

### Fault tolerance in Grid and Grid'5000

IFIP WG 10.4 on dependable Computing and Fault Tolerance

Franck Cappello INRIA Director of Grid'5000 fci@lri.fr

Fault tolerance in GridGrid'5000

### Applications requiring Fault tolerance in Grid

**Domains** (grid applications connecting databases, supercomputers, instruments, visualization tools):

- Finance,
- Health care,
- eScience, Cyber Infrastructure (EGEE, Virtual observatory, TeraGrid, etc.)
- Nature and industrial disasters prevention and management
- etc.

### Key technology:

• Web Services (with some extensions: WSRF)

# The EGEE project (Enabling Grid for E-Science)

•Building and Maintaining a large scale computing infrastructure • Provide support for Scientists using it.

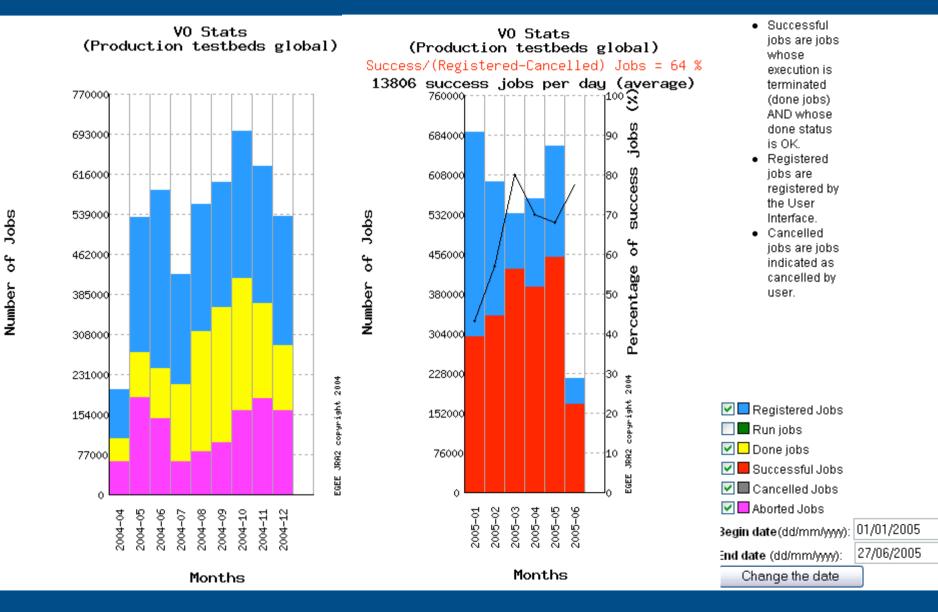
#### Size:

Users: 3000 Institutes: 70 Countries: 27 Sites: 148 CPU: > 13000Disk > 98 PB

Duration: 2 years Cost: 32M Protvino-IHEP SINP ITEP JINR Next: EGE BITLab-LCG SNJ CG2 NIKHEP HU-BEBLIN SARA POZNAN WARSAW WUPPERTAL BWTH GSI PRAGUE KBAKOW CEA-DAPNIA-SACLAY IPSL-IPGP PRAGUE-CESNET FZK SCAI NGS-OXFORDIN2P3-LAL BEIJING IN2P3-LPC INFN-PADOVA MILANO BO.01 IN2P3-CC IN2P3-CPPM TORINO CNAF INFN-CNAF BG01-IP BG02-IM BG-INBN USC INEN-B MAAPOLI INFILINE INTA-CAB HELLASGRID AUTH PieGR01-UoM GR-02-LbM GR-03-HEPNTUA GR-05-DEMOKRITOS TAIWA FORTHICS TAU GOG-ALBERTA **Pilot applications:** TRIUME TRIUMF-GC CARLETON PAKGRID LHC experiment (Alice, Atlas, CMS Sites: 148 PK-NCP TOPONTO UIOWA-LCG2 BNL-LCG  $\rightarrow$  Scale, high bandwidth data tran CFU: >13000 Disk: > 98 PB **Biomedical experiments:** HPTC-LCG2

 $\rightarrow$  Security, Ease of use, distributed data base

# Job Statistics in EGEE (Enabling Grid for E-Science)



# EGEE issues and problems

- Hardware / Software issues
  - Heterogeneous hardware, software, OS are a BIG problems !
  - Example: User Interface
  - Example: floating point accuracy
  - Example: dynamic libraries
  - Example: distributed application across different platforms
  - Revival of the interpreter, JIT ?
  - Security and accounting IntraGrid vs. InterGrid
  - Submission times ???
- Political Issues
  - Different communities different agendas / hidden agendas
  - coordination between partners
  - typical problems of large, heterogeneous organisations
  - small and dynamic vs. large and powerful organisations

# Job Efficiency in EGEE

Execution time : ET = D3-D2 , Waiting Time : WT = D2-D1Grid Efficiency : GE = ET/(ET+WT)

