does not have any capability to automatically select the best possible sites where the job can be run. With Globus the user must perform the site/resource selection manually or a program to perform it must be written for the intended application. This is the advantage of the EDG Resource Broker. The EDG RB performs the matchmaking between the required program resources and the testbed available resources.

To use the RB the user must describe the job requirements in a file written in JDL (not RSL). The Condor-G software that is used to perform the matchmaking requires the usage of JDL. Since the Gatekeepers don't understand JDL a translation to RSL is automatically performed by the RB after the matchmaking is performed and before the actual job submission to the Gatekeeper.

The first example shows how to list the testbed resources that match the job requirements without actually submitting the job.

```
$ cat > test1.jdl <<EOF
Executable      = "myscript.pl";
StdOutput       = "test1.out";
StdError        = "test1.err";
InputSandbox    = {"test1.pl"};
OutputSandbox   = {"test1.out", "test1.err"};
EOF

$ dg-job-list-match test1.jdl
```

The example shows first a simple JDL specification to execute a Perl script named *myscript.pl*. The output is written in a file named *test1.out* while the errors are sent to *test1.err*. The input sandbox contains the files that must be transferred to the remote site prior to execution while the output sandbox contains the files that must be transferred back at the end. In this case the Perl script is transferred to the remote site before the execution and the output and error files are transferred back in the end.

To actually submit the job the command *dg-job-submit* is used. If successful in transferring the job to the RB the command will print a job identifier in the form of an URL.

```
$ dg-job-submit test1.jdl
```

The job identifier is required to obtain information about the job, retrieve the job output, cancel the job or remove the output sandbox files produced.

```
$ dg-job-status \
https://lngrid06.lip.pt:7846/193.136.90.139/191119125120405?lngrid
06.lip.pt:7
```

A typical *dg-job-status* is shown below.

```
*****
BOOKKEEPING INFORMATION:

Printing status info for the Job : https://lngrid06.lip.pt:7846/193.136.90.139/1
91119125120405?lngrid06.lip.pt:7
---
dg_JobId           = https://lngrid06.lip.pt:7846/193.136.90.139/1911191
25120405?lngrid06.lip.pt:7
Status             = OutputReady
Last Update Time (UTC) = Fri Jan 17 19:12:34 2003
Job Destination    = zeus.cyf-kr.edu.pl:2119/jobmanager-pbs-workq
Status Reason      = terminated
Job Owner          = /C=PT/O=LIP/OU=Lisbon/CN=Jorge Oliveira Gomes
Status Enter Time (UTC) = Fri Jan 17 19:12:34 2003
*****
```

From the output it can be seen that the job was submitted to a PBS job manager at the gatekeeper *zeus.cyf-kr.edu.pl* and that the job status is “*outputReady*”. The “*outputReady*” status means that the job has ended and the output sandbox files have been transferred from the remote gatekeeper to the RB system and are now available to be transferred to the UI.

To transfer the output sandbox files from the RB to the UI the user must issue the command *dg-job-get-output*. The files are transferred to a directory under */tmp*.

```
$ dg-job-get-output \
https://lngrid06.lip.pt:7846/193.136.90.139/191119125120405?lngrid
06.lip.pt:7
```

To cancel a submitted job the user must issue the command *dg-job-cancel* with the job id as argument.

```
$ dg-job-get-output \
https://lngrid06.lip.pt:7846/193.136.90.139/191119125120405?lngrid
06.lip.pt:7
```

## 6 CERTIFICATION AUTHORITIES

Most certification authorities (CA's) required by CrossGrid are ready and issuing certificates. A major effort was carried out to deploy Grid CA's in countries where they didn't previously exist. At the same time the CrossGrid CA team has worked hard to verify the certification policies and practice statements from the several new CA's to make them trustable and acceptable and by all partners.

Close collaboration with the DataGrid certification authorities task force has been established aiming to coordinate CA activities and mutual cross acceptance of CA's used by the two projects. This was a major step towards the interoperation of both testbeds. DataGrid now accepts all CrossGrid CA's.

The extension of the CrossGrid testbed to new sites might still imply the creation of new Certification Authorities. This will be the case of the University of Linz since this will be the first Austrian grid site. CrossGrid will provide support for the deployment of national CAs whenever necessary contributing to the deployment of the grid building blocks and therefore to the extension of grid technologies across Europe.

The following table shows the list of CrossGrid CA's and their status.

