



LCG-France Tier-1 & AF

Réunion de Coordination

Fabio Hernandez

fabio@in2p3.fr

Lyon, 10 juillet 2008

dapnia

cea

saclay



CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE

► Table des Matières



- Réunions CCRC, GDB et MB de juin et juillet 2008
- Disponibilité, fiabilité, efficacité du site
- Chantiers en cours
- Thème du jour



- Agendas:
 - GDB: <http://indico.cern.ch/categoryDisplay.py?categId=31181>
 - MB: <http://indico.cern.ch/categoryDisplay.py?categId=666>
- Principaux sujets traités
 - Monitoring distribué
 - Benchmarking CPU
 - Tests du CE CREAM
 - Collecte des informations de capacité installée des sites

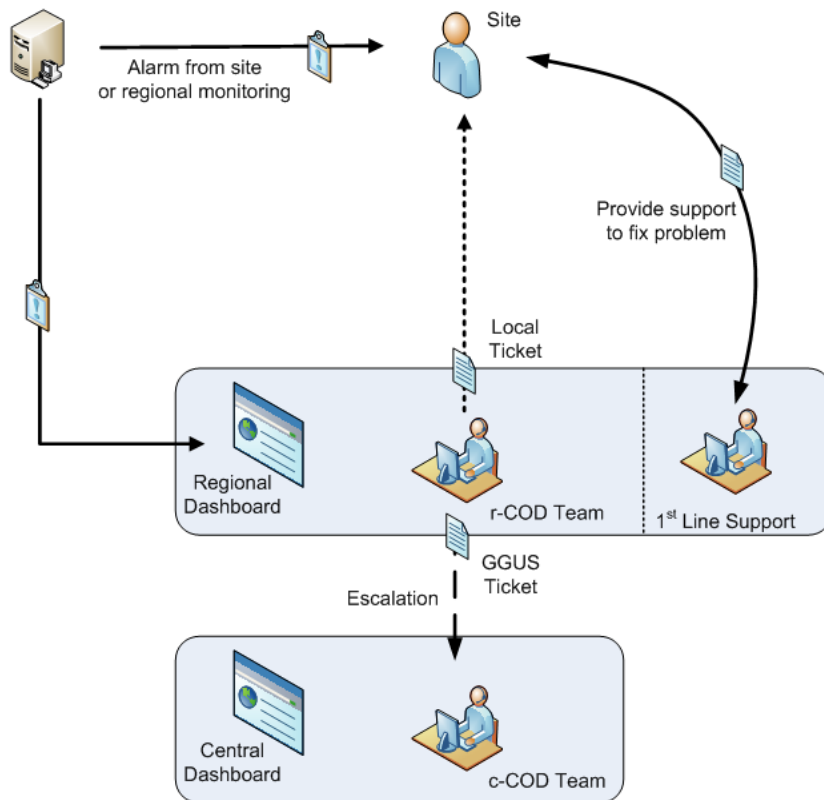
Monitoring distribué (1/2)