Overall

Month	Short jobs	Medium jobs	Long jobs	Infinite jobs
	WT=54.05 min	EG= 30.06 % WT= 54.71 min ET= 23.52 min	WT= 54.77 min	EG= 78.81 % WT= 312.42 min ET= 1162.22 min
		EG= 5.43 % WT= 364.81 min ET= 20.96 min	WT= 115.38 min	EG= 60.25 % WT= 682.46 min ET= 1034.21 min
	WT=18.72 min	EG= 19.47 % WT= 85.03 min ET= 20.56 min	WT= 109.18 min	EG= 77.38 % WT= 212.17 min ET= 725.83 min
	WT=21.28 min	EG= 16.14 % WT= 111.94 min ET= 21.55 min	WT= 154.33 min	EG= 73.22 % WT= 263.64 min ET= 720.90 min
	WT=62.89 min	EG= 7.17 % WT= 251.74 min ET= 19.44 min	WT= 326.08 min	EG= 75.79 % WT= 336.64 min ET= 1053.97 min
_	WT=41.46 min	EG= 10.85 % WT= 170.72 min ET= 20.78 min	WT= 211.58 min	EG= 71.56 % WT= 379.74 min ET= 955.28 min

# Software Status in TERA GRID 1/2



#### TeraGrid:

-integrated, persistent computational resource.

-Deployment completed in September 2004,

-40 teraflops of computing power
-nearly 2 petabytes of storage,
-interconnections at 10-30 gigabits/sec.
(via a dedicated national network.)

Summary of Common TeraGrid Software and Services 2.0 Page generated by Inca: 06/27/05 10:24 CDT

This page offers a summary of results for critical grid, development, and cluster test test results are available by clicking on the resource name in the "Site-Resource" of

Site-Resource	Grid Development		Compute	Total Pass		
anl-ia64	Pass: 7 Fail: 12	Pass: 4 Fail: 5	Pass: 2 Fail: 1	Pass: 13 Fail: 18		
	36% passed	44% passed	66% passed	41% passed		
anl-viz	Pass: 14 Fail: 5	Pass: 9 Fail: 0	Pass: 3 Fail: 0	Pass: 26 Fail: 5		
	73% passed	100% passed	100% passed	83% passed		
caltech-ia64	Pass: 13 Fail: 6	Pass: 9 Fail: 0	Pass: 3 Fail: 0	Pass: 25 Fail: 6		
	68% passed	100% passed	100% passed	80% passed		
indiana-avidd	Pass: 18 Fail: 1	Pass: 9 Fail: 0	Pass: 3 Fail: 0	Pass: 30 Fail: 1		
	94% passed	100% passed	100% passed	96% passed		
ncsa-ia64	Pass: 19 Fail: 0	Pass: 9 Fail: 0	Pass: 3 Fail: 0	Pass: 31 Fail: O		
	100% passed	100% passed	100% passed	100% passed		
<u>psc-qs1280</u>	Pass: 8 Fail: 11	Pass: 7 Fail: 2	n/a	Pass: 15 Fail: 13		
	42% passed	77% passed		53% passed		
psc-tcs	Pass: 12 Fail: 7	Pass: 8 Fail: 1	n/a	Pass: 20 Fail: 8		
	63% passed	88% passed		71% passed		
purdue-linux	Pass: 17 Fail: 2	Pass: 9 Fail: 0	Pass: 3 Fail: 0	Pass: 29 Fail: 2		

http://tech.teragrid.org/inca-prod/cgi-bin//primaryhtmlmap.cgi?mapfile=/var/www/tech.teragrid.org/inca/TG/html/preload.state&topkey=exec