Country	CrossGrid	DataGrid
Poland	Deployed and accepted	Accepted in October 2002
Netherlands	Accepted	Already in DataGrid
Germany	Deployed and accepted	Accepted in June 2002
Slovakia	Deployed and accepted	Accepted in October 2002
Ireland	Accepted	Already in DataGrid
Spain	Accepted	Already in DataGrid
Greece	Deployed and accepted	Accepted in October 2002
Portugal	Accepted	Already in DataGrid
Cyprus	Deployed and accepted	Accepted in February 2003

**Table 46. Certification authorities.**

The deployed CA's mentioned above are issuing X.509 authentication certificates to the CrossGrid users and systems involved in the current testbed activities. One of the most frequent problems found in the setup of the CrossGrid sites until recently was the installation of the certificates and CRL distribution points for the new CrossGrid CA's. The problem appeared since the new CA's were not being distributed nor installed within the standard EDG middleware. Therefore new RPMs had to be built to solve this problem until DataGrid could recognize the CA's. Currently all CA certificates and distribution URLs are installed properly. The CA's are generally behaving correctly regarding both the certificate issuance and the CRL issuance, which is essential for stable testbed operation. The CrossGrid CVS repository has been established at Karlsruhe and will play an important role as the main distribution point for all the CrossGrid RPMs including the CA RPMs.

---

## 7 SUPPORT

This section contains information about the support services available for initial testbed.

### 7.1 HELPDESK

Although the CrossGrid helpdesk is still being established by IFIC in Valencia minimum helpdesk services for the initial testbed are already available. Users and site administrators can obtain help by sending Email to the helpdesk mailing list [CrossGrid-wp4-support@lists.cesga.es](mailto:CrossGrid-wp4-support@lists.cesga.es). Problems and questions regarding the central services and site configuration can be sent directly to [grid.support@lip.pt](mailto:grid.support@lip.pt). The first version of the helpdesk software is being tested at selected sites.

### 7.2 VO SUPPORT

One of the most important central services is the Virtual Organizations server that contains the CrossGrid VO. To use the CrossGrid VO for user authorization some configuration steps have to be performed in all testbed systems. The support for VO site configuration is provided through the central services web pages at <http://www.lip.pt/computing/cg-services>. Users wishing to join the CrossGrid VO must send their request by Email with an explanation of their participation in the project and including the X.509 user certificate attached. A web interface is being developed to facilitate the VO registration process.

### 7.3 WEB PAGES

The central CrossGrid testbed web site is the WP4 web site at <http://grid.ifca.unican.es/CrossGrid/wp4>, however web pages containing information about the central testbed services are available at the LIP web server. The main link for these pages is <http://www.lip.pt/computing/cg-services>. These web pages also contain pointers to usage statistics and information on how to configure the testbed sites to use the central services. Useful information on how to diagnose problems and a list of common configuration problems is also provided. The pages are both a central place to find updated information about the testbed and a first level user support tool where users can cross check their site configurations and obtain helpful hints on how to solve common problems. A second web page dedicated to the validation testbed is available at <http://www.lip.pt/computing/cg-tv-services>.

### 7.4 MAPCENTER

Users and site administrators can obtain information about the testbed status from the mapcenter web pages available at <http://mapcenter.lip.pt>. Mapcenter is an important tool for testbed support since it provides quick global view of the testbed status and also historical data can be used to identify the origin of problems occurred in the past. The mapcenter web pages are extremely important for site administrators since they allow them to check the site services connectivity.

### 7.5 CVS

The CrossGrid CVS repository deployed at Karlsruhe is also an important service for the testbed deployment. The repository already contains the LCFG configuration profiles of

---

several deployed sites. New sites can use these profiles as a reference on how to configure and deploy LCFG successfully. The CVS repository main web page is <http://gridportal.fzk.de>.

## 7.6 STATISTICS

Web pages containing information about the RB and CE usage are available from the central services web page at <http://www.lip.pt/computing/cg-services> and also from the Mapcenter web site at <http://mapcenter.lip.pt>. These statistics are useful not only to have a glance of the testbed usage but also to detect problems in the testbed since they also contain information about the number of job submission errors.

## 7.7 HOST INSTALLATION AND CONFIGURATION VERIFICATION

A system installation and configuration verification tool was also developed. This tool uses several methods to verify the correct behavior and configuration of CE and SE systems. The resulting information is published as a set of web pages available at [http://www.lip.pt/computing/projects/crossgrid/crossgrid-services/site\\_check](http://www.lip.pt/computing/projects/crossgrid/crossgrid-services/site_check).

---

## 8 TESTBED DEPLOYMENT

This section contains information about the experience with the testbed middleware deployment.

### 8.1 INSTALLING THE MIDDLEWARE

The installation is performed through the LCFG installation software developed by the University of Edinburgh and modified by DataGrid. Once the LCFG installation server is correctly deployed installing the site systems consists basically in defining the required LCFG client profiles and adding the machine to be installed to the DHCP server. Therefore this section is centred on LCFG since the middleware installation is very similar for all grid systems differing only on the configuration.