eGee

Enabling Grids for E-science

Future operational model



Processus d'escalade des problèmes détectés par les outils de monitoring

Source: James Casey, GDB, 10/07/2008

EGEE-II INFSO-RI-031688

To change: View -> Header and Footer

8

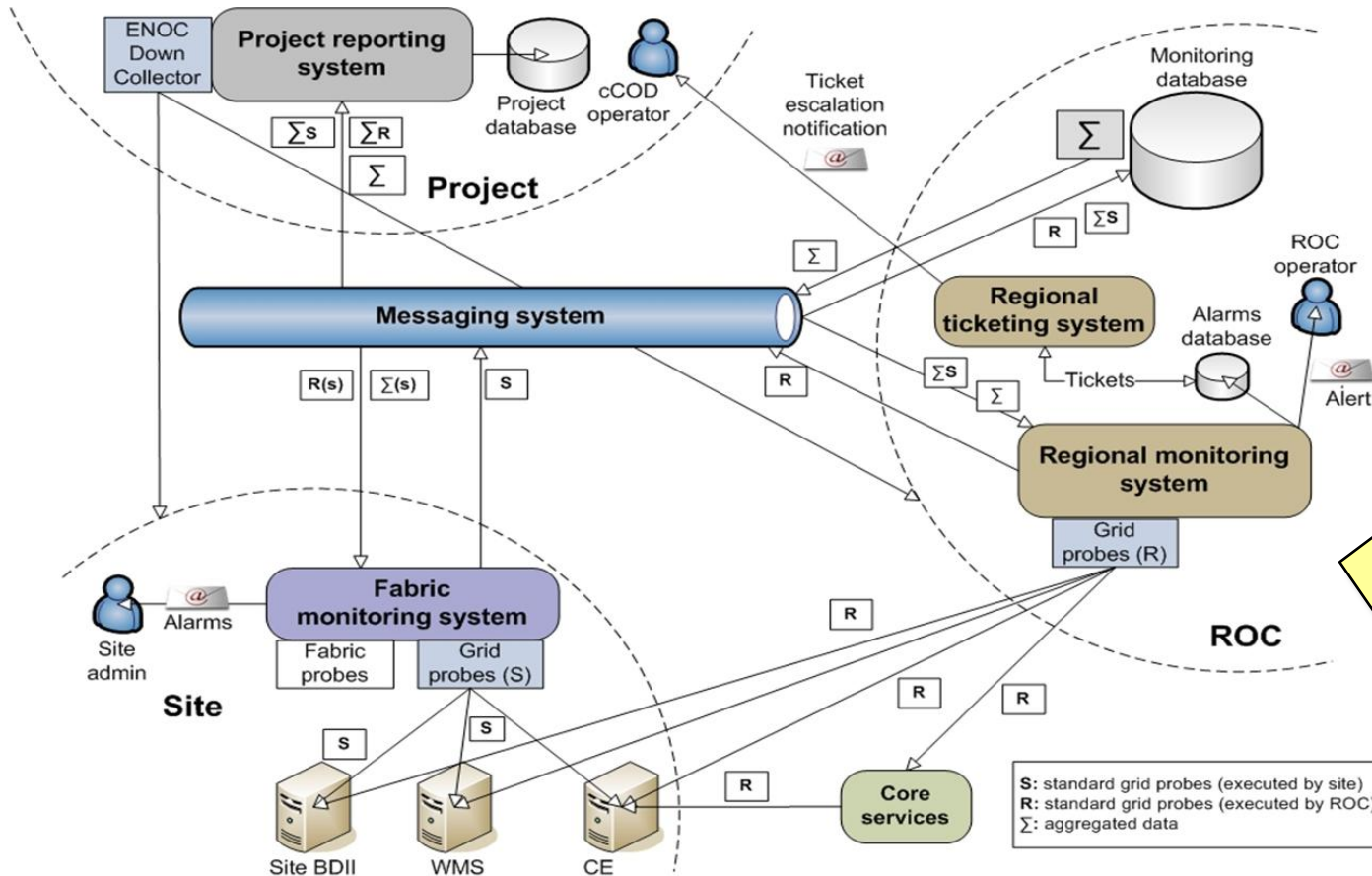
"Operations Automation Strategy": <https://edms.cern.ch/document/927171>

Monitoring distribué (2/2)



Multi level monitoring framework

Enabling Grids for E-science



Régionalisation et consolidation nécessaires des outils des opérations, y compris le portail CIC

Source: James Casey, [GDB, 10/07/2008](#)

CPU Benchmarking



Proposal

CERN IT
Department

- Adopt SPECcall_cpp2006 and environment presented
- CERN will make script available
- Sites to purchase a SPECcpu2006 licence
- Name: Call it HEP-SPEC or something like this?
- Transition of existing capacities: Run old benchmark the old way (whatever that was), and HEP-SPEC on samples of all HW configs in use
- Transition of existing requests/commitments: I presume these were based on some reference machine type; both benchmarks to be run there

Adoption d'une nouvelle suite de benchmarking et d'une nouvelle unité pour exprimer la puissance CPU (offre et demande)

Groupe de travail à constituer pour proposer un plan de transition

► Tests de CREAM (nouveau CE)



Test performed

- **Installation and configuration**
 - Followed the formal installation procedure, specially check the package dependency
 - Tested different installation scenario
 - Configured it by YAIM
- **Basic functionalities**
 - Submission through CLI, job status check, delegate proxy etc.
- **Stress testing**
 - Submission through CLI
 - 9800 jobs per day with 49 users
 - Jobs accumulated in queues as fast as possible