						on Affichage Favoris Outils ?
Soft	war	re S	tatus	s in <sup>-</sup>		10 globus victores est interpart provided and vide and vi
🕘 Inca Status Page - I	licrosoft Internet	Explorer				<ol> <li>globus unit test fqdn_hostname_to_jobmanager-batch</li> <li>globus unit test env_test_to_jobmanager-interactive</li> </ol>
Fichier Edition Affich	hage Favoris Outil:	s ?				19. globus unit test gridftp_to_ornl-login 20. globus unit test gridftp_to_sdsc-datastar
🔾 Précédente 🝷 🕤	- 🖹 🗈 🏠 🔎	Rechercher 🛛 📌 Favori	is 🙆 🔗 🔭 Adress	se 🕘 http://tech.teragrid.or		
1.6.2	<u>1.6.2</u> <u>1.6.2</u>	<u>1.6.2</u> <u>1.6.2</u>	<u>1.6.2</u> <u>1.6.2</u> <u>1.6</u>	<u>6.2</u> <u>1.6.2</u>	patch 1.6.2	<ol> <li>globus unit test gatekeeper_to_purdue-linux</li> <li>globus unit test gatekeeper-ping</li> </ol>
mpich-g2-gcc	[download] anl-	_ caltech- indiana	[help] - ncsa- psc-			<ol> <li>globus unit test env_test_to_jobmanager-batch</li> <li>globus unit test gridftp_to_purdue-linux</li> </ol>
version	ia64 anl-vi	z ia64 avidd	ia64 gs1280 ps	sc-tcs purdue-linux	. sp datastar	
2.4.3 <u>(2</u> subpackages)	<u>2.4.3</u> <u>2.4.3</u>	<u>2.4.3</u> <u>2.4.3</u>	<u>2.4.3</u> <u>2.4.3</u> <u>2.4</u>	<u>4.3</u> 2.4.3	<u>2.4.3</u> <u>2.4.3</u>	<ol> <li>globus unit test gridftp_to_anl</li> <li>globus unit test duroc_mpi_helloworld_to_jobmanager-batch</li> </ol>
unit tests	anl- ia64 anl-vi	z caltech- indiana ia64 avidd	i- ncsa- psc- ia64 gs1280 p:	sc-tcs purdue-linux	puraue- sasc-	<ol> <li>globus unit test gridftp_to_caltech-ia64</li> <li>globus unit test gatekeeper_to_tacc-lonestar</li> </ol>
mpicc	error passed		passed passed pa	ssed passed	passed passed	<ol> <li>globus unit test gatekeeper_to_anl</li> <li>globus unit test gridftp_to_ncsa-cobalt</li> </ol>
mpich-p4-gcc	[download] anl-	_ caltech- indiana	[help] - ncsa- psc-		purduo dec	35. globus unit test gatekeeper_to_sdsc-ia64 36. globus unit test gridftp_to_indiana-ia32
version	ia64 <sup>ani-vi</sup>	z ia64 avidd	ia64 gs1280 <sup>p:</sup>	sc-tcs purdue-linux	sp datastar	<ol> <li>globus unit test date_test_to_jobmanager-interactive</li> <li>globus unit test gatekeeper_to_psc-tcs</li> </ol>
>=1.2.5.2	anl-	<u>1.2.5.2</u> <u>1.2.6</u> caltech- indiana	1.2.5.2 error error	ror <u>1.2.6</u>		39. globus unit test gatekeeper_to_ncsa-cobalt 40. globus unit test gatekeeper_to_tacc-viz
unit tests	ia64 <sup>ani-vi</sup>	1a64 avidd	ia64 gs1280 <sup>p:</sup>	sc-tcs purdue-linux	sp datastar	<ol> <li>globus unit test gatekeeper_to_ornl-login</li> <li>globus unit test giis_query</li> </ol>
mpicc	passed passed	<u>passed</u> passed	passed error en [help]	nor passed		43. globus unit test gridftp_to_tacc-viz 44. myproxy version check
version	anl- ia64 anl-vi	z caltech- indiana ia64 avidd	i- ncsa- psc- ia64 gs1280 p	sc-tcs purdue-linux	purdue- sdsc-	45. openssh unit test openssh_to_sdsc-ia64 46. openssh unit test openssh_to_psc-gs1280
>=0.6.2 (4	4			=0.6.2 >=0.6.2		<ol> <li>openssh unit test openssh_to_purdue-linux</li> <li>openssh unit test openssh_to_tacc-viz</li> </ol>
openssh (do	errors wilload]		[help	)]		49. openssh unit test openssh_to_caltech-ia64 50. openssh version check
version	anl- ia64 anl-vi	z caltech- indiana ia64 avidd	- nesa- nec-	sc-tcs purdue-linux		51. openssh unit test openssh_to_ncsa-ia64 52. openssh unit test openssh_to_purdue-sp
any	error 3.7.1p			8.1p1 Debian-		53. openssh unit test openssh_to_indiana-ia32 54. openssh unit test openssh_to_psc-tcs
		caltoch, indiana		8.sarge.4.rcac2		55. openssh unit test openssh_to_ncsa-cobalt 56. opensch unit test opensch, to, sdee detector
unit tests	ia64 anl-vi	<sup>z</sup> ia64 avidd	ia64 gs1280 <sup>p:</sup>	sc-tcs purdue-linux		
16 tests	errors <u>1 error</u>	rs <mark>2 errors</mark> <u>1 errors</u>	<u>1 errors</u> <u>2 errors</u> <u>2 e</u>	errors <u>1 errors</u>	<u>2 errors</u> <u>2 errors</u>	7 errors 2 errors 2 errors
openssl [dov	vnload]	_ caltech- indiana	[help	)]	purdue- sdsc-	[back to top]
version	anl- ia64 anl-vi	z ia64 avidd	i- ncsa- psc- ia64 gs1280 p:	sc-tcs purdue-linux	sp datastar	sdsc- tacc- ia64 lonestar tacc-viz
0.9.* python [dow	0.9.6q 0.9.6q nload]	<u>0.9.6i 0.9.6q</u>	0.9.6 <u>m 0.9.6m 0.9</u> [help]	<u>9.6m 0.9.6m</u> 1	<u>0.9.6j 0.9.6j</u>	0.9.6q 0.9.7d
version	anl- ia64 anl-vi	z <mark>caltech- indiana</mark> ia64 avidd	- ncsa- psc- n	sc-tcs purdue-linux	purdue- sdsc- sp datastar	sdsc- tacc- tacc-viz
>=2.2 softenv [dow	2.2.1 2.2.1 (nload)	<u>2.3.3</u> <u>2.2.1</u>	<u>2.2.1 2.2.3 2.2</u> [help]	2 <u>.2 2.3.5</u> )]	<u>2.4.0 2.2.0</u>	2.2.1 2.3.4 2.3.3 [back to top]
version	anl- anl- anl-vi	z caltech- indiana	- ncsa- psc-		purdue- sdsc-	sdsc- tacc-viz
<u>e</u>						Internet

http://tech.teragrid.org/inca-prod/cgi-bin//primaryhtmlmap.cgi?mapfile=/var/www/tech.teragrid.org/inca/TG/html/preload.state&topkey=stack\_compute