#### 8.1.1 LCFG

LCFG requires the installation of a dedicated machine to host the installation server that in turn will be used to install all grid systems at the site. A basic system with RedHat 6.2 must first be installed on the machine where LCFG is to be deployed. LCFG is installed over this system that must satisfy some software requisites such as the presence of an Apache web server to publish the client profiles, a DHCP and PXE servers to boot the client systems, a NFS server to export the installation mini root and the directories containing the software packages to be installed in the clients.

The DHCP server is an essential component of the LCFG installation server. DHCP is responsible for the distribution of IP addresses and IP configuration information to the clients systems. DHCP can also be used with PXE to boot the client systems directly from the network instead of using a floppy disk. Using PXE makes the installation simpler and in some cases the installation can even be performed remotely without any physical contact with the system being installed. Using PXE also makes easier to support systems with network adapters not included in the standard LCFG boot floppy disk.

Once the LCFG software is installed the installation repository must be filled with the standard RedHat 6.2 RPMs and the EDG RPMs. The installation of a client is performed in two steps.

First the client system boots and obtains an IP address and eventually the kernel from the DHCP and PXE servers. Then the client mounts the LCFG server directory containing a mini root file system that is used to install the RedHat operating system into the local disk. When the first step is completed the system reboots and is ready for phase two, where the specific software packages corresponding to the desired system functionality are installed.

In phase two the system fetches a specific profile from the LCFG server and installs all required RPMs, then the software is configured through LCFG object scripts. The scripts use the LCFG client profile to perform all the required changes to the system configuration files.

The main problems with this procedure are the retrieval of the correct versions of all required RPMs, create the RPM lists in such a way that no dependency problems may appear and create the LCFG client profiles. The creation of the profiles is the most difficult step; this is

---

due to the syntax of the profiles that use the C pre-processor language and to the lack of good documentation and intrinsic complexity of the software. In fact all system configuration (and not only the grid middleware) is performed through LCFG making the creation of the profiles difficult and error prone. The creation of the profiles is done usually using a try and error approach where profiles changes are performed followed by a test installation, if the installation fails (which usually occurs) a new round of corrections and trial is performed. However once LCFG is installed and the profiles are created successfully installing new client systems is just a matter of copying the correct profile and reconfigure the DHCP server.

## 8.2 CONFIGURING THE MIDDLEWARE

As explained in the previous section the systems configuration is mostly performed through the definition of LCFG profiles. However even using an automated installation and configuration system some site-specific configuration issues have to be taken into account, which are:

- The usage of the same UNIX account names and UNIX UIDs across the Gatekeeper, Storage Element and Worker Nodes. This is required for NFS file sharing and to allow users to have the same home directory across all site systems.
- The share of the user home directories across the Gatekeeper, Storage Element and Worker Nodes. This is required to allow users to see the same home directory in all the site systems and each time they submit a job.
- The share of the gatekeeper directory `/etc/grid-security/certificates` with the WN's. This is required since the WN's default configuration does not start the certificates CRL upgraded daemon hence either the daemon is enabled in the WN's or preferably the Gatekeeper directory containing the CRL's must be exported through NFS to all WN's.
- Certificates must be issued for the RC, CE, SE, RB and MyProxy servers.
- The CA certificates from all the recognized Certification Authorities must be installed. For CrossGrid this means that additional certificates to the ones in the EDG distribution must be available in CE, SE, UI, WN's, RB, MyProxy, and RC.
- The configuration file `mkgridmap.conf` in CE, SE and RB systems must contain the authorized VO's. For CrossGrid these are the `crossgrid` VO for users and the `gdmppservers` VO for SE systems running GDMP.

The following sections explain more deeply the configuration of the grid middleware for each type of system.

### 8.2.1 CE/Gatekeeper

The CE is installed and configured mostly from LCFG. Changes had to be performed to the standard configuration to use the CrossGrid central services such as the CrossGrid VO server. Some configuration steps must be performed manually these are:

- The PBS batch scheduler configuration.
- The polled accounts configuration and creation of empty files in `gridmapdir`.
- Insert the machine certificate.

### 8.2.2 WN

The WN is installed and configured completely from LCFG. The installation of grid middleware in the WN is minimal and it doesn't require any daemons or inbound network connectivity. The WN requires outbound network connectivity to transfer the files in the input and output sandboxes for jobs submitted through the RB. The only grid service required is the Globus URL copy file transfer client and the replica manager client. However the WN requires the installation and configuration of the support libraries for the intended applications also performed through LCFG.

### 8.2.3 SE

The SE is installed and configured completely from LCFG. Changes had to be performed to the standard configuration to use the CrossGrid central services such as the RC and VOs. Besides the CrossGrid VO the SE also uses the gdmbservers VO to authorize file replications with other SE's via GDMP. The support for GDMP file replication is the most difficult component to configure in the SE requiring separated authorization and configuration files per VO. Besides GDMP support, RFIO and the MDS schema the SE configuration is very similar to a CE. Most SE's even have a gatekeeper fork job manager that allows debug access to the system

### 8.2.4 UI

The UI is installed and configured completely from LCFG. Changes had to be performed to the standard configuration to use the CrossGrid central services such as the RB, LB and RC.