Tests de CREAM (2/2)



Enabling Grids for E-science

Test results

- **Basic functionality tests passed**
- **Dependency missing on some packages**
 - tomcat, mysql-server and mysql-connector-java
- **Configuration issues**
 - Configurations of BLPaser server and blah
 - LCAS and LCMAPS generate too many logs
 - Issue with the upgrade of Glexec
 - fixed
- **Other issues**
 - Need to restart tomcat if CAs are updated
 - problems in trustmanager
 - too many files left under the home directories of pool accounts
 - Authentication fails with new type of VO attributes in VOMS proxy
 - fixed

EGEE-III

CREAM CE certification and testing - Qing

Proposition de passage en PPS afin de tester CREAM dans d'autres conditions et avec d'autres systèmes de batch

Déploiement généralisé prendra encore quelques mois.

Pouvons-nous reprendre le travail d'interfaçage de CREAM et BQS?



Enabling Grids for E-science

Stress test results

- **System load**
 - CPU load is quite low, even less than 1 for most of time, only when submitting massive jobs to cream CE, it can reach 9
 - Memory usage is low too, less than 2GB
 - Disk usage can increase
 - Can be solved by purging jobs and limiting the log level of services
- **Job submission**
 - Job submission to CREAM CE can fail
 - Happened in last two days test, even more than 3% of jobs could not be submitted
 - Job success rate
 - More than 99.5% jobs succeeded in 6 of 10 days test
 - The worst failure rate is about 9.5%
 - *Most of failures only give error message, "blah error"*

EGEE-III

CREAM CE certification and testing - Qing

6

Source: Di Qing, [GDB](#), 10/07/2008

Resources: installées vs. utilisées



The requirements



■ The goal

- Provide the management with information about installed capacity (per VO).
- Provide the management with information about resource usage (per VO).
 - This info can also be used by VO operations people in order to "monitor" their usage of the resources

■ Main focus

- Storage resources
- Computing resources

■ Dynamic view

- Information should be retrieved as dynamically as possible in order to be reliable
- The information will be made available through the WLCG accounting system
- Use the information system as source



WLCG Grid Deployment Board, CERN 9 July 2008

3

Source: Flavia Dono, GDB, 10/07/2008

Ressources: installées vs. Utilisées (cont.)



The Storage Resources



■ GlueSA Sizes

- GlueSAReservedOnlineSize
 - Space physically allocated to a VO (or a set of VOs) – Installed capacity
- GlueSATotalOnlineSize (in GB = 10^9 bytes)
 - Total Online Space available at a given moment (it does not account for broken disk servers, draining pools, etc.)
- GlueSAUsedOnlineSize (in GB = 10^9 bytes)
 - Size occupied by files that are not candidates for garbage collection – Resources usage
 - For Castor this is equal to GlueSATotalOnlineSize for T1D0 (they differ for T0D1)
- GlueSAFreeOnlineSize (in GB = 10^9 bytes)
 - Free = Total - Used
- GlueSACapability
 - Key/value pair to publish agreed characteristics of Storage Area, i.e. access=read, network=WAN-LAN, etc.
 - Experiments can use this field to select dCache pools



WLCG Grid Deployment Board, CERN 9 July 2008

The Computing Resources



- Ongoing work to make published information about installed capacity more reliable (see presentation of Steve Traylen at GDB)
- Current tools publish always 1 GlueCluster and 1 GlueSubCluster when a GlueCE is published (problematic CPU counting)
- Difference between SubCluster PhysicalCPUs and GlueCETotalCPUs
- Is Subcluster PhysicalCPUs coherent with SubCluster Host BenchmarkSI00?
- When SubCluster PhysicalCPUs=0 consider GlueCETotalCPUs: Number of slots per queue. Compatible with Subcluster Host BenchmarkSI00? Is this number correct (extracted from batch system)?
- VOView TotalCPUs deprecated. Difficult to guess VO shares.