# Why FT in Grid is difficult (1/2)

Grids are installed, administered and controlled by humans
 -local priority may lead to stop or freeze jobs
 -modifications and updates take times and introduce
 configuration inconsistencies
 -upgrades and modifications may introduce errors

- Heterogeneity (hardware and software, availability)
- Instability (hardware and software)

#### + Resources belong to different administration domains!

# Why FT in Grid is difficult (2/2)

# Vertical complexity and consistency

Site1

Site2

Application
Application Runtime
Grid Middleware (WS)
Services
Operating System
Networking

Application

Application Runtime

Grid Middleware (WS)

Services

**Operating System** 

Networking

Horizontal interoperability AND consistency

→ When running applications on dynamic and heterogeneous Grid, we may experience many software failures

# Research in Grid Fault Tolerance (some aspects)

#### Computing models (application runtimes):

Very few work (RPC-V, MPI: MPICH-V, MPICH-GF)

#### Infrastructure:

- Server fault tolerance (GridServices, Webservices, WSRF)
- Fault detectors (few results, Xavier'talk)
- High performance protocols (content distribution: BitTorrent)
- Resource discovery (DHT: Kadelmia)

#### FT techniques:

- Self stabilization (crash may append during stabilization)
- Consensus (impossibility result on asynchronous network)
- Majority voting (decisions may apply to a majority of nodes absent during the vote...)

#### Fault tolerance is one research topic of the CoreGrid NoE

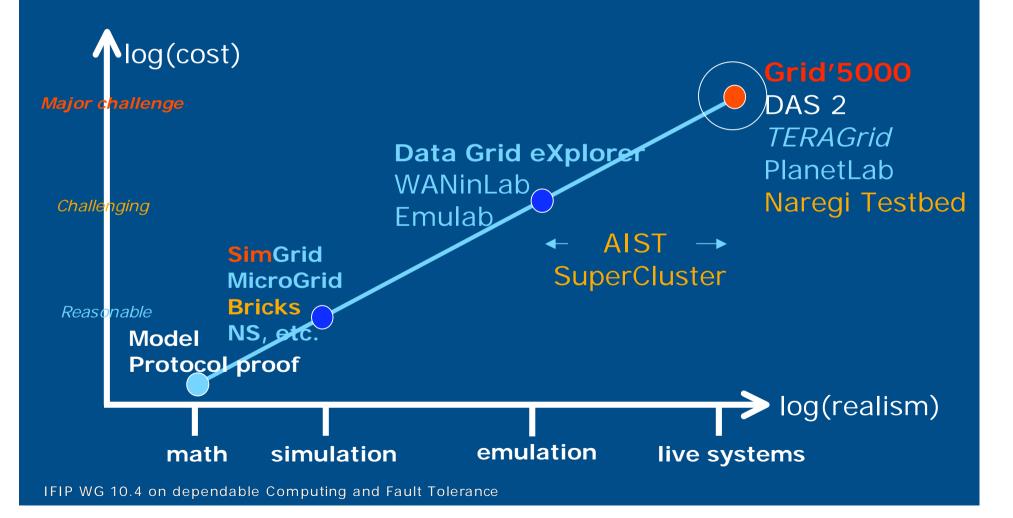
Grid still raises many issues on fault tolerance, BUT also on other topics: performance, scalability, QoS, resources usage, accounting, security, etc.