### 8.2.5 RB

The RB is installed and partially configured from LCFG. Changes had to be performed to the standard configuration to use other CrossGrid central services such as the CrossGrid VO, the RC and MyProxy servers. The addition of testbed sites to the RB is described in the maintenance section further ahead.

Post install configuration is done manually:

- Creation of the initial databases:
  - MySQL for the LB service;
  - PostgreSQL for the RB and JSS service;
- Setup of CondorG.
- Configuration of the daemons for the LB, RB and II services.

### 8.2.6 MyProxy

MyProxy is installed using the standard LCFG procedure however configuration steps such as the automatic start-up of the daemon had to be configured manually.

### 8.2.7 RC

The Replica Catalogue installation is performed with LCFG up to the end of the first installation phase. The configuration must be performed manually since there aren't LCFG object scripts to configure the RC. The daemon is started manually, and after one has to create the first file collection in the DB. Another problem is the lack of information on how to setup the RC completely. The installation of the CrossGrid RC was possible due to the help and teaching provided by DataGrid RC experts.

### 8.2.8 VO server

The VO server presents a situation similar to the RC. The VO server doesn't have RPMs or LCFG installation scripts. The CrossGrid VO server was installed over an already running system, LCFG was not used and is not required for the installation or configuration of the server. The installation documentation is available in the form of a white paper unfortunately the document is centred on the VO management tools and not on the VO server setup. The VO server was installed by analysing the management tools and the data in the existing DataGrid VO servers. Several problems and insufficiencies have been found in the management scripts, the tools have been corrected and enhanced to satisfy the CrossGrid needs.

### 8.2.9 Mapcenter

Mapcenter is installed in a similar way as the RC. A base system is installed with LCFG and the Mapcenter software is then installed and configured manually. There isn't a profile for the Mapcenter system therefore a profile had to be written for it. The created profile is basically a mix between a CE and a UI and it has been chosen to enable the Mapcenter system to function as a grid-monitoring box with the capability of running future grid diagnostics that are to be integrated into Mapcenter.

The installation of Mapcenter was performed in close contact with DataGrid WP7 network monitoring experts that also provide the source code (not previously available) and additional support on the configuration and usage of undocumented features. The collaboration with DataGrid also had as result the addition of new functionalities into Mapcenter allowing the integration of the statistics modules developed at LIP.

---

## 9 TESTBED MAINTENANCE

The testbed activities aren't limited to the site and services deployment. After the deployment systems and services must be maintained and monitored to assure high availability and smooth operation as possible.

### 9.1 VO MAINTENANCE

For users to get authorized to use the CrossGrid testbed they must be registered into the CrossGrid VO. Thus the maintenance of the VO database is an essential activity for the testbed operation. The addition of new users to the VO follows a procedure where VO membership requests containing the user certificate are sent by Email to [vo.admin@lip.pt](mailto:vo.admin@lip.pt) and verified. The verification is performed by checking the certificate authentication against the corresponding CA certificate, if found authentic the certificate holder name and email are checked against the list of CrossGrid members available from the main crossgrid web server <http://www.eu-crossgrid.org>. When the user is unknown his team leader or site responsible are contacted to verify whether the user actuality is a CrossGrid project member. Once the checks are passed successfully the user is added to the VO server and receives an Email confirming his membership and explaining that due to the characteristics of the VO authorization system the update of all the testbed authorization files may take up to 24 hours therefore only at that point the user will be accepted at all testbed sites. Currently more than 30 users are registered into the crossgrid VO.

Since the VO server protection is a key point for the testbed security it also requires careful maintenance and monitoring. The machine hosting the VO server contains a limited set of services and its operating system and software are kept updated. The machine is installed in a separated network properly protected by a firewall, all system activities and log files are closely verified.

### 9.2 RB MAINTENANCE

The two major RB maintenance activities are the addition of new sites and the monitoring to verify that the RB is behaving correctly.

The addition of new testbed sites follows a validation process performed by members of the testbed quality assurance (task 4.4). The validation process verifies the site configuration and functionality through the submission of jobs first through Globus and later through the resource broker, the site information published through MDS and the GSI file transfer service are also verified. During the whole validation process the network connectivity is monitored. When problems are found they are reported back to the site responsible with information about the possible problem origin. This process implies a strong cooperation between the test and validation team and the site responsible. The test and validation team reports problems and the site responsible take actions to correct them. Once the reported problems are corrected a new iteration is performed to validate the site, if problems aren't found the site is declared correctly deployed and is added to the resource broker.