WLCG Grid Deployment Board, CERN 9 July 2008

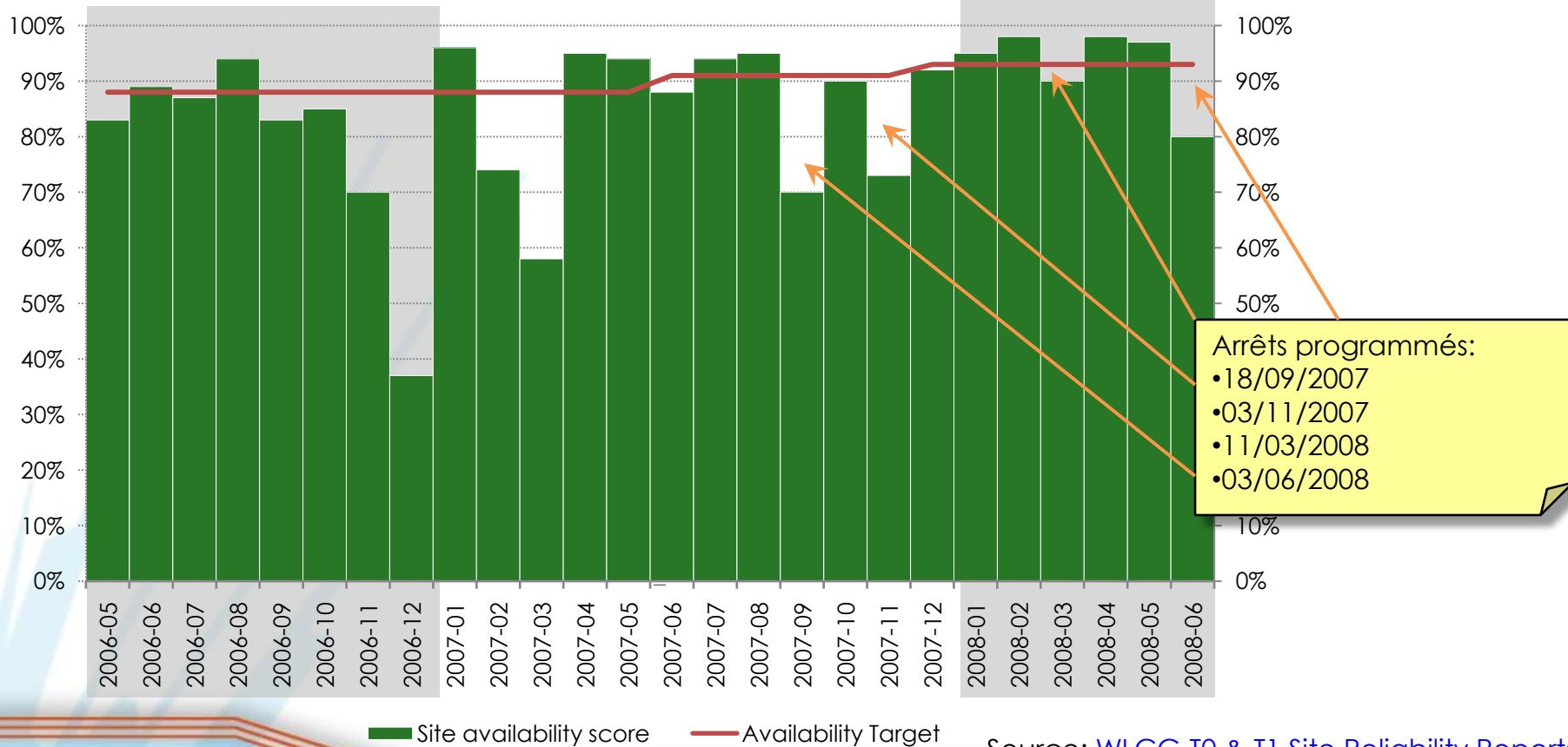
11

Source: Flavia Dono, [GDB, 10/07/2008](#)

Disponibilité du tier-1



CC-IN2P3 Tier-1: monthly availability score
(VO OPS)