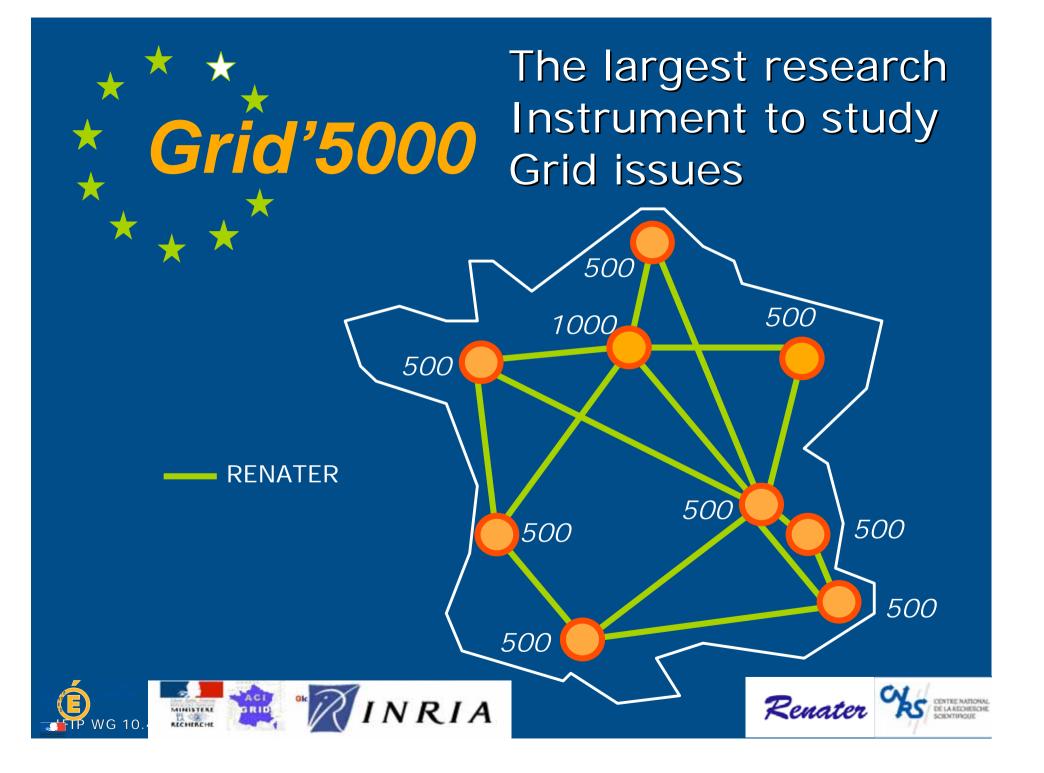
No environment or tool to test REAL Grid software at large scale

### We need Grid experimental tools

In the first ½ of 2003, the design and development of two Grid experimental platforms was decided:

→ Grid'5000 as a real life system





### Grid'5000 foundations: Collection of experiments to be done

#### Networking

- End host communication layer (interference with local communications)
- High performance long distance protocols (improved TCP)
- High Speed Network Emulation
- Middleware / OS
  - Scheduling / data distribution in Grid
  - Fault tolerance in Grid
  - Resource management
  - Grid SSI OS and Grid I/O
  - Desktop Grid/P2P systems
- Programming
  - Component programming for the Grid (Java, Corba)
  - GRID-RPC
  - GRID-MPI
  - Code Coupling
- Applications
  - Multi-parametric applications (Climate modeling/Functional Genomic)
  - Large scale experimentation of distributed applications (Electromagnetism, multi-material fluid mechanics, parallel optimization algorithms, CFD, astrophysics

