Enabling Grids for E-science in Europe

Athens, March 20th-21st, 2006

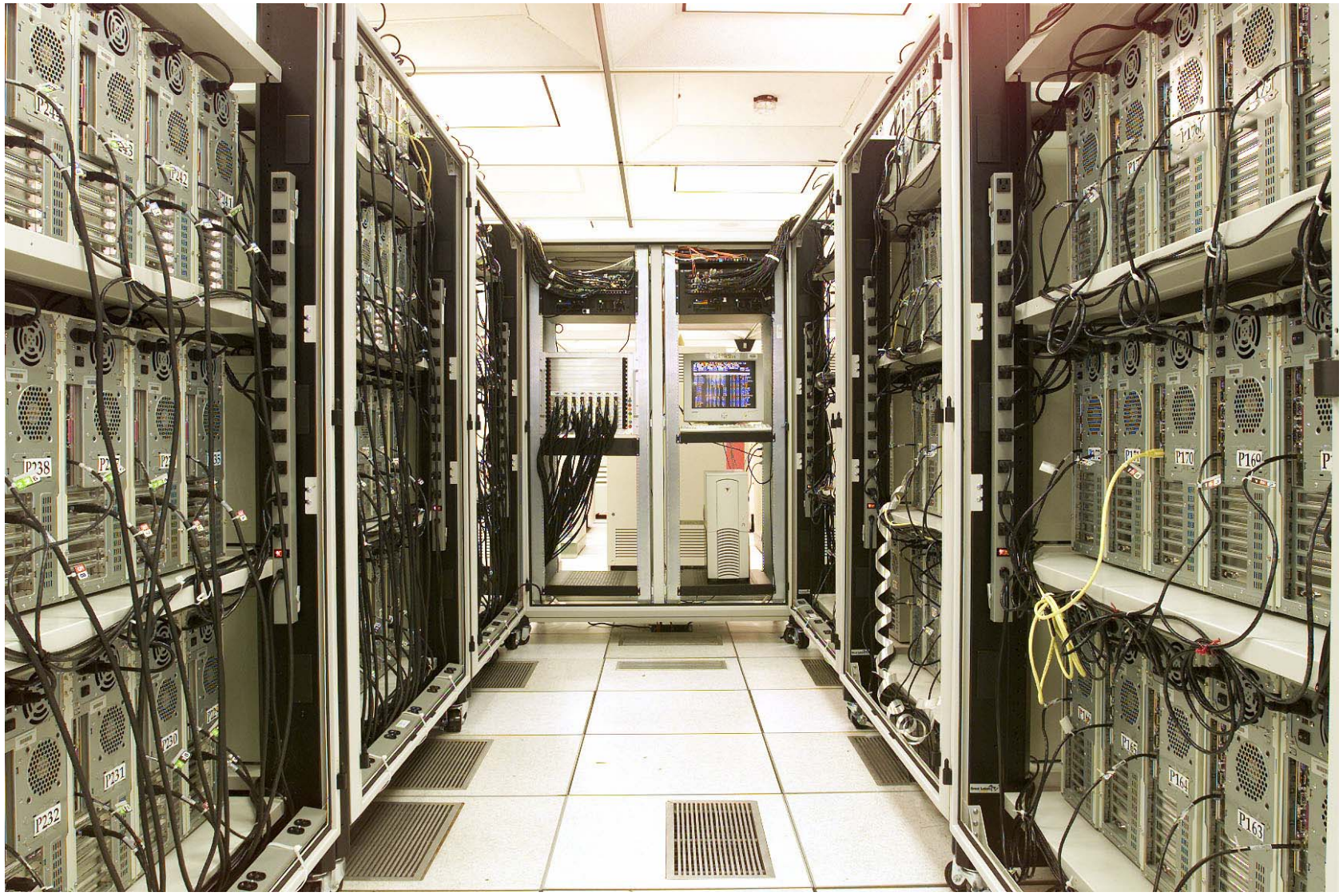
LCG/EGEE Grids & System Administration

Fotis Georgatos
Trainer, GRNET

EGEE is a project funded by the European Union

Watch out what you wish for!

Enabling Grids for E-science in Europe

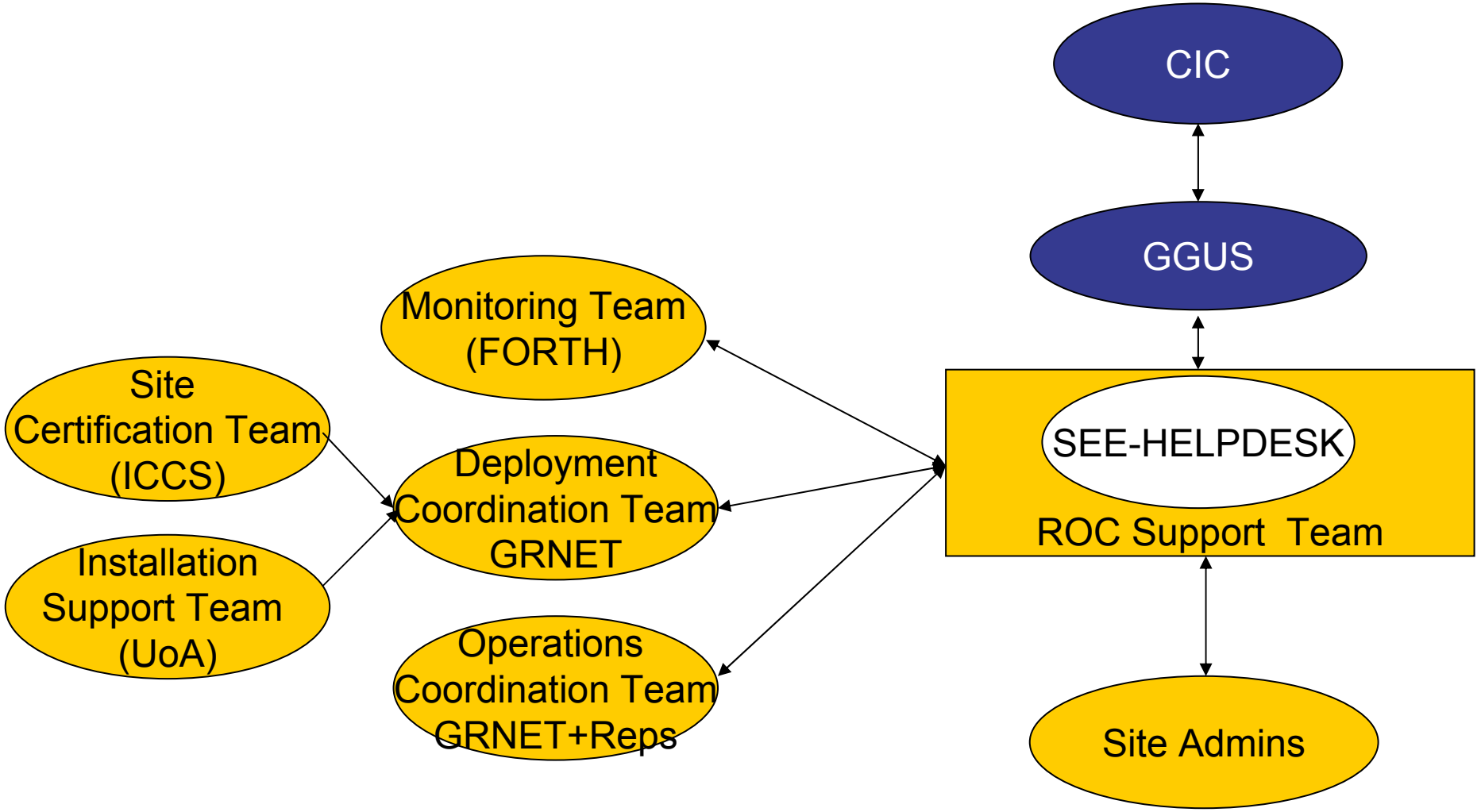


Scope

Enabling Grids for E-science in Europe

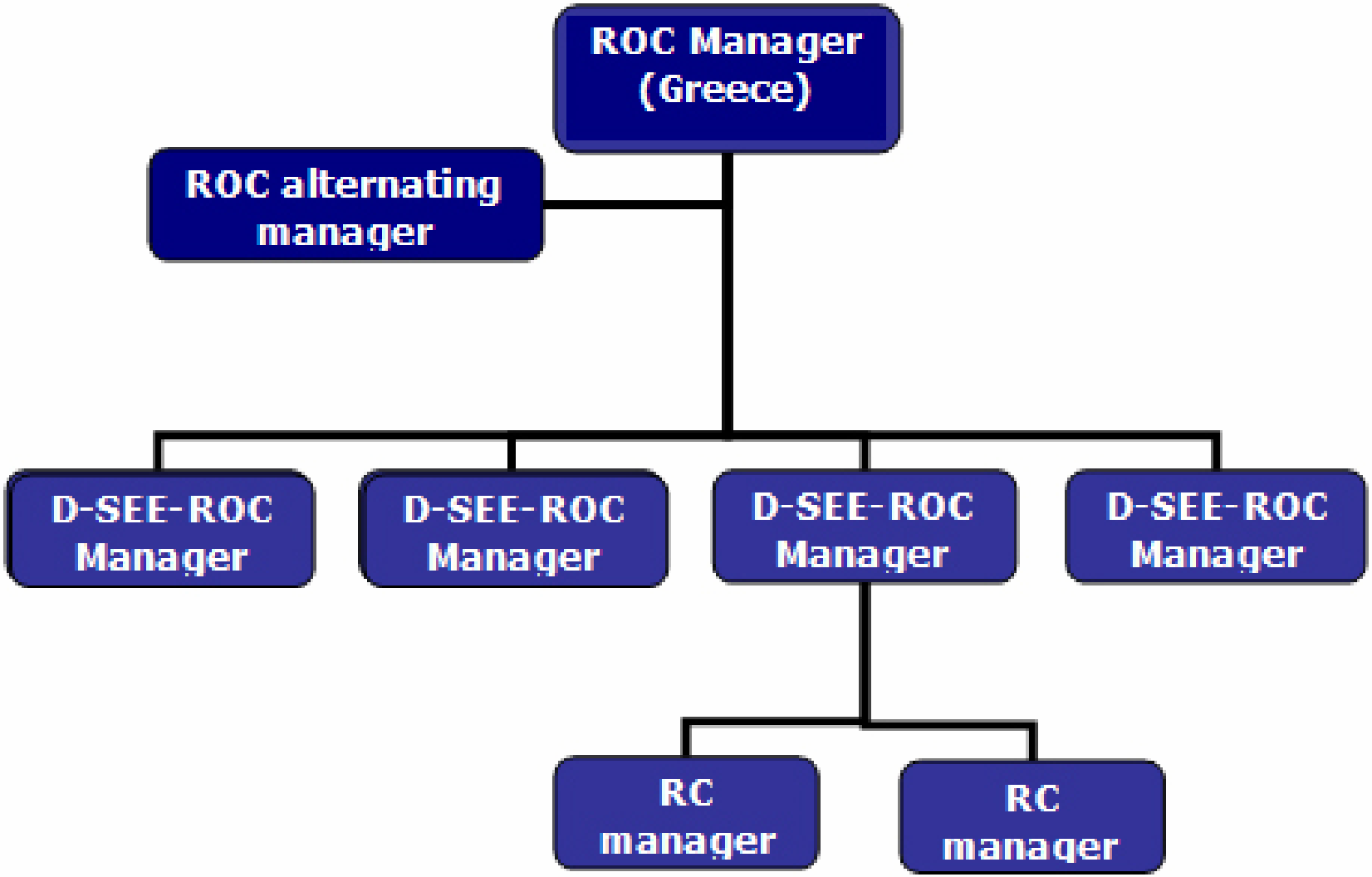
- Explain runtime operations and how the workflow will evolve
- Explain new site deployment and upgrade procedures
- Transfer experience from the HellasGrid I & II phases
- What is GGUS, the egee-see helpdesk and other friends
- Runtime operations include
  - Site problem solving (ROC Support Team)
  - Minor upgrades (Operations Coordination Team)
  - Configuration changes (Operations Coordination Team)
  - Security incidents (Security Coordination Team)

SEE Support Model



The ROC structure within SEE

Enabling Grids for E-science in Europe



Site deployment

Enabling Grids for E-science in Europe

- Involves:
  1. Site registration, done in coordination with the Deployment Coordination Team (DCT) – 'candidate site'
  2. Site installation, done with the guidance and assistance of the Installation Support Team (IST) and the DCT - 'uncertified site'
  3. Site certification, in coordination with the CIC-on-duty and the Site Certification Team (SCT) - 'certified site'

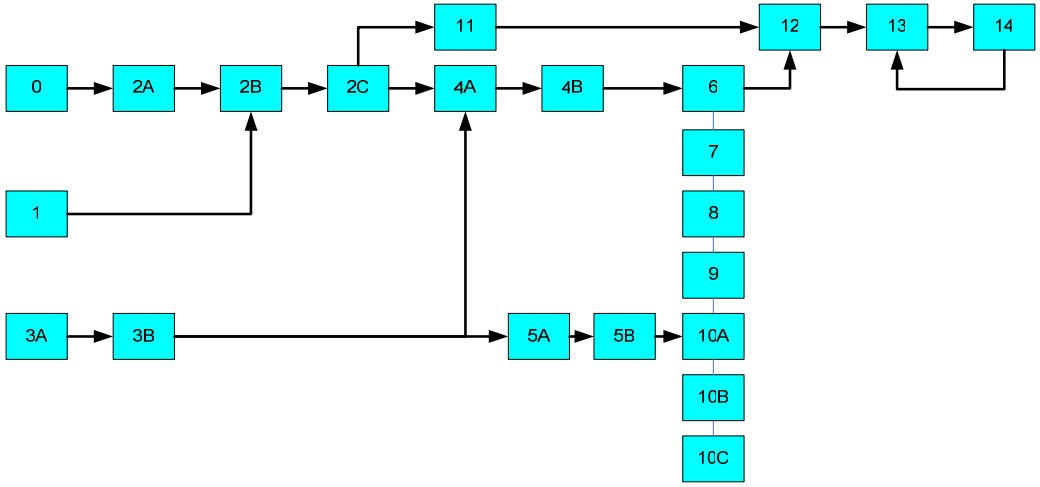
Site induction procedure

Enabling Grids for E-science in Europe

- Follow the 10-step registration procedure, as described here: http://www.egee-see.org/Site_registration.php
- Perform a site installation according to the instructions here: http://lcg.web.cern.ch/LCG/Sites/releases.html
- Follow the 5-step certification sequence: http://www.egee-see.org/Site_certification.php
- Successful Site Functional Tests for five consecutive days, imply the site is stable enough to be considered "Production"
- DCT changes site status in the GOC database to *'certified'*
- From now on the site will submit the relevant weekly reports, also known as RC reports, Friday – Monday 11:00AM(GMT): https://cic.in2p3.fr/index.php?id=rc&subid=rc_report&js_status=2