The monitoring of the RB is verified manually through the usage of the workload management UI tools and automatically through the submission of jobs by a verification tool that is still being improved. A tool to verify the testbed site information has also been developed. This tool verifies the information published by CE and SE systems through the MDS GRIS and it can dynamically remove from the RB sites publishing incorrect or

---

incomplete information. The tool can also automatically add the site back into the RB when the situation is corrected by the site administrator. This is a very useful tool since remote MDS GRIS servers behaving incorrectly constitute the cause of most RB problems. Mapcenter is a very useful tool to detect such failures in the services internal to the RB and in remote testbed sites that can potentially disturb the RB behavior. The RB statistics are also a valuable tool to detect problems and their origin. The statistics contain counters for the several steps of a job execution and thus allow the identification of the step where a given job is being lost.

### **9.3 RC MAINTENANCE**

The RC maintenance consists mainly in adding new users and SE's and enable them to write into the RC LDAP server. This is done manually by editing the LDAP server configuration file and restarting the RC service.

### **9.4 MAPCENTER MAINTENANCE**

The Mapcenter maintenance is a task that consists mainly in the addition of new sites and services to the configuration files. Sites are added to Mapcenter when the site administrator reports the site as ready to be integrated. Mapcenter is then used to monitor the site services connectivity and store historical information about alarms. The monitoring of the sites by checking the Mapcenter web pages is also a maintenance task that must be carried out frequently however this is a general testbed maintenance task. To reduce the burden of checking the web pages a new feature is starting to be explored in the CrossGrid Mapcenter. This is the capacity to send Email alerts to site contacts and testbed administrators automatically.

### **9.5 SITE MAINTENANCE**

After a site is deployed its maintenance consists mainly in performing the required and frequent software release upgrades and package updates. The maintenance of a grid site is basically equal to the maintenance of a normal computing infrastructure with a few differences. The configuration and maintenance is mostly performed through changes in the profiles and RPM lists centrally maintained at the local LCFG installation server. Another important maintenance task is to verify frequently the site health status. This is accomplished by looking at the log files, GRIS published information, the process status and by submitting jobs to the site through Globus and through the RB and checking whether the jobs finish successfully. In the same way performing file transfers allows to check the site GSI ftp servers. This task will be performed automatically with scripts that are being developed for test, validation and monitoring of the testbed.

---

## 10 SITE DEPLOYMENT PROBLEMS

This section contains some commonly found configuration problems in newly deployed sites.

### 10.1 DNS AND REVERSE LOOKUPS

This is not a Grid specific problem but a general configuration problem that has affected several CrossGrid sites and that should be taken into account by new sites. While validating remote sites it was noticed that sometimes the DNS name and its corresponding IP reverse mapping do not match. This means that a node is registered into DNS with a name but the reverse translation of the IP address into a name does not match the initially specified name. When this happens the X.509 authentication will fail preventing users to access the system. Great care should be taken when creating or changing DNS records always making sure that the direct and reverse mappings match.

**Applies to: production testbed, test and validation testbed**

### 10.2 NUMBER OF PROCESSORS AND STORAGE SIZE

These are also frequent configuration problems where site administrators do not configure correctly the amount of storage that is published through the SE MDS. The storage capacity number should be specified in GB but frequently is published in MB creating the illusion that much more storage is available than what really is. The same happens with the number of processors. Sometimes changes in the site configuration are not reflected in the local PBS batch system configuration and as a consequence the number of processors announced by the CE MDS is different from the real number of processors. This is mostly due to the lack of an automatic system that should collect information about the real number of processors and publish it into the CE MDS. In CrossGrid a system information data collector is being developed that will verify among other things this configuration problem.

**Applies to: production testbed, test and validation testbed**

### 10.3 POLLED ACCOUNTS

This is a method to map automatically new certificate distinguish names into UNIX accounts without the intervention of the local system administrator. The major problem with this feature is the configuration since the EDG instructions fail to specify that a directory called "gridmapdir" is used to keep track of the mappings established between distinguish names and accounts, and that this directory must contain empty files owned by the root user with names matching the UNIX accounts that belong to the pool. The documentation also fails in specifying that the "gridmapdir" directory must be shared with the Worker Nodes and local Storage Elements through NFS.

**Applies to: production testbed, test and validation testbed**

### 10.4 CA CERTIFICATES

Another common problem is the administrator's failure to install all the require CA certificates and CRL distribution points by using the most updated CA RPMs. This prevents users with certificates signed by the missing CA's from accessing the systems where the CA certificates are missing.

**Applies to: production testbed, test and validation testbed**

---

## 10.5 FIREWALLS

Problems with firewalls are common and not restricted to the Grid. Nowadays firewalls are installed in all sites connected to the Internet and are generally considered a key element of the local network security. The problems appear when the TCP ports used by the grid services are filtered preventing outside systems from connecting to the local grid services. Since the TCP ports used by the grid middleware are blocked by default in all firewalls, they must be specifically opened. The situation gets more complex when the firewall is not under the control of the people involved in the grid site deployment. Another issue is that some grid services weren't developed with network security in mind thus the TCP ports involved are numerous and not contiguous making the firewall configuration more complex and less clean.