IFIP WG.10. Medical index Collaborating tools in virtual 3D environment

### Grid'5000 goal: Experimenting fault tolerance and many other topics on all layers of the Grid software stack

Application

**Programming Environments** 

**Application Runtime** 

Grid Middleware

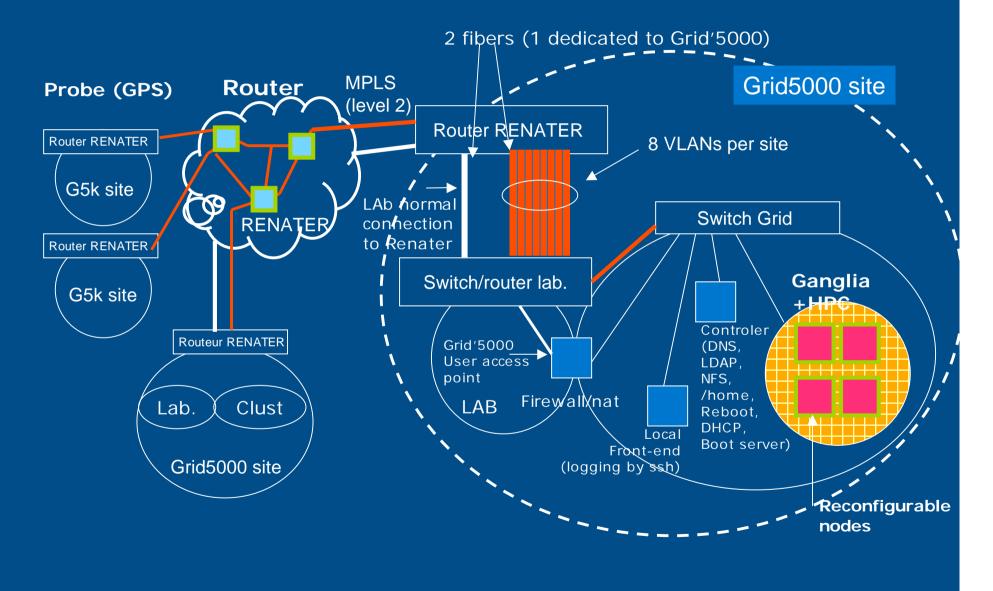
**Operating System** 

Networking

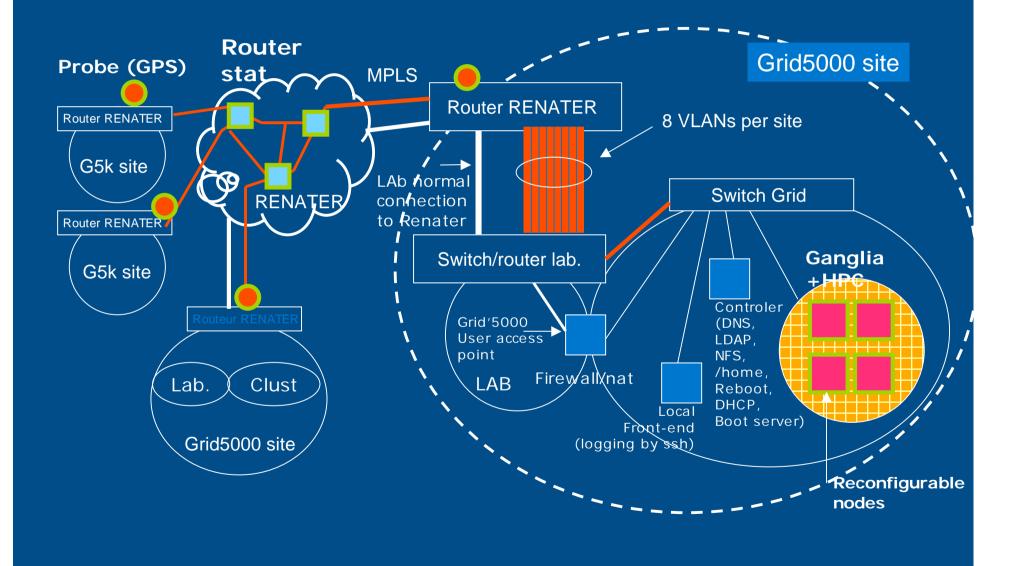


A highly reconfigurable, controllable and monitorable experimental platform

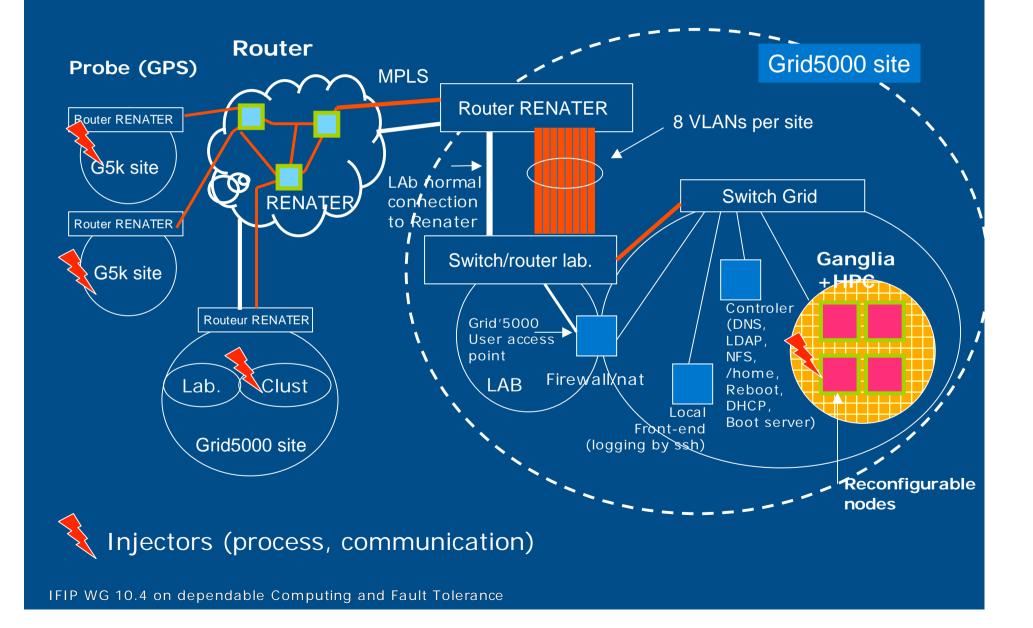
### **Confinement / isolation**



### **Observation & Monitoring**

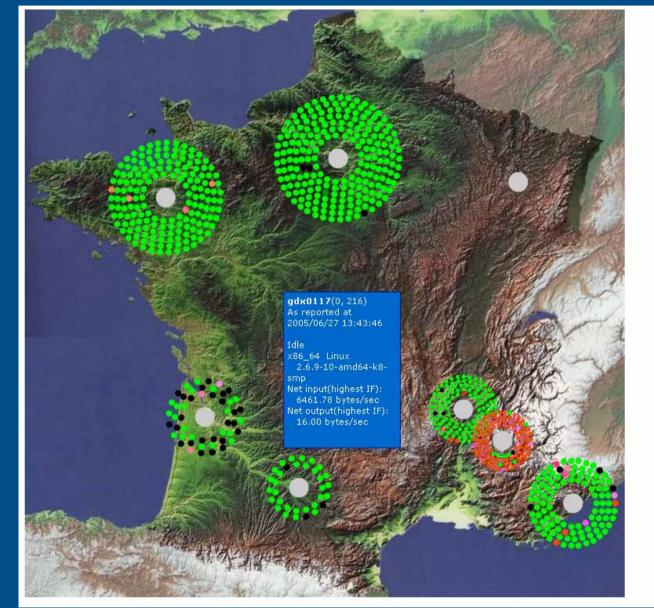