The complexities of a full site setup

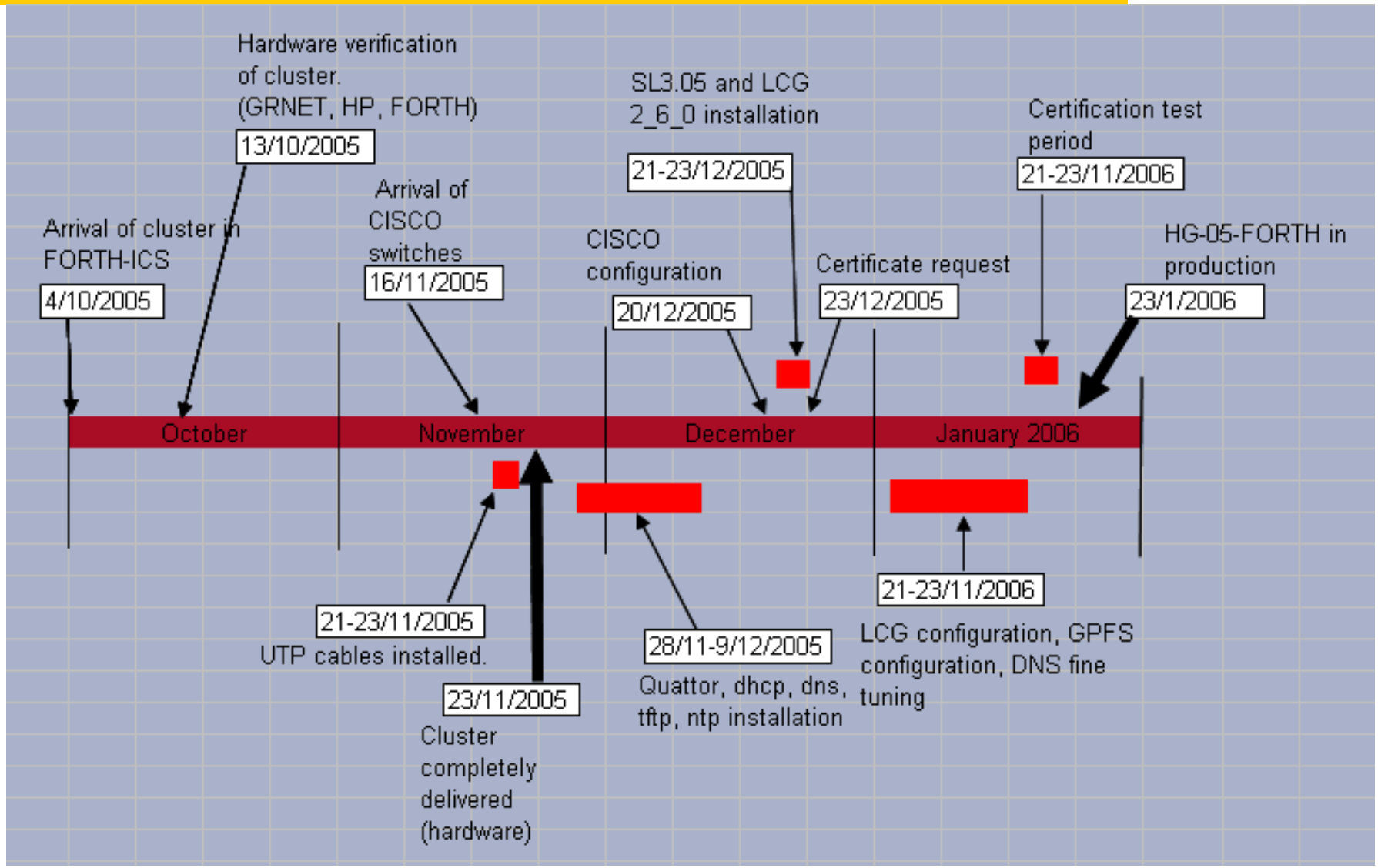
Enabling Grids for E-science in Europe



0) UI account setup (Isabella or local)
1) An academic sends a fax to Hellasgrid CA, mentioning sysadmin's name
2A) User certificate is created by sysadmin(s), for himself
2B) Accept user certificate against CA, within a week's time!
2C) convert user certificate to pkcs12, add to browser and mail client
3A) Decide for IP, NTP and DNS configuration (addresses, names, servers)
3B) Verify NTP, DNS (forward and reverse) with host, www.dnsstuff.com etc
4A) Host certificates are created by sysadmin(s), (for CE, SE, MON ...)
4B) Accept host certificates against CA, within a week's time!
5A) Istall SL30x to your systems
5B) Setup Java, NTP, lcg-yaim
6) Configure CE
7) Configure SE
8) Configure WNs
9) Configure MON 9B) & send its IP address to RGMA DB
10) Configure UI and/or 10B) BDII 10C) RB
11) Add site to GOCdb & inform ROC manager, update GOCdb
12) Ask Isabella's team for SFTs
13) Identify any erroneous entries with SFTs
14) Perform corrective actions for errors found in 13)

Timeline: a real case of installation

Enabling Grids for E-science in Europe



Site registration procedure

Enabling Grids for E-science in Europe

- Resource admin at candidate site obtains a certificate from Certification Authority (CA) eg. http://ca.hellasgrid.gr or http://ca.egee-see.org
- Site manager contacts the ROC Deployment Coordination Team (DCT)

  Providing:

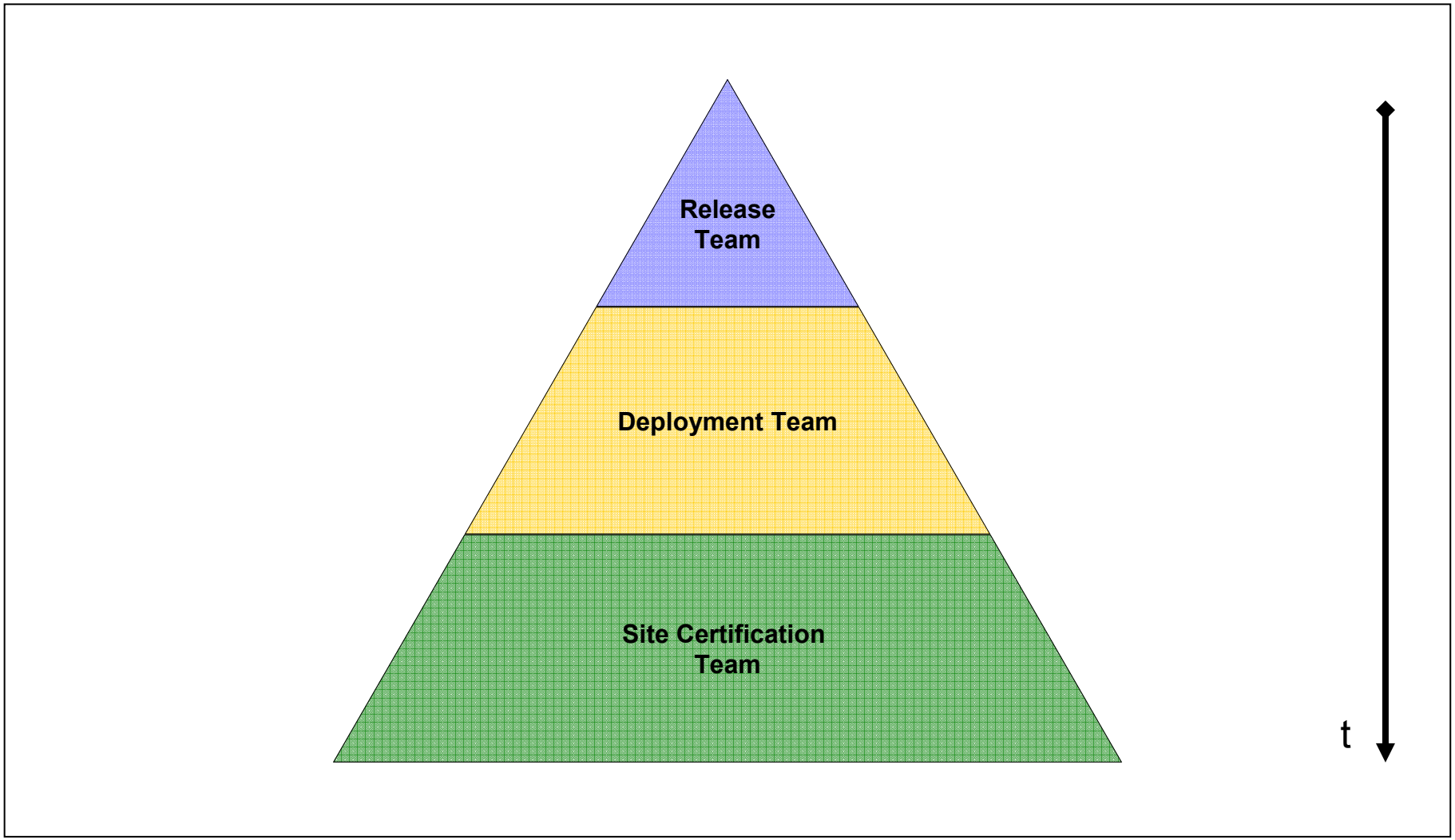
  1. Statement of acceptance of policy documents. http://www.egee-see.org/aup
  2. Site contact information
  3. HW details in the Infrastructure.xls
  4. Statement of agreement with LCG/EGEE Security Incident Response procedures https://edms.cern.ch/file/428035/LAST_RELEASED/LCG_Incident_Response.pdf
- DCT validates the information and creates the new site's record in the GOC database; site status is '***candidate***'.
- Site to register with the helpdesk https://helpdesk.egee-see.org/
- Resource Admin at the site enters the remaining information in the GOC database, and then requests validation by the DCT
- The DCT validates the information and changes the site status to '***uncertified***'

Site installation

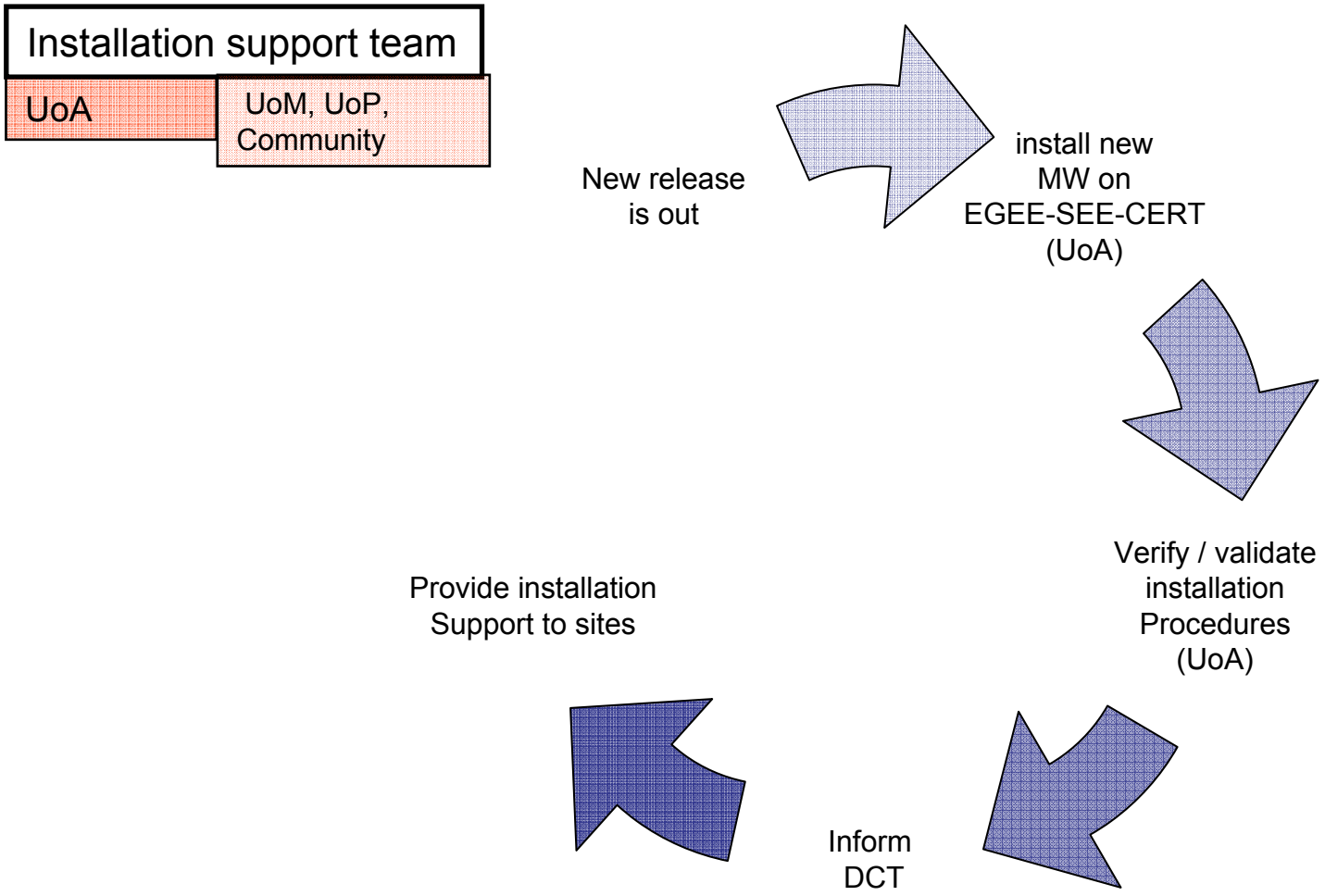
Enabling Grids for E-science in Europe

- Done with the guidance and assistance of the IST and the DCT
  - Installation support through egee-see helpdesk
- Using the Installation Guidelines at
  - http://lcg.web.cern.ch/LCG/Sites/releases.html
- Supported OS: SL 305, SLC306, RHEL3 or similar