**Applies to: production testbed, test and validation testbed**

## 10.6 MDS CONFIGURATION

Configuration problems that lead to the absence of information in the Computing Elements and Storage Elements MDS/GRIS servers are very common. Most information published by the CE's and SE's into the MDS is required by the RB to make matchmaking decisions. If the information is not present or the MDS GRIS server is not responding the site will never be selected as candidate to run jobs.

**Applies to: production testbed, test and validation testbed**

## 10.7 VO SETUP

For users to be authorized to use a site, their X.509 certificate distinguish name must be mapped into a local UNIX account at the site. The retrieval of the authorized distinguish names to build the site local mappings is done through the VO server that provides the list of the user distinguish names. The correct setup of the configuration files used to specify the authorized VO's and the creation and set-up of the corresponding UNIX accounts is fundamental. These steps require additional configurations to be performed but since LCFG doesn't have support to perform these changes automatically they must be performed manually. Unfortunately sometimes administrators either forgot to setup all the VO's or make mistakes in its configuration.

**Applies to: production testbed, test and validation testbed**

## 10.8 GDMP SETUP

GDMP is one of the file replication systems. GDMP will be abandoned in favor of the Replica Manager software however it is still being used namely by some HEP experiments such as CMS. Besides other limitations (such as only allowing replication between SE's) GDMP is difficult to configure and it must be configured manually for each supported VO.

**Applies to: production testbed, test and validation testbed**

## 10.9 INPUT AND OUTPUT SANDBOX TRANSFER

Each job submitted through the RB has an associated input and output sandbox. The sandboxes are sets of input or output files that must be transferred at the start or end of the job. Frequently jobs fail because the sandbox transfer also fails. The major reason for these failures is connected to the misunderstanding of how the sandbox transfer is performed. In fact a job wrapper script performs the sandbox transfer from the RB to the CE. The script is

---

generated by the RB Job Submission Service and performs the transfer of the files mentioned in the input sandbox prior to execute the real user job, it also transfers the files mentioned in the output sandbox back to the RB at the end of the job. The file transfer is performed using GSIFTP. The wrapper script is part of the submitted job and therefore is run in the Worker Node. This means that all Worker Nodes need outbound connectivity, the CA certificates and updated CRL's. Since the daemon that transfers the CRL's is not run by default in the Worker Nodes the directory containing the CA certificates and CRL's should be mounted through NFS from the Gatekeeper. This is not explained in the EDG installation documentation.

**Applies to: production testbed, test and validation testbed**

### 10.10 LCFG AND SYSTEM BOOT

A frequent problem with LCFG is the minimal support for network devices contained in the boot floppies. Since the size of the boot floppy is reduced only a minimal set of network drivers is available by default. If a site administrator has systems with different network adaptors he has to make its own boot floppy which can be a difficult task. LCFG has support to boot directly from the network unfortunately not all systems being deployed at the several sites have support for network boot. Also when DHCP and PXE servers are already installed in the network they may prevent the usage of the network boot feature by answering to the network boot requests before the installation server.

**Applies to: production testbed, test and validation testbed**

### 10.11 LCFG AND DISK PARTITIONS

One annoying feature of LCFG is its requirement of destroying the system partitions content and not support more than 4 partitions per physical disk. This is a big limitation that also prevents the deployment of other operating system versions in the same disk.

**Applies to: production testbed, test and validation testbed**

### 10.12 RPM LISTS

The RPM lists published with each new EDG release frequently have dependency problems that have to be checked and corrected manually.

**Applies to: production testbed, test and validation testbed**

### 10.13 SITE CONFIGURATION

With the current installation system and instructions the EDG installation procedure is still very prone to configuration errors. It is virtually impossible to install a site at the first attempt without ending with major site configuration problems. Installation and configuration is therefore a tedious process of trial and error.

**Applies to: production testbed, test and validation testbed**

#### **10.14 MPI AND PRIVATE IP ADDRESSES**

A major problem for CrossGrid is the impossibility of using MPICH-G2 between sites when the WNs are using private IP addresses. In the current CrossGrid testbeds all WNs have been configured with public IP addresses circumventing the problem. However this insufficiency can be a major obstacle to scalability since most of the large farms existing today use private IP addresses. This problem is being carefully evaluated and a possible solution is being studied.

**Applies to: production testbed, test and validation testbed**

---

## 11 DEVELOPMENTS AND CORRECTIONS

This section covers some developments that have been performed in the context of the testbed quality assurance activities.

### 11.1 CHECKING SITES AND SERVICES

Tests are being developed to verify the correct configuration and functionality of the middleware. Most of these tests are still under development however early versions of some programs are already available. One of these is a tool to verify the MDS data being published by CE and SE systems that is being actively used to remove and insert sites into the RB index server accordingly with their MDS configuration status.