### Workload/Traffic & Fault injection



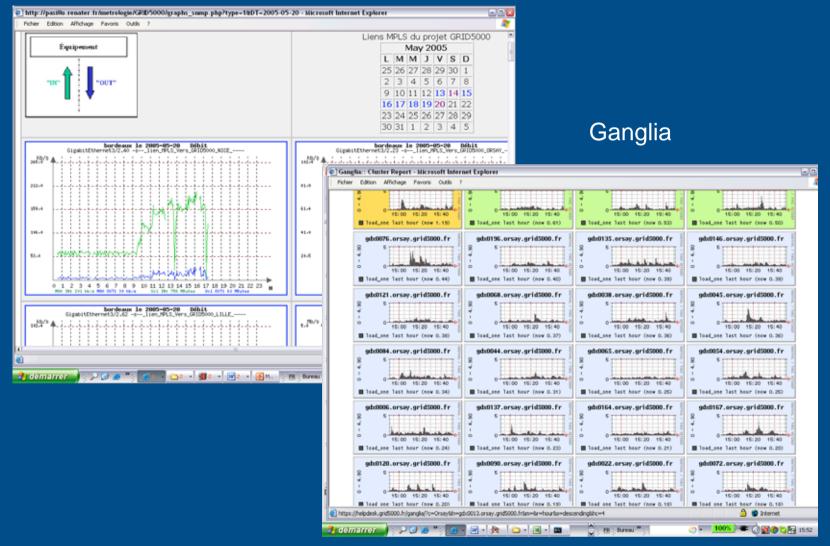


### Grid'5000 Global Observer



### Grid'5000 Monitoring tools

#### Network traffic



🔄 Grid5000 status - Microsoft Internet Explorer

# Grid'5000 Reservation Cluster Name and reconfiguration

Cluster Name	gdx	idpot	lyon	paraci	parasol	sophia	tartopom	toulouse	all
Site	orsay	grenoble	lyon	rennes	rennes	sophia	rennes	toulouse	
Турө	opteron	xeon	opteron	xeon	opteron	opteron	g5	opteron	
Free Nodes	164	6	62	1	23	65	32	23	376
Busy Nodes	48	13	0	0	0	0	0	0	61
All Nodes	216	22	62	64	64	105	32	31	596

Free

ine. Ter

ree Tee ree .....