Site deployment/upgrade workflow



Deployment/upgrade workflow (Release Team)



Deployment/upgrade workflow (Deployment Team)

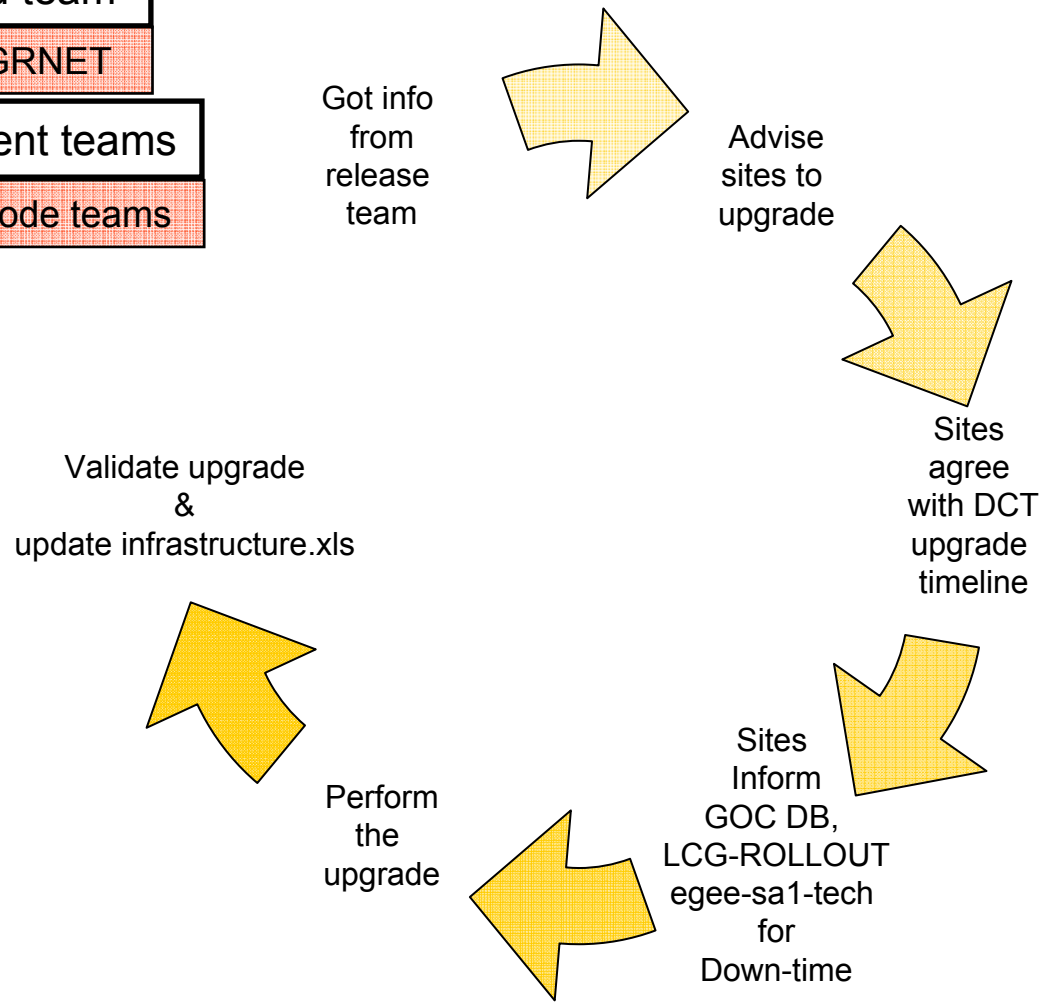
Enabling Grids for E-science in Europe

Deployment coord team

GRNET

Per-node deployment teams

Node teams



Deployment/upgrade workflow (Site Certification Team)

Enabling Grids for E-science in Europe

Deployment coord team

GRNET

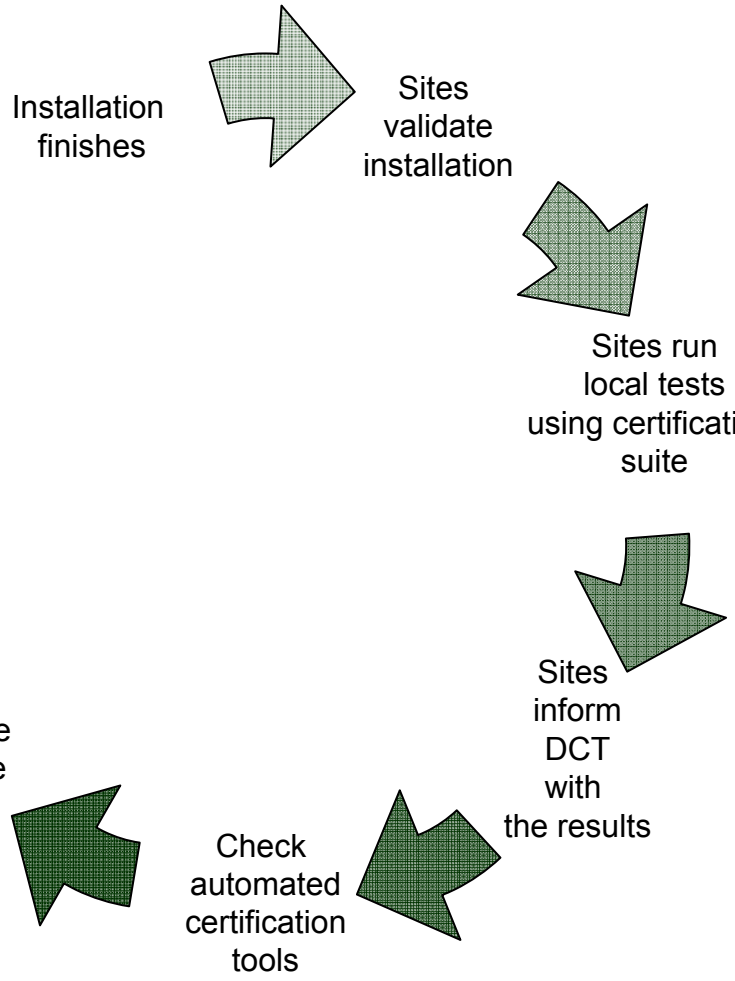
Per-node deployment teams

Node teams

Site certification team

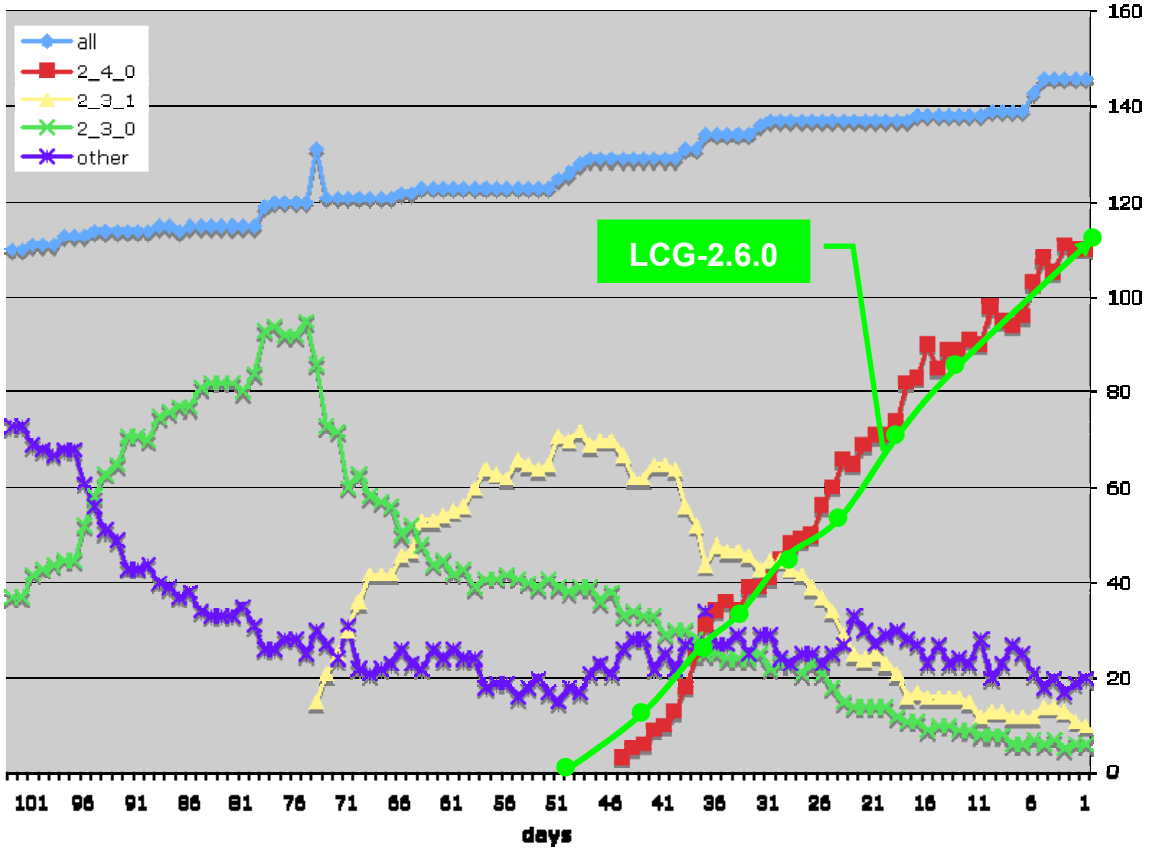
ICCS, Demokritos

GRNET



Upgrade cases of the LCG Testbed

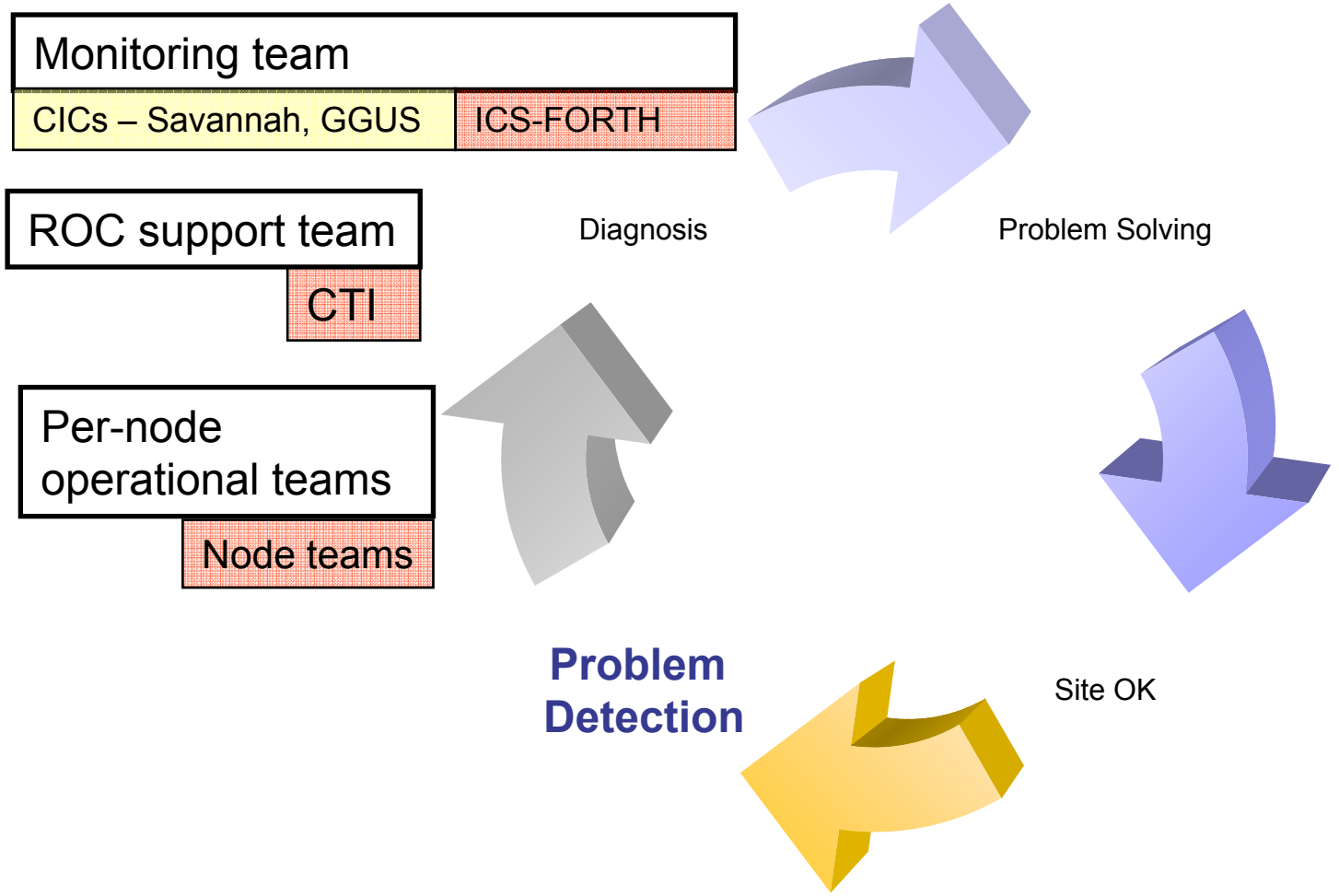
Enabling Grids for E-science in Europe



- **Time to upgrade ~constant (~2.5 sites/day)**
- **Takes a long time, O(#sites), to upgrade entire infrastructure**
- **Better now than it was – site functional tests and operational oversight**
- **Need to move away from the need to do full upgrades more than 1-2 times / year**
  - But need to be able to deploy updates, new tools, security patches, etc.