A system installation and configuration verification tool was also developed. This tool uses LDAP searches, gsiftp transfers and globus job submissions to verify the correct behavior of CE and SE systems. The tool also collects configuration information about the hosts and performs several checks to verify its installation and configuration. The tool currently verifies:

- DNS hostname and reverse mapping
- System reachability through ICMP echo
- Several system and grid middleware configuration files
- The gridmapdir directory
- Presence of required processes
- File system space
- Presence of required CA files
- CRL validity
- Job submission
- File transfer with GSIFTP
- PBS configuration
- Presence of worker nodes
- Configured VO's

### 11.2 STATISTICS

Two statistics modules have been developed in Perl and Shell to obtain information about the usage of CE and RB systems and publish it in a web server. The RB statistics module collects information from the Logging and Bookkeeping database while the CE statistics module collects information from the Globus gatekeeper log files. Information is collected twice a day. The information published contains counters with the total values and the variations since the last data retrieval. Collaboration with DataGrid on this subject has started with the manifestation of interest by DataGrid of using the CrossGrid statistics modules.

### 11.3 TEST LIBRARY FUNCTIONS

To cover the needs of the above test and validation tools library functions are being developed. They perform operations that are required by several tools and encapsulate operations that might suffer changes in the future.

---

#### 11.4 MAPCENTER

The Mapcenter tool has been enhanced with the capacity of probing of new services, and the addition of links to web pages containing CE and RB statistics developed at LIP. These statistics pages are described above. Collaboration with DataGrid on the improvement of Mapcenter has started with the implementation of some corrections and changes.

#### 11.5 VO TOOLS

The EDG VO management tools have been modified to support the CrossGrid testbed needs and to correct several problems and insufficiencies. Some new scripts have been written to ease and automate the VO management process.

#### 11.6 VOMS

VOMS is the new EDG software to replace the currently used LDAP based VO servers. VOMS is a more sophisticated system that uses a relational database to store the users rights and roles within a Virtual Organization. VOMS is being tested at LIP.

#### 11.7 LCFG

The LCFG installation profiles provided by DataGrid have been reviewed, cleaned and enhanced. These changes have been reported back so that they can be used in a future EDG release. Several CrossGrid partners have submitted LCFG bug reports and enhancements to DataGrid. An extensive manual about site deployment with LCFG has been written by Demokritos covering all the site deployment steps. The new EDG distribution of LCFGng has been deployed and tested in a small cluster of three machines at LIP using RedHat 7.2. A report of the problems found has been submitted to DataGrid.

---

## 12 TESTS

This section covers the tests being performed over the initial testbed and the perspectives regarding future tests with program prototypes.

### 12.1 BASIC TESTS

Basic tests covering the Globus and EDG middleware functionalities have been performed using the testbed. These tests cover the job submission through Globus, job submission through the EDG RB, file transfer with GSI ftp, file replication with GDMP, file replication with the Replica Manager, the VO server and the MDS information system. These tests were performed on both EDG 1.2.x and 1.4.x and they are repeated for each middleware release. Additionally stress tests were successfully performed in the validation testbed aiming to validate the robustness and performance of EDG 1.4.x.

In addition the first tests with MPI in the testbed have been performed. Since the EDG RB doesn't support parallel applications the program executions have been performed using the Globus job submission capabilities directly. The MPI package used for the tests was MPICH-G2 a variant of the MPICH implementation using Globus IO for inter-process communication and Globus DUROC for resource co-allocation. The first major obstacle was the lack of some required Globus features in the EDG distribution of Globus that prevent the compilation of MPICH itself. The only solution found was to compile and install a full release of Globus in a User Interface and compile MPICH with this complete distribution. However this solution presents an inconvenience. It means that the MPI programs must also be compiled with the complete Globus distribution that is unavailable at other testbed sites. Two solutions were possible, copy the shared libraries to all testbed sites or compile the program statically, the second option was chosen since it allowed proceeding faster and easier. Several small test programs were compiled and run using this approach. Resources in Poland, Germany, Spain and Portugal were used to run different tasks of the same applications with success.

In the same context tests were performed on how to start the parallel applications. MPICH provides its own program start-up command named mpirun. The MPICH-G2 mpirun command has some limitations. It assumes that the executable is already available at all sites and that its pathname is the same everywhere. In most circumstances both these assumptions are incorrect, therefore other approaches were tried first by transferring the files to selected target sites with GSI ftp and storing their locations and later by using the Globus GASS cache mechanism to transfer the executables to the remote sites automatically prior to execution. The last option is the most simpler and effective. Unfortunately for large and complex applications it may take some additional time to transfer the required files.