ree Free Free ree ree

Abst Ter 149 Free 1120

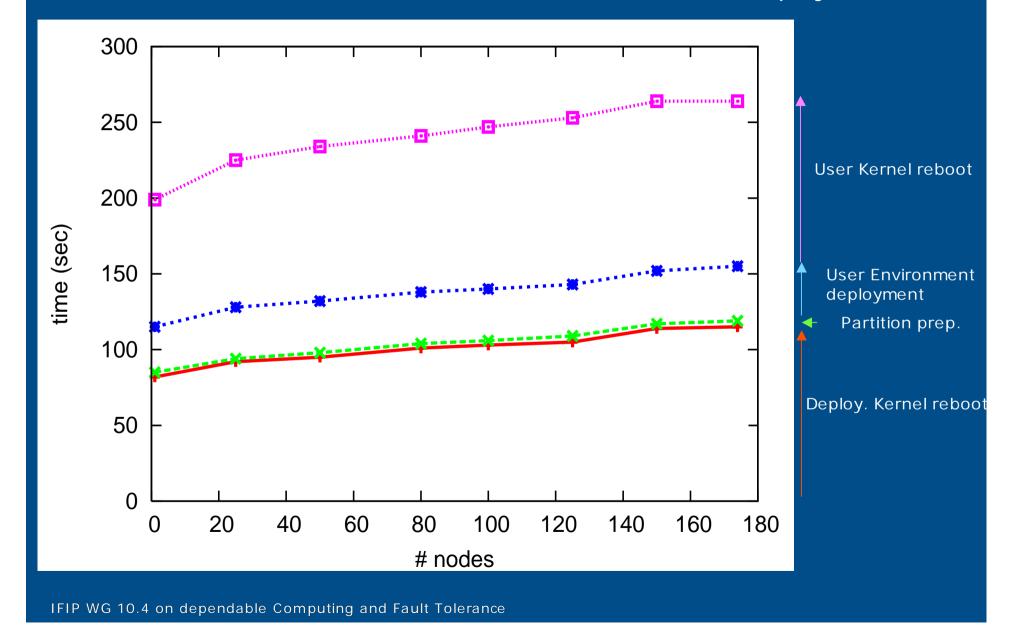
148 rer ree ree

Free Free

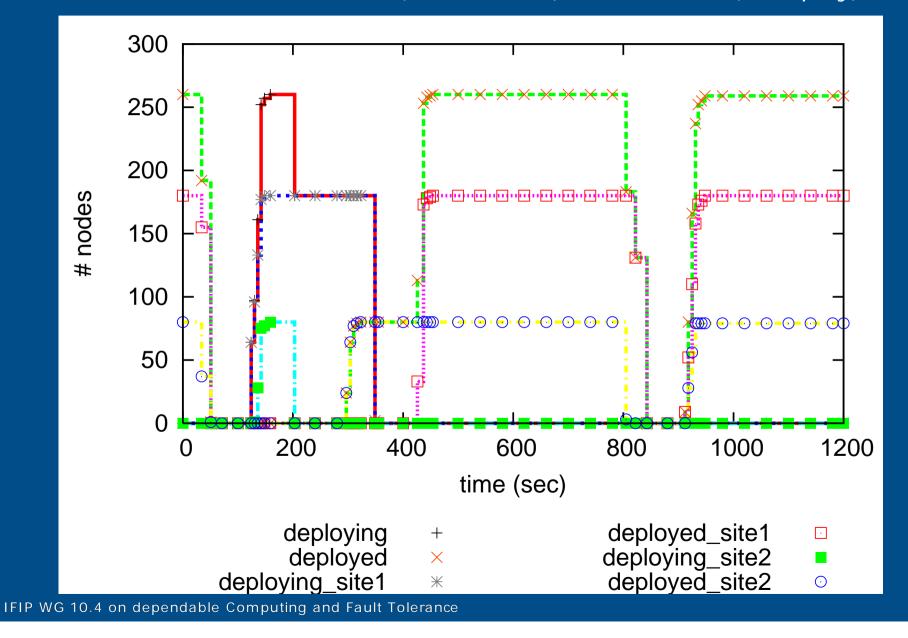
🚊 🙂 Internet 〇 - 100% - @ 図のる場局 15:46

GanttChart - Microsoft Internet Explorer     Fichier Edition Affichage Favoris Outils ?		orsay: gd	X					
③ Précédente       ▼  <		- E	ree Free 1148 ree Free Free ree 1148 Free	1148 Free	148 Free Fre	e Free Fre	e Free F	ree
			ree 1148 Free	1148 Free	ree Free Fre	e Free Fre	e 1148	ree
	1145		ree Free Free ree Free Free	1148 1148 1148 Free	ree Free Fre Free Free Fre	e Free Fre e Free Fre	e Free 1 e Free F	148 iree
84x0152.0rs30.8r165000.Fr 84x0152.0rs30.8r165000.Fr 84x0152.0rs30.8r165000.Fr	1145	Fr	ree Free Free Tree Free Free	Free Free F	ree Free Fre	e Free Fre e 1148 Fre	e Free 1	148 ree
	1145		ree 1148 Free		ree Free Fre	e 1148 Fre	e Free 1	148
	1145	1		1148 1120			e Free F	ree ree
84x0166ors89.87168000.Fr 84x0166ors89.87168000.Fr 84x0166ors89.87168000.Fr		E.	ree Free Free	Free Free	ree Free Fre	e Free Fre	e 1120 A	bst
		Terniné	2 🔎 🕑 🍎 " 2		🎘   🗀 - 🛛 🗷		2 60 2 6.4	
		Juliana de la coma de					IS THE S DOM	
8000242,007384,871182000,171 8000248,007281,871182000,171 8000248,007281,871182000,171 80002488,0072884,871182000,171	1145		1148					
	1120		1148		-			
gdX0186, -07384, gr 1d5000 + fr gXX0187, -07384, gr 1d5000 + fr gXX0188, -07584, gr 1d5000 + fr gXX0188, -07584, gr 1d5000 + fr gXX0188, -07584, gr 1d5000 + fr	1120							
	1145							
			1148					
	1145							
8400204.or536(.8r1d5000.+r 8400206.or534.8r1d5000.+r 8400206.or534.8r1d5000.+r	1145							
	1120 1103 1120		Absent					
( \$300213:07-307:87:185000.Fr 8dx0214.orsad.8rid5000.Fr								>
ê					🔒 🥝	Internet		

### Grid'5000 Reconfiguration time Time to reboot 1 cluster of Grid'5000 with Kadeploy



### **Grid'5000 Reconfiguration time** Time to reboot 2 clusters (Paris + Nice) of Grid'5000 (Kadeploy)



## Grid'5000 Fault Generator: Fail

#### Objectives

- •Probabilistic and deterministic (reproducible) fault injection.
- •Expressiveness of scenarios.
- •No code modification.

IMPACT OF FAULTS

•Scalable.

#### Concepts

100

80

60

20

0.2

0.4

Probability of fault appearance

0.6

execution time (in seconds)

A dedicated language for fault scenario specification (FAIL: FAult Injection Language).
Fine control of the code execution (through a debugger)

> Added 2 Completed 1

> > 0.9

#### Daemon ADV2

01001 10011

01001

10011

nd Fault Toleran fail-exec.run

1

```
{
time_g timer = 5;
node 1 :
    always int rand = FAIL_RANDOM (1,10);
    timer && rand < 2 -> halt, goto 2;
node 2 :
    always int rand = FAIL_RANDOM (1,10);
    timer && rand > 7 -> restart, goto 3;
node 3 :
```

01001

10011

01001

10011

01001

10011

01001

10011

### Summary:

- Grid still raises many issues about fault tolerance
- Grid'5000 will offer a large scale infrastructure to study some of these issues (operational in September 2005)

**Grid'5000** 

• Grid'5000 will be opened to international collaborations