Vicious circle of problem solving

Enabling Grids for E-science in Europe



Problem detection

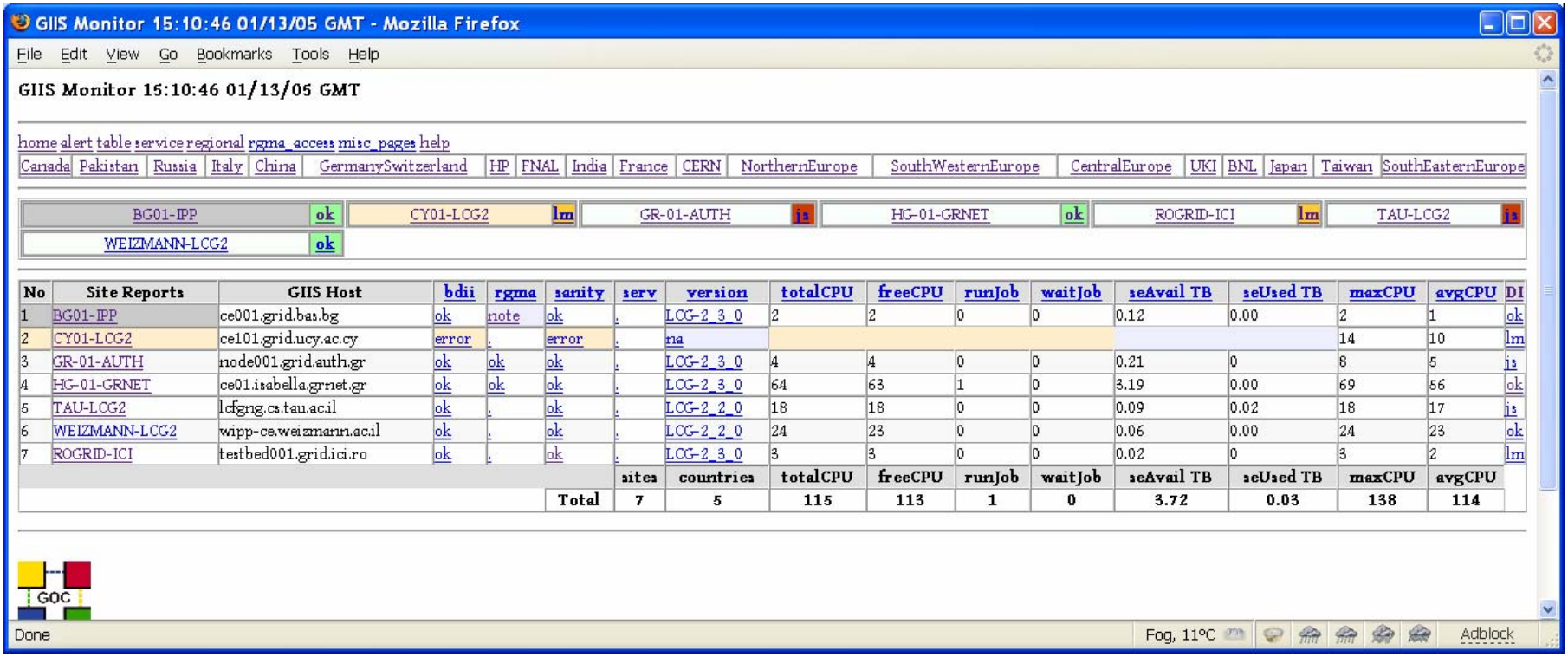
Enabling Grids for E-science in Europe

- Done mostly by CIC-on-duty, but also ROC
- 1. Problem spotted using:
  - **GIIS Monitor**: http://goc.grid.sinica.edu.tw/gstat/
  - **Site Functional Tests**: https://lcg-sft.cern.ch/sft/lastreport.cgi
  - **Grid ICE**: http://gridice2.cnaf.infn.it:50080/gridice/site/site.php
  - **GOC Database**: https://goc.grid-support.ac.uk/gridsite/db
  - **GOC Job Monitor**: http://goc02.grid-support.ac.uk/googlemaps/lcg.html
  - **GOC Live Job Monitor**: http://gridportal.hep.ph.ic.ac.uk/rtm/
- 2. CICs open a Ticket at Savannah (https://savannah.cern.ch/projects/lcg2sites) and sends an e-mail to site Administrators and the ROC **support team** grid-support@egee-see.org.
- Site admins should look at:
  - A) the GIIS Monitor and
  - B) the Site Functional Tests and start working on a problem ASAP

Monitoring tools: GIIS Monitor

Enabling Grids for E-science in Europe

- GIIS Monitor (http://goc.grid.sinica.edu.tw/gstat/)



GIIS Monitor 15:10:46 01/13/05 GMT - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

GIIS Monitor 15:10:46 01/13/05 GMT

home alert table service regional rgma_access misc_pages help

Canada | Pakistan | Russia | Italy | China | GermanySwitzerland | HP | FNAL | India | France | CERN | NorthernEurope | SouthWesternEurope | CentralEurope | UKI | BNL | Japan | Taiwan | SouthEasternEurope

BG01-IPP ok	CY01-LCG2 lm	GR-01-AUTH js	HG-01-GRNET ok	ROGRID-ICI lm	TAU-LCG2 js
WEIZMANN-LCG2 ok					

No	Site Reports	GIIS Host	bdii	rgma	sanity	serv	version	totalCPU	freeCPU	runJob	waitJob	seAvail TB	seUsed TB	maxCPU	avgCPU	DI
1	BG01-IPP	ce001.grid.bas.bg	ok	note	ok	.	LCG-2_3_0	2	2	0	0	0.12	0.00	2	1	ok
2	CY01-LCG2	ce101.grid.ucy.ac.cy	error	.	error	.	na							14	10	lm
3	GR-01-AUTH	node001.grid.auth.gr	ok	ok	ok	.	LCG-2_3_0	4	4	0	0	0.21	0	8	5	js
4	HG-01-GRNET	ce01.isabella.grnet.gr	ok	ok	ok	.	LCG-2_3_0	64	63	1	0	3.19	0.00	69	56	ok
5	TAU-LCG2	lcfgng.cs.tau.ac.il	ok	.	ok	.	LCG-2_2_0	18	18	0	0	0.09	0.02	18	17	js
6	WEIZMANN-LCG2	wipp-ce.weizmann.ac.il	ok	.	ok	.	LCG-2_2_0	24	23	0	0	0.06	0.00	24	23	ok
7	ROGRID-ICI	testbed001.grid.ici.ro	ok	.	ok	.	LCG-2_3_0	3	3	0	0	0.02	0	3	2	lm
						sites	countries	totalCPU	freeCPU	runJob	waitJob	seAvail TB	seUsed TB	maxCPU	avgCPU	
					Total	7	5	115	113	1	0	3.72	0.03	138	114	

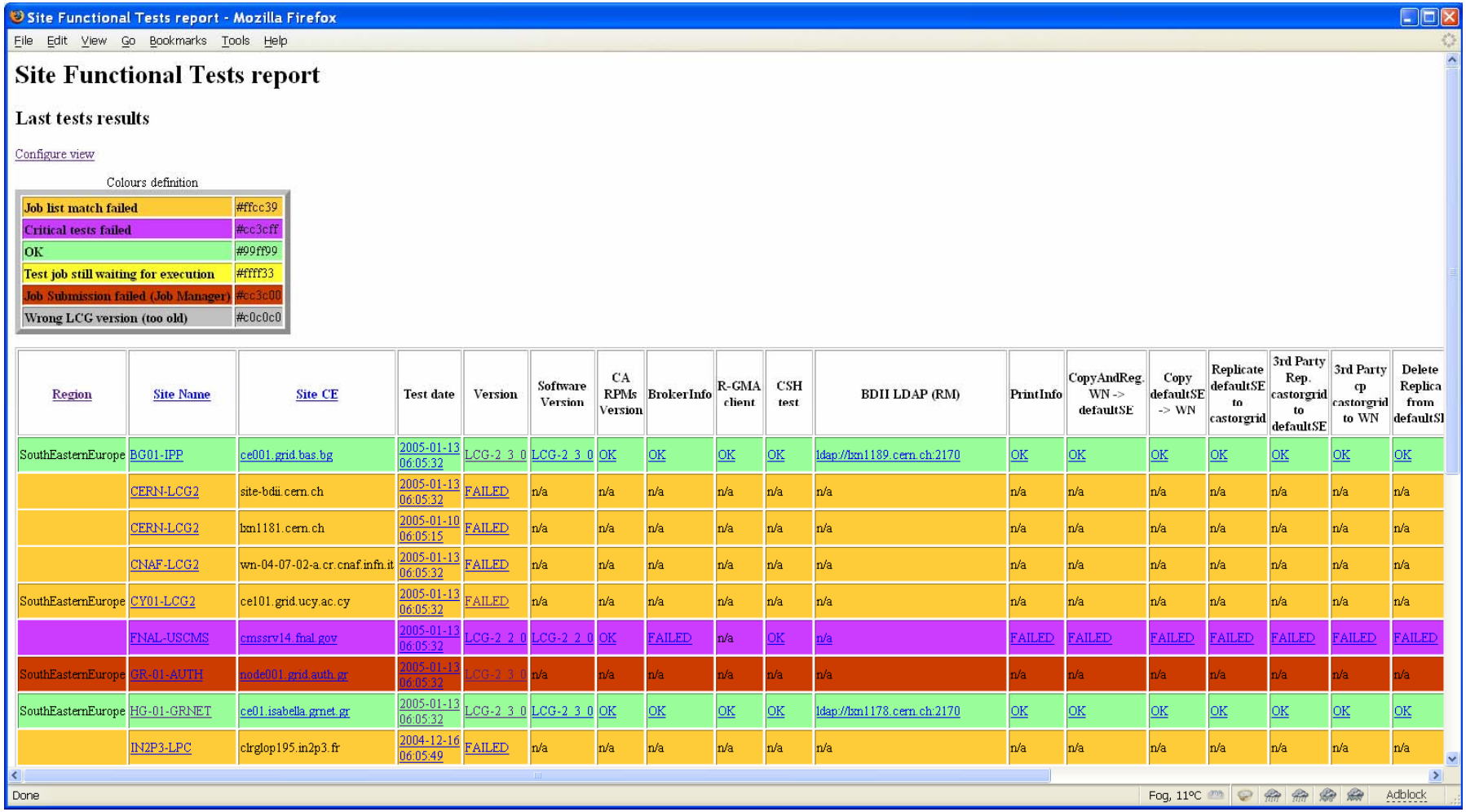
GOC

Done Fog, 11°C Adblock

Monitoring tools: SFTs

Enabling Grids for E-science in Europe

- Site Functional Tests (https://lcg-sft.cern.ch/sft/lastreport.cgi)



Site Functional Tests report - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

Site Functional Tests report

Last tests results

Configure view

Colours definition

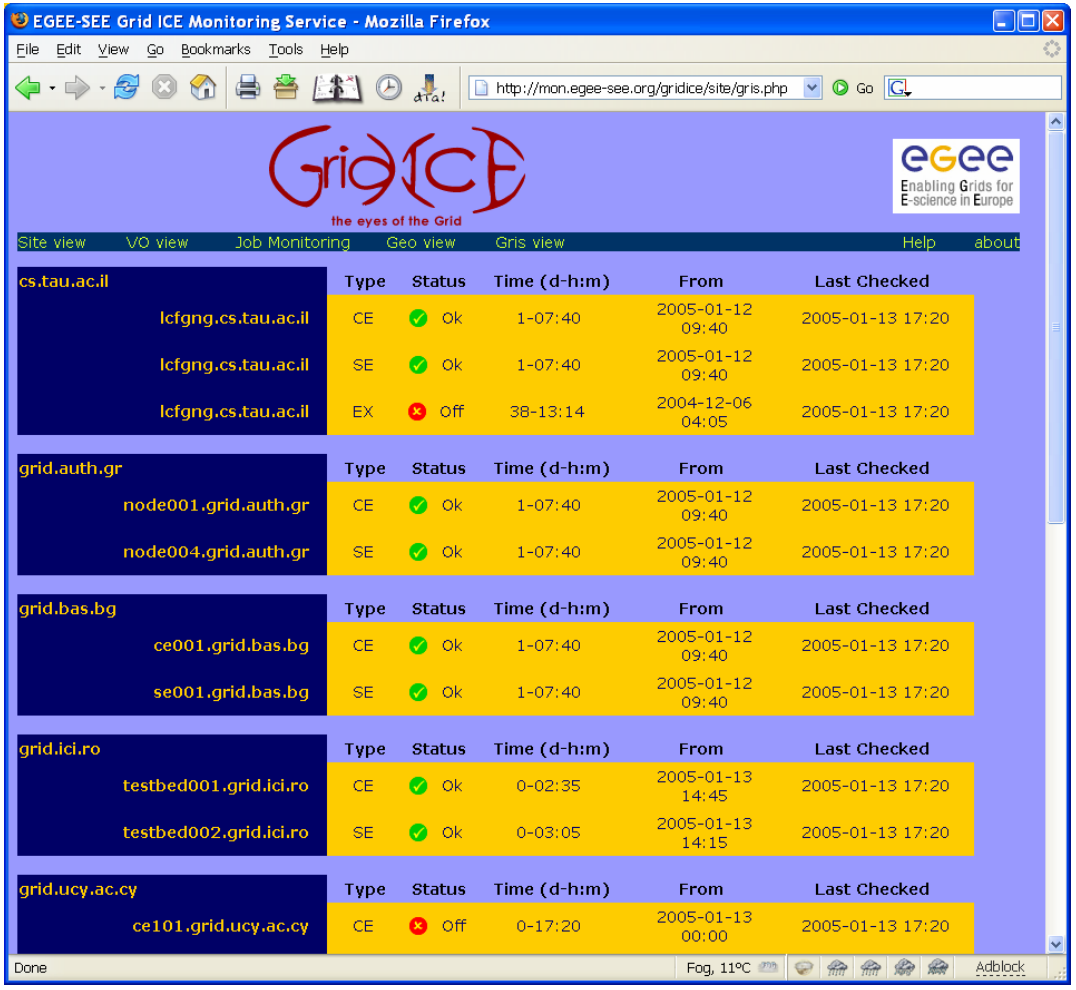
Job list match failed	#ffcc39
Critical tests failed	#cc3cff
OK	#99ff99
Test job still waiting for execution	#ffff33
Job Submission failed (Job Manager)	#cc3c00
Wrong LCG version (too old)	#c0c0c0