### 12.2 FIRST PROGRAM PROTOTYPES

The first program prototypes are being developed. A prototype of the HEP requiring MPI and very low latency was released in 2002. Tests aiming to validate the prototype architecture and the testbed middleware with a real CrossGrid application have been conducted in the testbed using MPICH-G2 and MPI traffic across several up to seven sites. At the CrossGrid Santiago workshop the real application was demonstrated running in five testbed sites and producing graphical output. The HEP application is currently being improved and a new round of tests is expected soon. Other applications CrossGrid applications are also being submitted to testbed sites but at this stage without using MPICH-G2 hence running only inside the clusters.

### 12.3 FIRST INTEGRATION

The CrossGrid software integration recently accomplished in February has put heavy load on the production and validation testbeds for several days. The load introduced by the integration work has shown that the current testbeds are capable of supporting the work of many simultaneous users using both parallel and non-parallel applications and tools. The testbed was also used to demonstrate the new software components and applications running over a real pan-European grid infrastructure.

The software integration results will be described separately.

### 12.4 IST DEMONSTRATION

Three CrossGrid sites (FZK, IFIC and LIP) have participated at the IST 2002 demonstration event showing the interoperation of the CrossGrid sites with other testbeds. During the demonstration several jobs have been successfully executed in the three participating CrossGrid sites.

#### 12.4.1 IFIC CE Usage statistics for the IST demonstration period

<b>Total user DNs</b>	<b>3</b>
<b>Total different Ips</b>	<b>5</b>
<b>Total Jobs submitted</b>	<b>118</b>
<b>Total Ping Jobs submitted</b>	<b>5</b>
<b>Total Jobs successfully submitted</b>	<b>91</b>
<b>Total failed jobs</b>	<b>22</b>
- LCAS failed jobs	1
- GSS failed jobs	0
- CRL failed jobs	17
- Jobman failed jobs	0
- Unknown reason	4

Table 47. IFIC CE Usage statistics.

#### 12.4.2 LIP CE Usage statistics for the IST demonstration period

<b>Total user DNs</b>	<b>4</b>
<b>Total different IPs</b>	<b>5</b>
<b>Total Jobs submitted</b>	<b>453</b>
<b>Total Ping Jobs submitted</b>	<b>125</b>
<b>Total Jobs successfully submitted</b>	<b>211</b>
<b>Total failed jobs</b>	<b>117</b>
- LCAS failed jobs	1
- GSS failed jobs	0
- CRL failed jobs	0
- Jobman failed jobs	0
- Unknown reason	116

Table 48. LIP CE usage statistics.

12.4.3 FZK CE Usage statistics for the IST demonstration period

Total user DNs	6
Total different IPs	8
Total Jobs submitted	146
Total Ping Jobs submitted	0
Total Jobs successfully submitted	78
Total failed jobs	68
- LCAS failed jobs	20
- GSS failed jobs	45
- CRL failed jobs	0
- Jobman failed jobs	0
- Unknown reason	3

Table 49. FZK CE usage statistics.

### 13 CONCLUSIONS

Ten sites have been successfully deployed within the CrossGrid production testbed together with all required support services. Several other sites are under acceptance tests or being deployed. The production testbed was used to evaluate the existing middleware developed by Globus and EDG as well as other packages such as MPICH-G2 and simple MPI applications. Two new production sites are under test.

A test and validation testbed infrastructure with three sites was deployed and is being used to evaluate new middleware.

A basic Grid monitoring system, and an installation and configuration system verifier is in place allowing the monitoring of both the production and test and validation testbeds.

Although several problems and limitations have been found it must be understood that the EDG and the Globus middleware distributions are extremely complex and still under active development. Both distributions rely on a number of underlying packages and utilities that also have intrinsic problems that frequently impose limitations. However even with the current limitations the middleware is still quite usable and satisfies many of the CrossGrid requirements.

Three CrossGrid sites in Germany, Portugal and Spain have successfully participated in the IST 2002 worldwide demonstration showing the testbed interoperability and the coordination and management capacity of CrossGrid.

---

## 14 REFERENCE DOCUMENTS

CrossGrid documents containing relevant information referenced in this document, are listed in this section.

### **CrossGrid Middleware Test Procedure**

#### **CG-4-D4.1-004-TEST-1.1**

URL: <http://www.eu-CrossGrid.org/Deliverables/M3pdf/CG-4-D4.1-004-TEST.pdf>  
Jorge Gomes (LIP)

### **CrossGrid Test and Validation Testbed Architecture**

#### **CG4.4-D4.2-v1.0-LIP011-ValidationOfTestbedArchitecture**

URL: <http://www.eu-CrossGrid.org/Deliverables/M6pdf/CG4.4-D4.2-v1.1-LIP011-ValidationOfTestbedArchitecture.pdf>  
Jorge Gomes (LIP), Mario David (LIP)

### **Globus Quick Start Guide**

URL: <http://www.globus.org/toolkit/documentation/QuickStart.pdf>  
The Globus project