Region	Site Name	Site CE	Test date	Version	Software Version	CA RPMs Version	BrokerInfo	R-GMA client	CSH test	BDII LDAP (RM)	PrintInfo	CopyAndReg. WN -> defaultSE	Copy defaultSE -> WN	Replicate defaultSE to castorgrid	3rd Party Rep. castorgrid to defaultSE	3rd Party cp castorgrid to WN	Delete Replica from defaultSE
SouthEasternEurope	BG01-IPP	ce001.grid.bas.bg	2005-01-13 06:05:32	LCG-2_3_0	LCG-2_3_0	OK	OK	OK	OK	ldap://lxn1189.cern.ch:2170	OK	OK	OK	OK	OK	OK	OK
	CERN-LCG2	site-bdii.cern.ch	2005-01-13 06:05:32	FAILED	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	CERN-LCG2	lxn1181.cern.ch	2005-01-10 06:05:15	FAILED	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	CNAF-LCG2	wn-04-07-02-a.cr.cnaf.infn.it	2005-01-13 06:05:32	FAILED	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
SouthEasternEurope	CY01-LCG2	ce101.grid.ucy.ac.cy	2005-01-13 06:05:32	FAILED	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	FNAL-USCMS	cmssrv14.fnal.gov	2005-01-13 06:05:32	LCG-2_2_0	LCG-2_2_0	OK	FAILED	n/a	OK	n/a	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED	FAILED
SouthEasternEurope	GR-01-AUTH	node001.grid.auth.gr	2005-01-13 06:05:32	LCG-2_3_0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
SouthEasternEurope	HG-01-GRNET	ce01.isabella.grnet.gr	2005-01-13 06:05:32	LCG-2_3_0	LCG-2_3_0	OK	OK	OK	OK	ldap://lxn1178.cern.ch:2170	OK	OK	OK	OK	OK	OK	OK
	IN2P3-LPC	clrglop195.in2p3.fr	2004-12-16 06:05:49	FAILED	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Done Fog, 11°C Adblock

Monitoring tools: GridICE

- Grid ICE (http://gridice2.cnaf.infn.it:50080/gridice/site/site.php)



Diagnosis

Enabling Grids for E-science in Europe

- Read the description of the task / problem
  - mail received from the CIC-on-duty
- Report in Savannah that you are dealing with the task
- Look (yes!) at the “Sites functional tests” for more details
- Get assistance / guidance from
  - Knowledge databases with information about the typical problems
    - http://goc.grid.sinica.edu.tw/gocwiki/SiteProblemsFollowUpFaq
    - http://grid-it.cnaf.infn.it/index.php?knowledgebase
    - http://www.gridpp.ac.uk/tb-support/faq/index.html
  - The pools of expertise that reside in
    - **LCG-ROLLOUT@LISTSERV.RL.AC.UK**
    - egee-sa1-tech@grnet.gr

Problem solving

Enabling Grids for E-science in Europe

- Try to **reproduce** the problem and **identify** its cause
- Try to **solve** the problem
- If needed request assistance from egee-sa1-tech@grnet.gr and or the CIC-on-duty
- If problem is solved
  - Report to Savannah that its is solved and explain briefly what caused it
  - CIC-on-duty or the ROC support team updates the ticket to “Site OK”
  - Your site will be in quarantine for 3 working days
- Else
  - Report to **Savannah** and grid-support@egee-see.org the reason why the problem is not solved.
  - Keep on trying … ☺

Where to get information from

Enabling Grids for E-science in Europe

- ICCS team Grid Certification Guide (you should be familiar!)
  http://www.egee-see.org/content/modules/downloads/Certification_v2.pdf
- HOW TO TEST AN LCG2 SITE (LCG2-Site-Testing.pdf)
  http://grid-deployment.web.cern.ch/grid-deployment/documentation/LCG2-Site-Testing.pdf
- DESY's Test Guide (find examples and debug commands)
  http://grid.desy.de/tests/
- GridPP: FAQ for LCG Site Administrators
  http://www.gridpp.ac.uk/tb-support/faq/
- GOCwiki: Grid Administration FAQs and Troubleshooting
  http://goc.grid.sinica.edu.tw/gocwiki
- Information System Troubleshooting (GRISs, GIISs, BDIIs)
  http://lfield.home.cern.ch/lfield/trouble.html (if you have to debug the MDS)

The SEE-VO

Enabling Grids for E-science in Europe

- Operational since end of March 2005
- Purpose:
  act as catch-all VO for new applications of Regional Users
- How to enroll : Get certificate and go to
  https://www.grid.auth.gr/services/voms/SEE/request.php
- Evaluation committee for new user/applications (SEE-EGAAP). Formal evaluation procedure will be established based on EGAAP existing policies but with relaxed requirements.

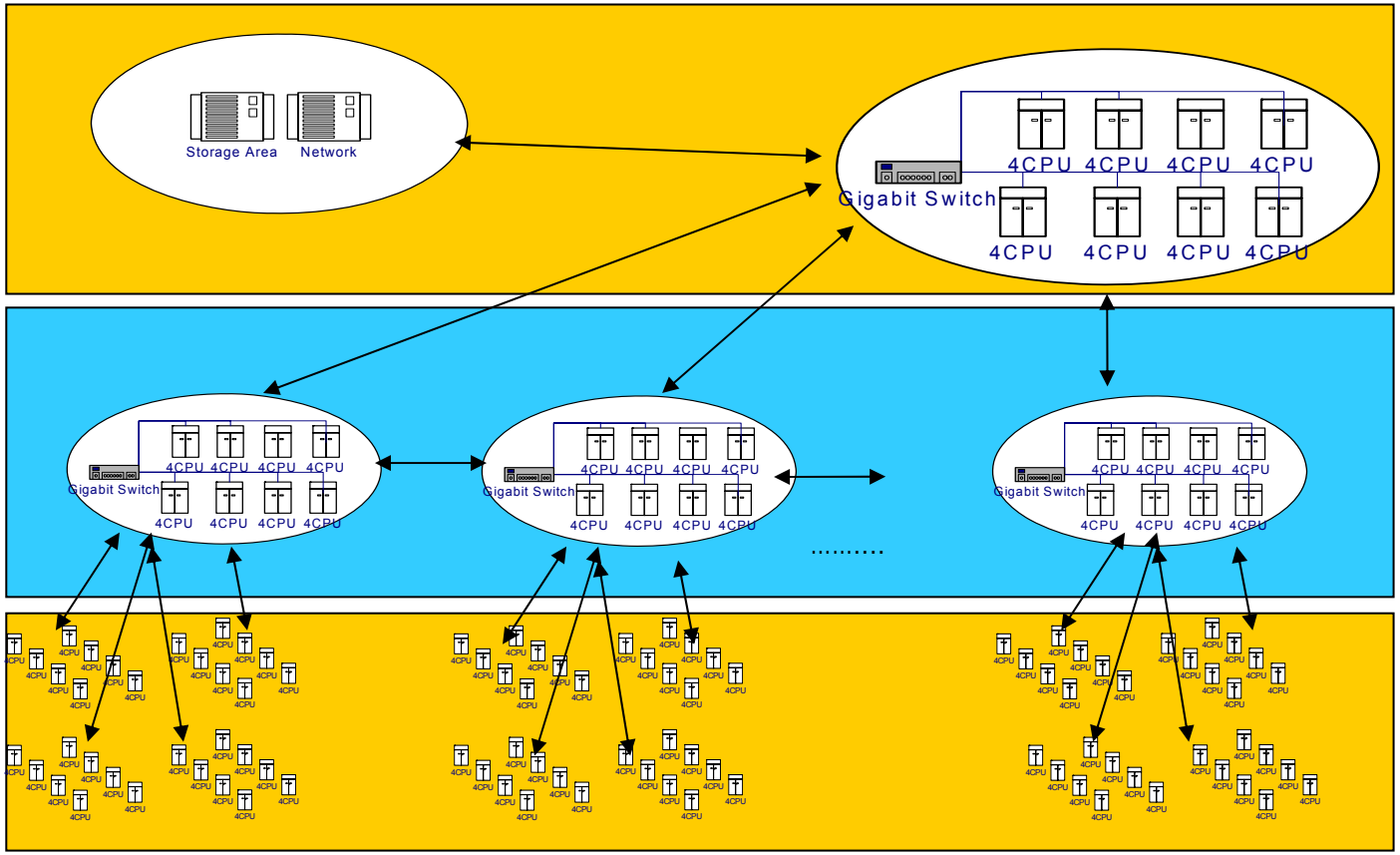
Hierarchical National Infrastructure

Enabling Grids for E-science in Europe

Επίπεδο 0 –
Κεντρικός Κόμβος ΕΔΕΤ
(64 CPUs-10TB SAN)

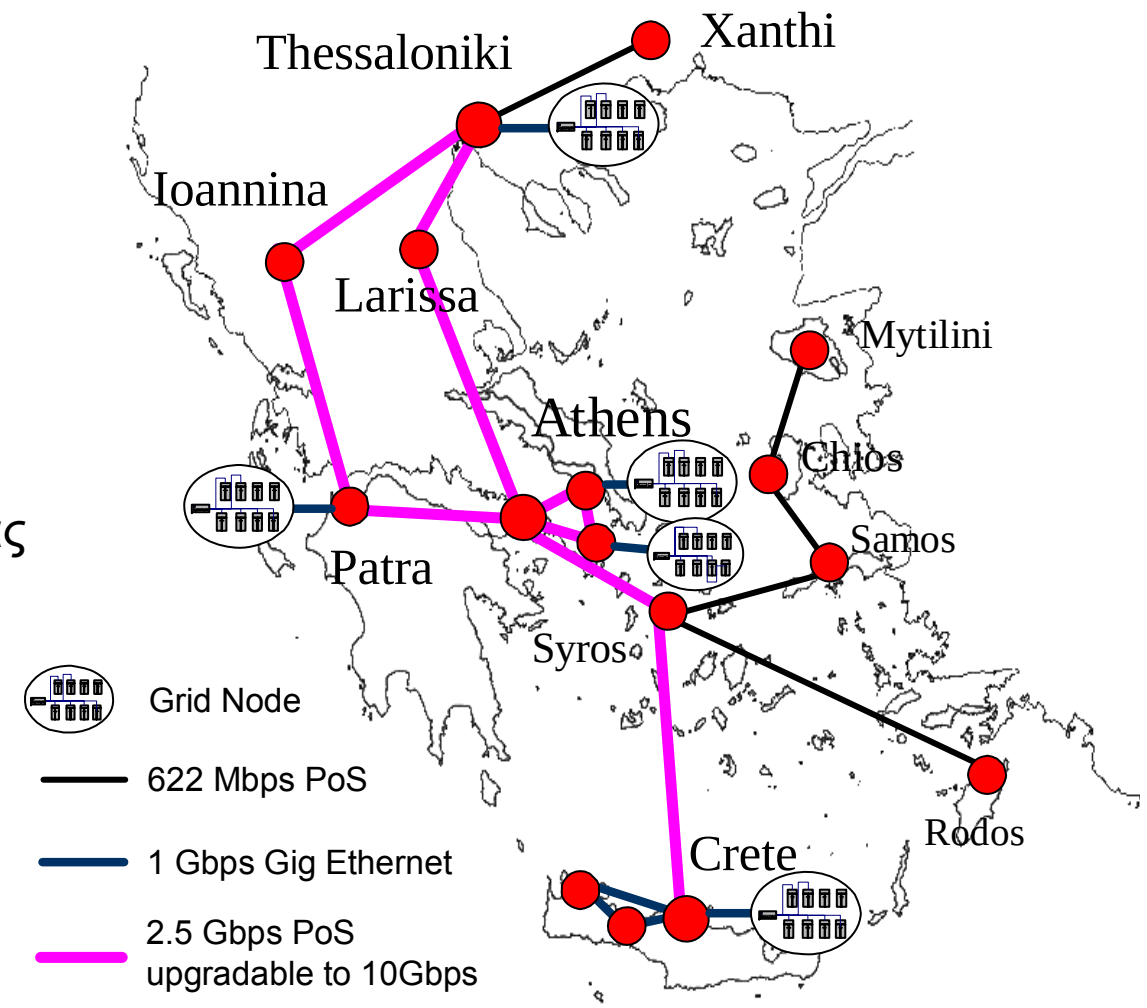
Επίπεδο 1 –
5 Περιφερειακοί Κόμβοι Αθήνα (2), Πάτρα, Θεσσαλονίκη, Κρήτη

Επίπεδο 2 –
Ολοκλήρωση άλλων υποδομών (ΠΣΔ, Εργαστήρια)

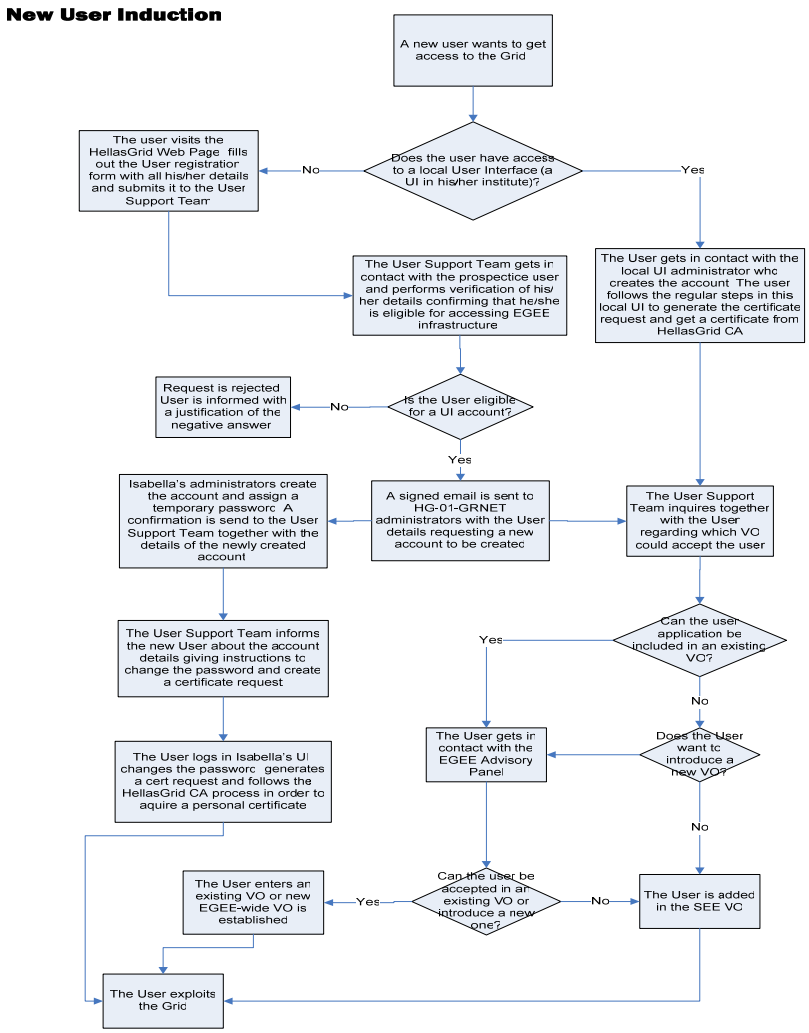


Networking Infrastructure

- Εκμετάλλευση ευρυζωνικών υποδομών ΕΔΕΤ2
- Τοποθέτηση κόμβων Grid μέσα σε Ακαδημαϊκά Ιδρύματα
- Δημιουργία ανθρώπινου δικτύου (Επιστημονικές και Ερευνητικές κοινότητες - ΕΔΕΤ)
- Δημιουργία Πόλων Τεχνογνωσίας ενσωματωμένων στα Ιδρύματα
- Χρήση AccessGrid για επικοινωνία και διαχείριση

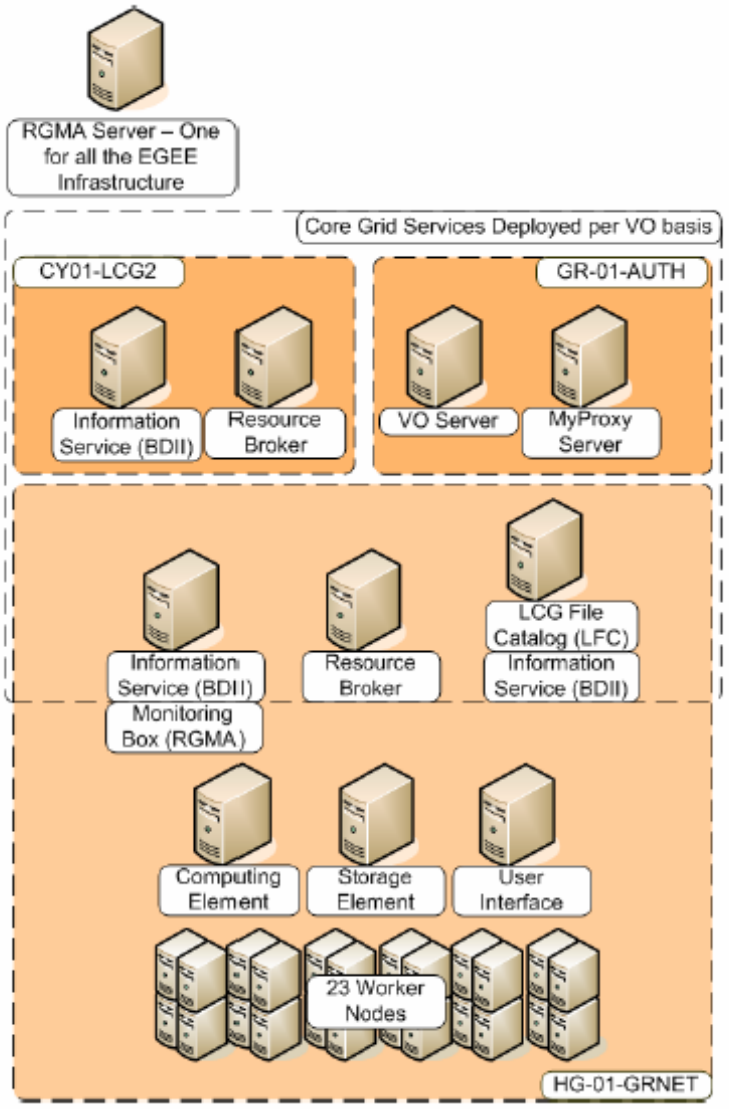


Policies and Procedures



- Policies exist for nearly any grid related activity, such as:
  - Adding a site (ROC-related)
  - Having a user get a certificate
  - Adding a user to a VO (AUP)
  - Adding a system administrator
  - Handling a security incident
  - …more we don't know…
- LCG Grid evolves faster than the policies themselves, so do "handle with care" ☺

Ανατομία ενός VO: SEE VO



- Κατάλογος χρηστών:
  - VO server & Myproxy
- Κατάλογος πόρων:
  - BDII (LDAP based!)
- Υπολογιστικοί πόροι:
  - Resource Broker (RB)
- Χωρητικοί πόροι:
  - LCG File Catalog (LFC)
- Τοπικές Υποδομές:
  - CE & WNs, SE, UI κλπ.

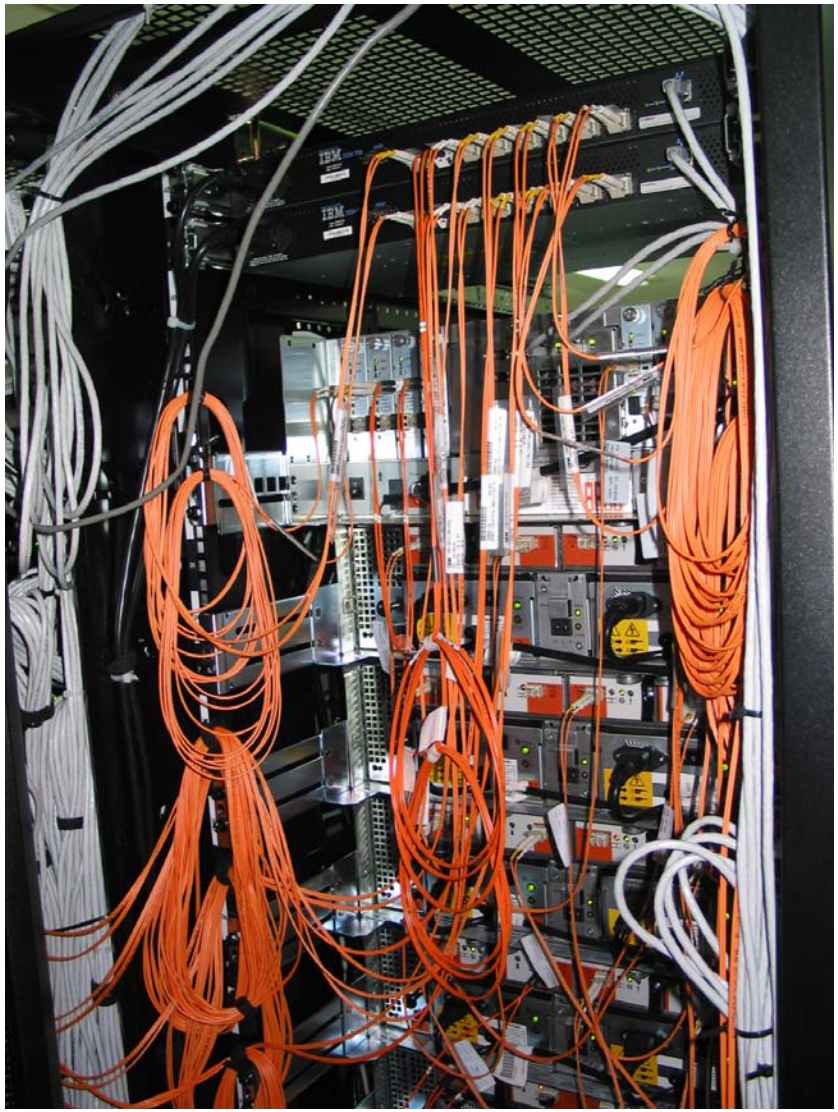
Support Resources

Enabling Grids for E-science in Europe

- Various channels have been established for User Support
- Problem Reporting
  - EGEE-SEE Helpdesk (https://helpdesk.egee-see.org/)
  - GGUS (https://gus.fzk.de) EGEE-wide
- Localized Documentation, FAQs, Tips, Instructions
  - EGEE-SEE Wiki (http://wiki.egee-see.org/index.php/Users)
  - EGEE-SEE Web Site (http://www.egee-see.org/). Recently restructured to provide easier navigation for the users.
- Other ad-hoc resources (various web sites, documents etc.)

Υποδομή HellasGrid I, Isabella

Enabling Grids for E-science in Europe



Production Level Grid Infrastructure

Enabling Grids for E-science in Europe

- High Availability
  - Through HW and SW redundancy
- Security aspects thoroughly examined
- Timely Resolution of problems
  - Efficient collaboration between team members
  - Close cooperation with VOs
- Aid in the deployment of New Sites / certification

Hardware Redundancy

Enabling Grids for E-science in Europe

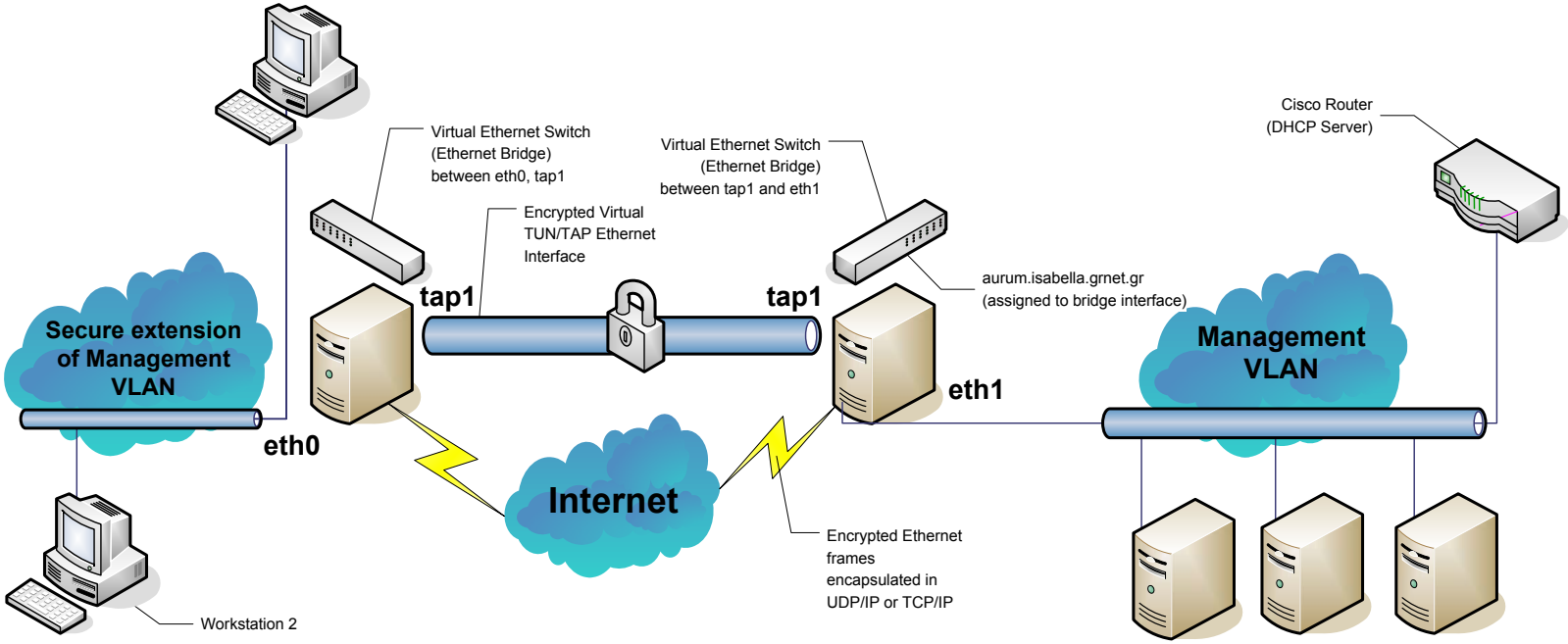
- Hardware Redundancy
  - Redundant disks on Service Nodes and WNs (RAID1)
  - Redundant Storage Infrastructure for SE at multiple levels
    - RAID5 Volumes on storage array
    - Redundant FC disk controllers
    - Redundant FC links in failover mode for GPFS storage nodes
    - Node redundancy at the GPFS level

Software Redundancy

Enabling Grids for E-science in Europe

- Redundant GPFS storage nodes
  - One primary / one secondary per Network Storage Device (NSD)
- Redundant network service instances
  - For DNS two on-site, two off-site servers
  - Similar redundancy in handling the NTP protocol

Security: OpenVPN



- Management interfaces unreachable from the outside
- Secure remote access to management VLAN using the free OpenVPN tool
  - Certificate-based authentication
  - SSL-based encryption

Security: Node hierachy

Enabling Grids for E-science in Europe

- Nodes are kept in a security hierarchy with different levels
  - Platinum: Backup server, Remote Console Access
  - Gold: Management Server
  - Copper: Worker Nodes for the Grid
- Explicitly defined trust relationships

Security: Intrusion detection, logging

Enabling Grids for E-science in Europe

- System-based Intrusion Detection Systems
  - OSS tools (e.g. chkrootkit)
  - Custom-written scripts (chksetuid with md5sum)
- Network based Intrusion Detection Systems
  - Snort

- *Factorized* Logging infrastructure based on syslog-ng
  - Logs collected centrally at the management server
  - Logs replicated at off-site location

Timely resolution of HW / SW problems

Enabling Grids for E-science in Europe

- Support contracts with vendors
  - IBM is the provider of the hardware of the HellasGrid I node
  - HP is the provider of the hardware of the HellasGrid II nodes
  - Support contracts are in place, in order to ensure timely replacement of components that are found to malfunction. It has been used in practice, equipment needs maintenance after all.
  - Support contracts extend to equipment such as the Uninterruptible Power Supply and the Air-Conditioners
  - Vendors are willing to keep up their promises, but the complexity of the equipment and its configuration setup should not be underestimated

Day-to-day Operations: Shifts

Enabling Grids for E-science in Europe

- Two shifts per day
  - 09:00 – 15:00, 15:00-21:00, Monday to Friday
- Ensures proper response in case of an emergency
- Handling of tickets coming from EGEE
- Timely service of requests from the User Support team

Day-to-day ops: Monitoring Tools

Enabling Grids for E-science in Europe

- EGEE Site Functional Tests
- Local monitoring tools
  - Ganglia
  - MRTG
- Vendor-specific tools
  - IBM Cluster Systems Management
    - Monitors various node health parameters
    - Sends e-mail alerts which can be routed to mobiles

Collaboration of Team Members

Enabling Grids for E-science in Europe

- **Request Tracker**
  - Web-based Ticketing System
  - Used for day-to-day collaboration
  - Permanent archive of information on all events during shifts
  - Facilitates integration of new team members
  - Knowledge base for all of HellasGrid
- **Weekly meetings**
  - Review of open tickets
  - Planning of future activities

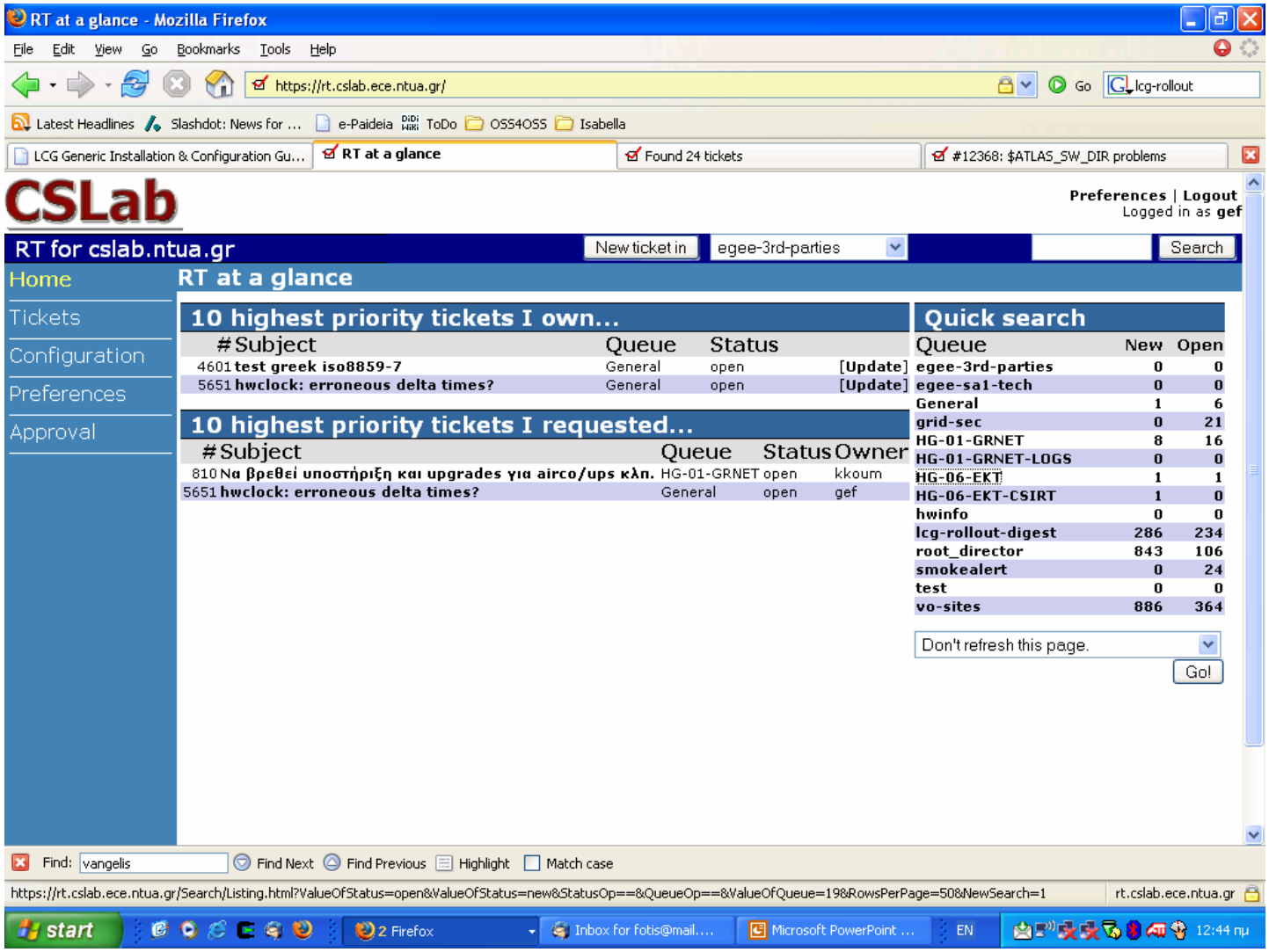
Introduction of new sites in HG

Enabling Grids for E-science in Europe

- Streamlining of new site installations
  - Guide for new HW installations
  - Customized instructions for OS deployment
- Certification Period
  - Certification SFTs run by the HG-01-GRNET team for all yet uncertified sites
  - Site enters production when the tests have run without problems for 5 days

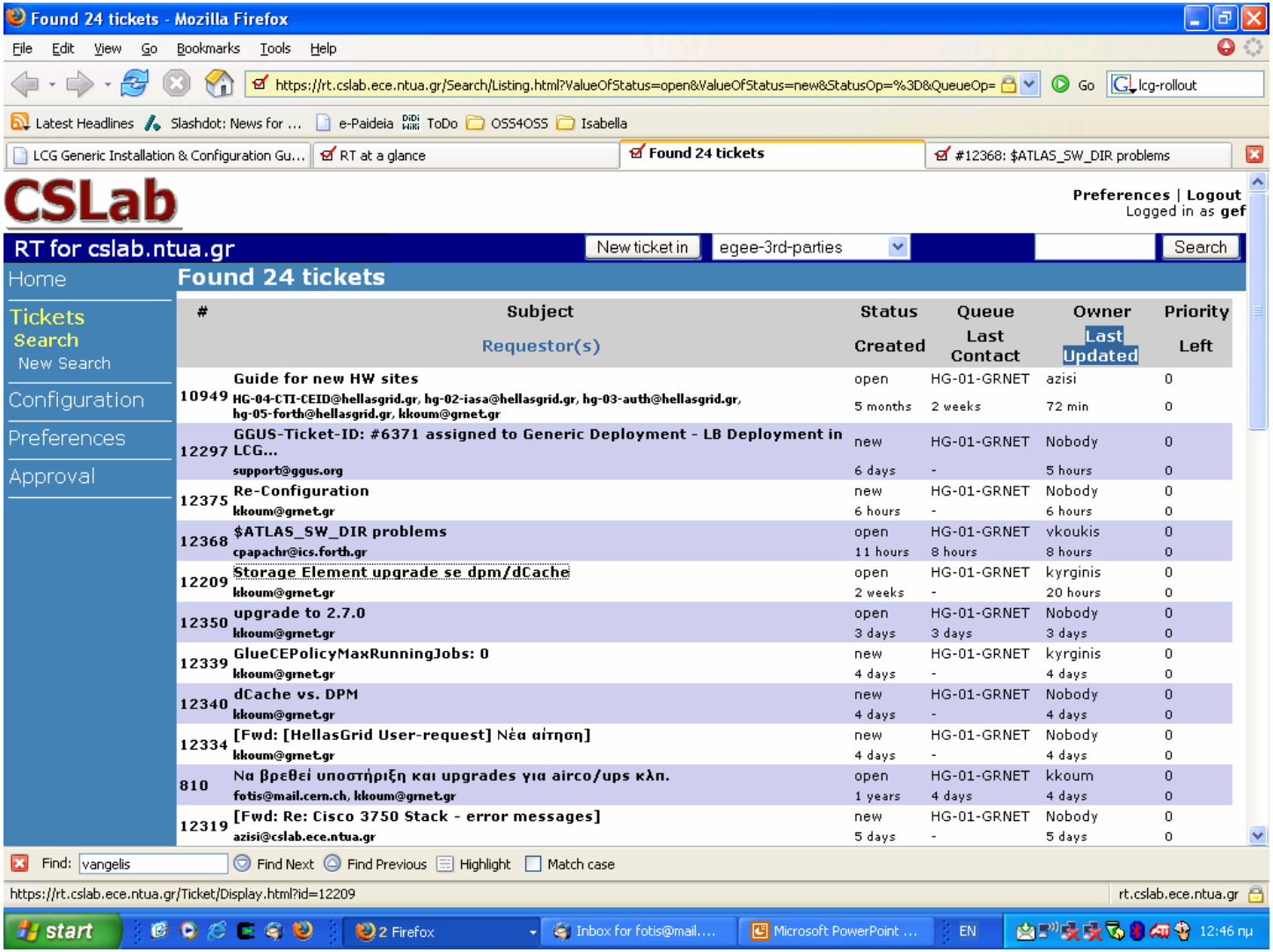
RT ticketing system: the big picture

Enabling Grids for E-science in Europe



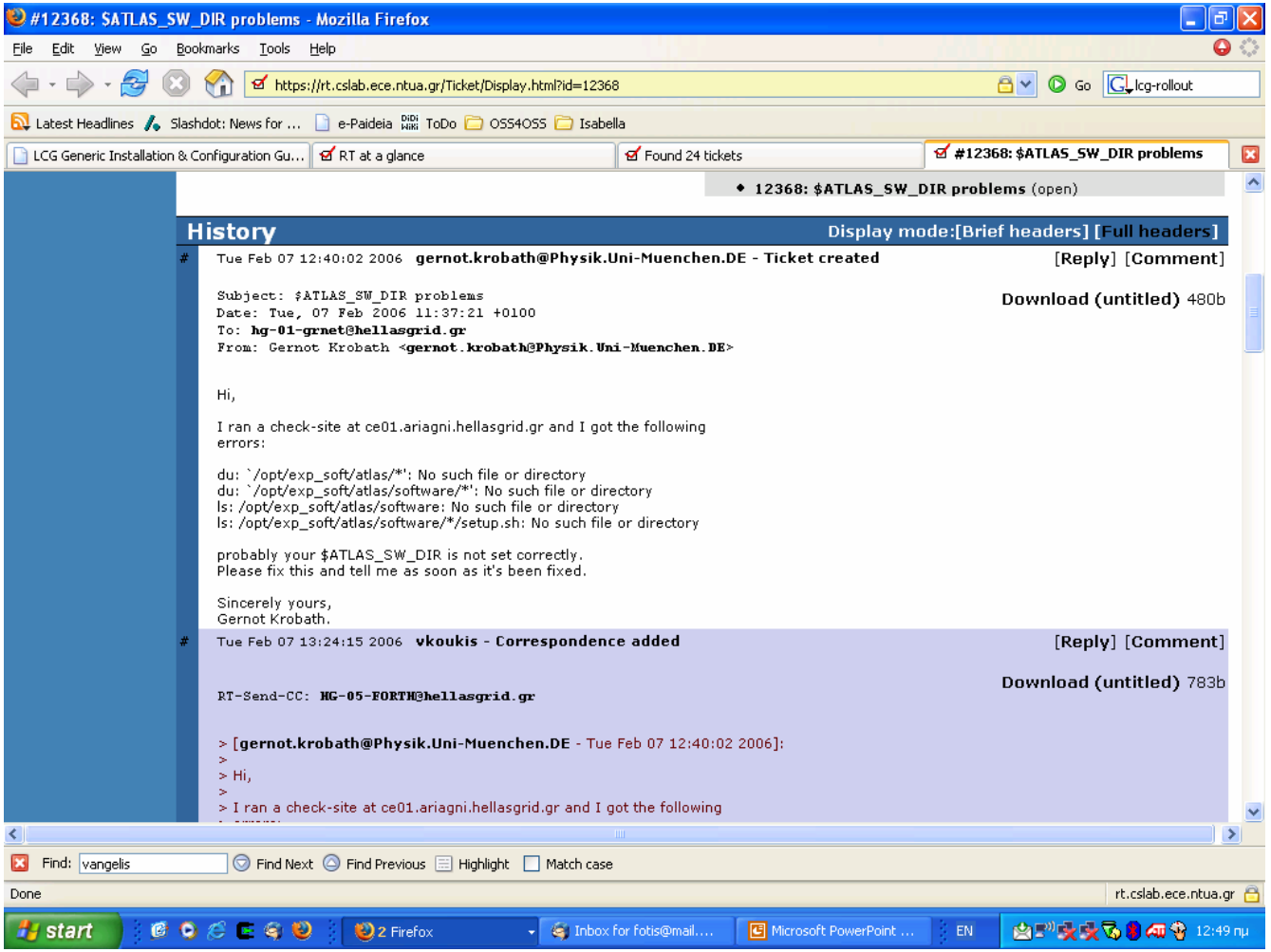
RT ticketing system: Queue status

Enabling Grids for E-science in Europe



RT system: managing a ticket

Enabling Grids for E-science in Europe

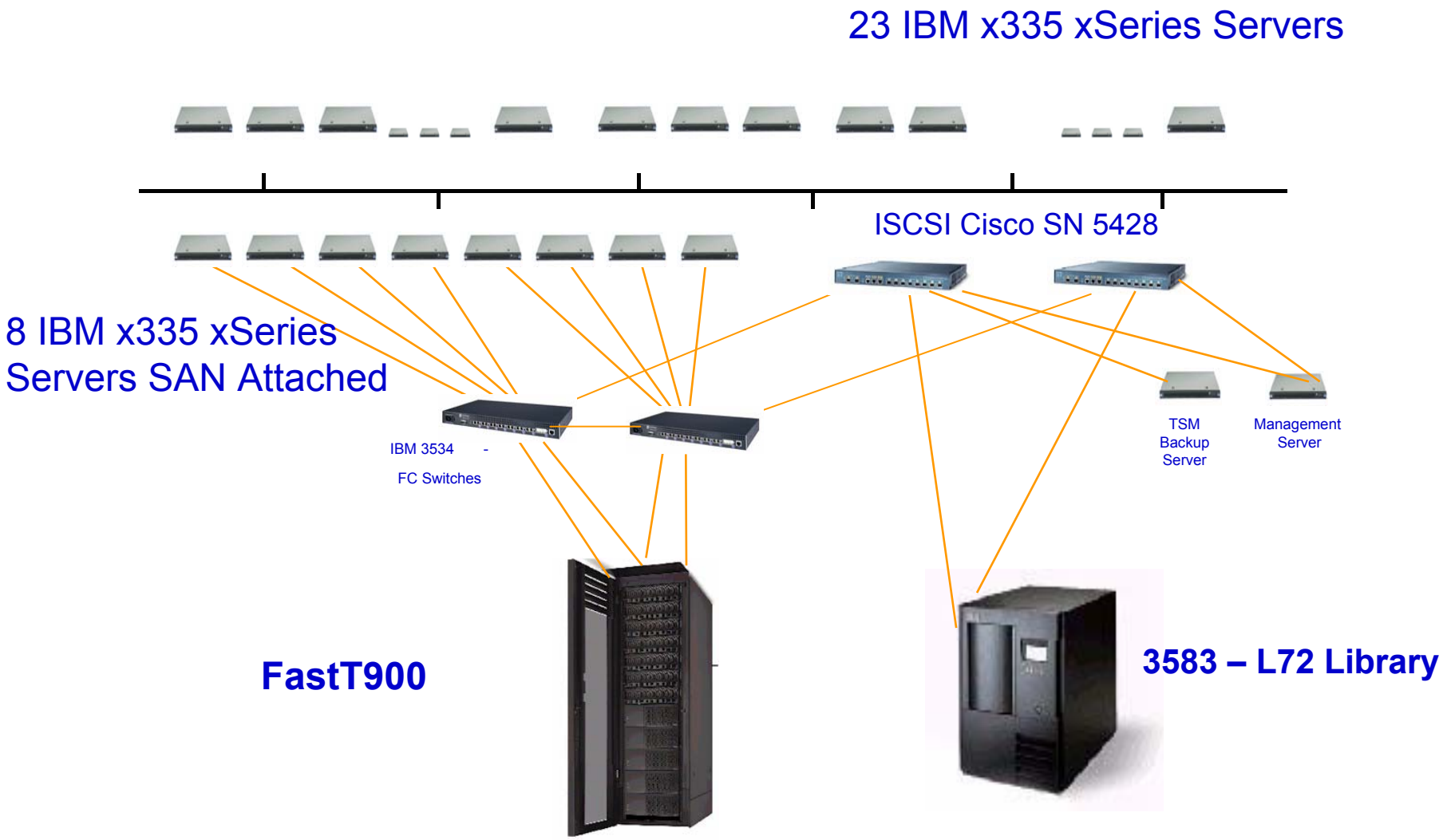


Notorious Tickets of the past

Enabling Grids for E-science in Europe

- #62: Evaluating GPFS suitability for our case
- #304: Optimizing configuration of CISCO switch stack
- #809: Document startup, boot, halt shutdown procedures
- #888: Daily check-list of the hg-01-grnet node
- #942: GPFS debugging & tiobench tests
- #2078: mprime and kernel-compiles, cpu temperature tests
- #4804: Heavy stress tests of HG-01-GRNET (minor repairs)
- #7200: Getting mpirun to work correctly with torque
- #7293: VPN architecture and OOB management
- #8460: Unscheduled downtime due to power
- #8617: Unscheduled downtime due to air-condition failure

Putting it alltogether



GPFS Capabilities

Enabling Grids for E-science in Europe

- High-performance parallel, scalable file system for Linux/AIX cluster environments
- Full Load Balancing: Shared-disk filesystem where every cluster node can have concurrent read/write access to (**meta-**)data
- Fail-over: High availability through automatic recovery from node and disk failures

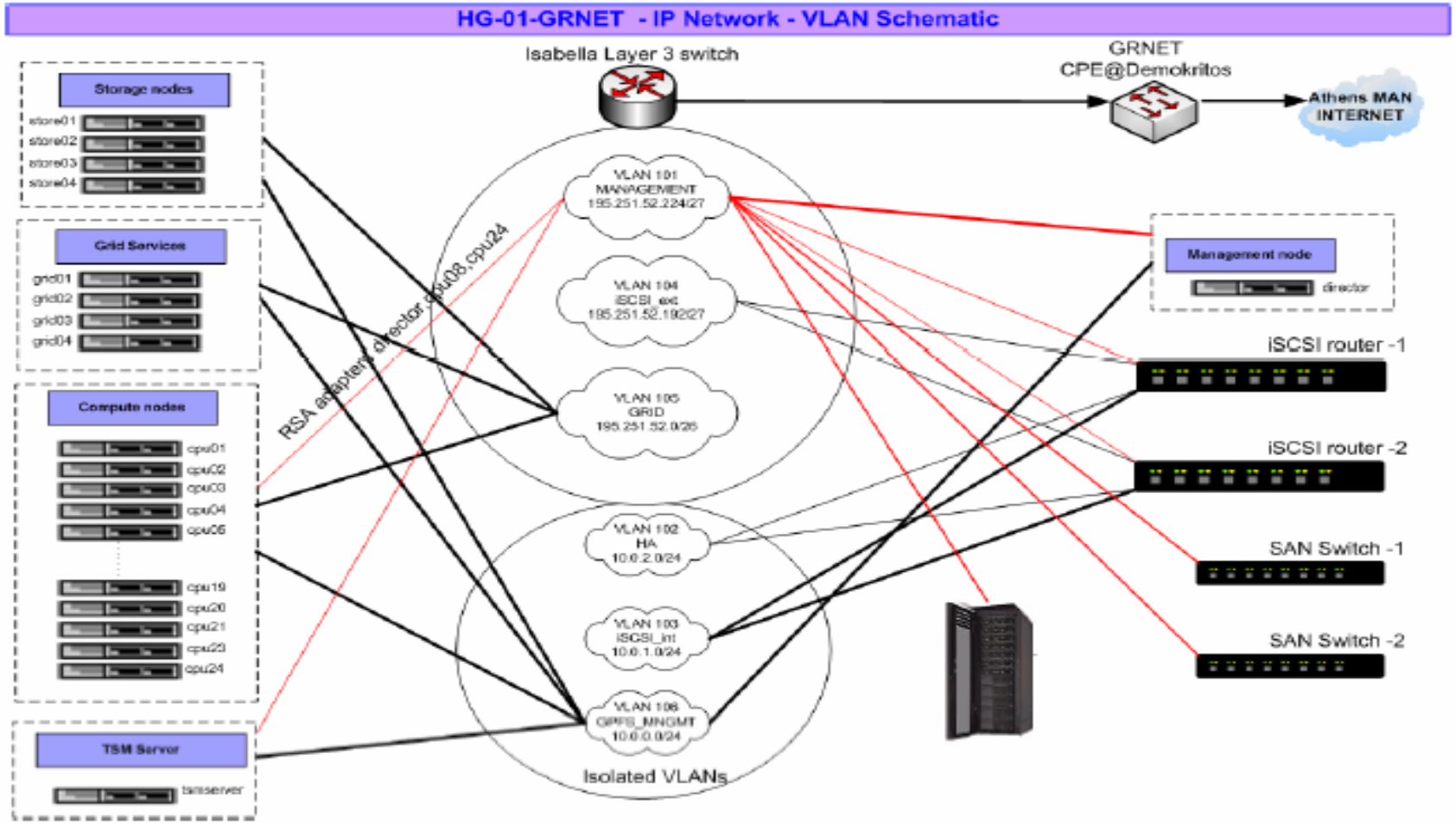
Η Υποδομή HellasGrid, Φάση I & II

Enabling Grids for E-science in Europe

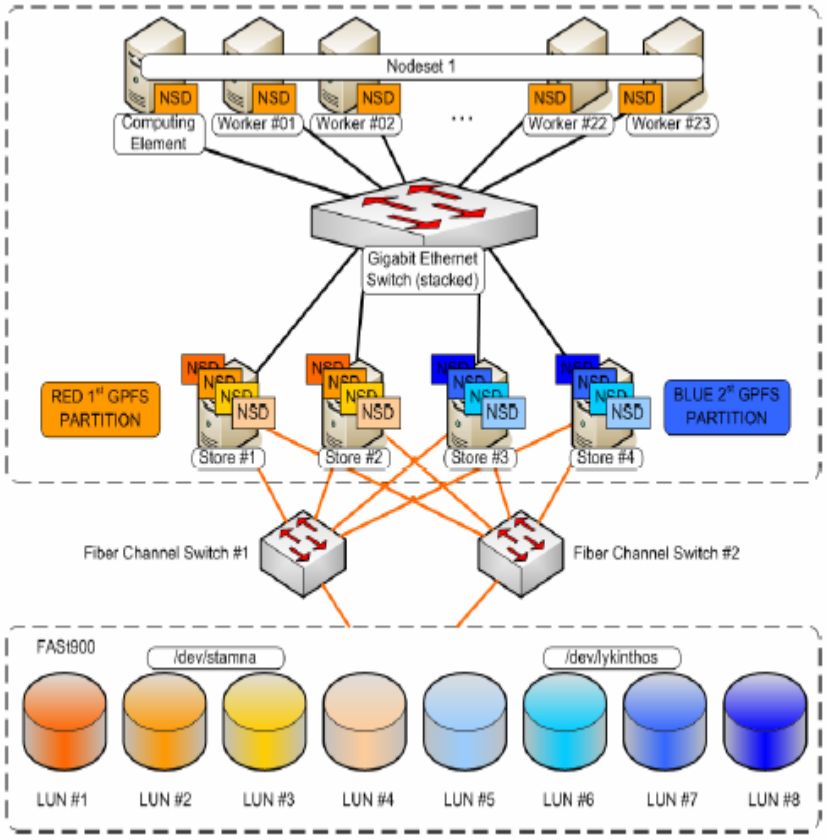
- HellasGrid I (500.000 €)
  - Βρίσκεται στον Δημόκριτο, Αγία Παρασκευή (a.k.a. Isabella)
  - 34 dual Intel P4 Xeon @ 2.8GHz, 1GB RAM, 2x 70GB SCSI HDD, 2x Gbit
  - IBM FAStT900 Storage Area Network, ολοκληρωμένο σύστημα
    - 2x Redundant Fiber Channel Controllers with 1Gbyte Cache each
    - 70x146.8GB= **10,276TB raw storage capability,** over 5 disk shelves
    - Πλήρως αυτοματοποιημένη λύση, **hot spare + hot swap + monitoring**
  - Tape Library με δυνατότητα έως ~30 TBytes, integrated monitoring
  - Παραδόθηκε στο ΕΔΕΤ από την IBM κατά τον Δεκέμβριο του 2004
- HellasGrid II (1.100.000 €)
  - 5 φυσικοί κόμβοι: ΕΚΤ (256), ΙΕΣΕ (128), ΑΠΘ (128), ΙΤΕ (128), ΙΤΥ (128)
  - ~700 Επεξεργαστές x86_64, 2 GB RAM, 1x 80GB SATA HDD, 2x Gbit
  - ~20 TBytes συνολικός αποθηκευτικός χώρος σε τεχνολογία SAN (5x 4TBs)
  - ~50TBytes Tape Library, το οποίο πρόκειται να εγκατασταθεί στο ΕΚΤ
  - Υπό εγκατάσταση (έχει ήδη γίνει η ποσοτική παραλαβή)

Υποδομή HellasGrid I, Isabella

Enabling Grids for E-science in Europe



Υποδομή HellasGrid I, Isabella



- Ο πρώτος κόμβος της Ελληνικής υποδομής Πλέγματος αποτέλεσε ένα εξαιρετικό εργαλείο για την συλλογή τεχνογνωσίας.
- Αυτή θα αξιοποιηθεί στην δεύτερη φάση του έργου, εις όφελος των νεώτερων κόμβων και χρηστών.
- Εξαιρετικά πρωτότυπη η οργάνωση του SAN και των συστημάτων αρχείων.

Monitoring of HG-01-GRNET

Enabling Grids for E-science in Europe

- hg-01-grnet@hellasgrid.gr, hwinfo(…)@rt.cslab.ntua.gr
- Tivoli Storage Manager – TSM backup interface
- CISCO switch stack
- CISCO iSCSI switches
- Cluster Systems Management (CSM) utilities, RSA, RCM
- GPFS and filesystems monitors
- Storage Area Network GUI client

Q & A

Enabling Grids for E-science in Europe