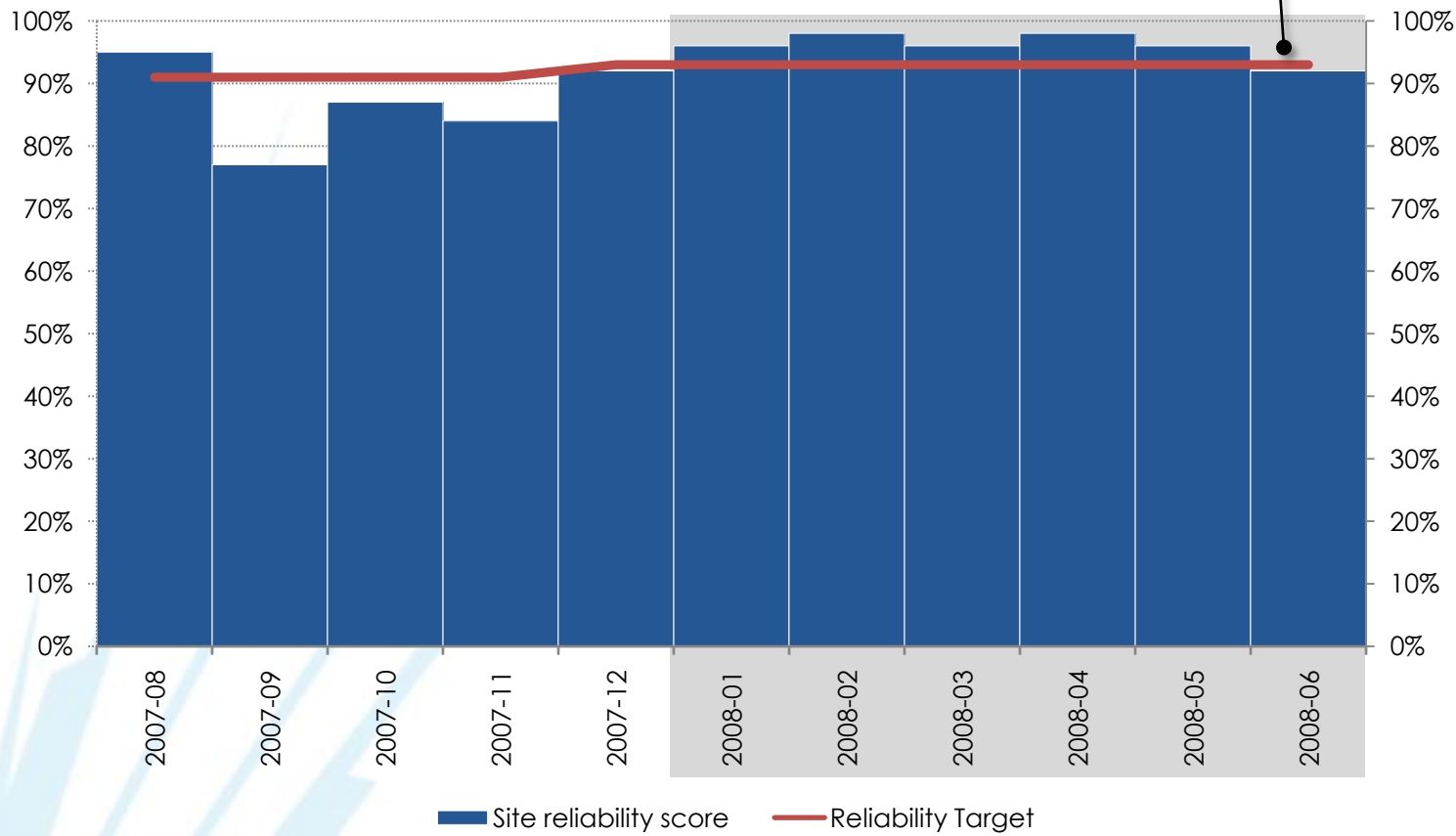


Source: [WLCG T0 & T1 Site Reliability Reports](#)

Fiabilité du tier-1



CC-IN2P3 Tier-1: monthly reliability score
(VO OPS)



Fiabilité en
juin 2008:
92%

Source: [WLCG T0 & T1 Site Reliability Reports](#)

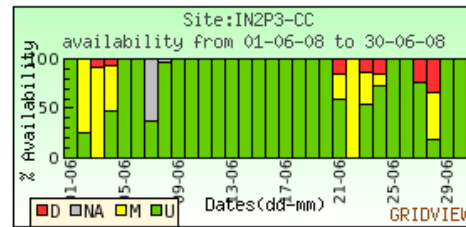
Disponibilité des sites: tier-1 (1/4)



- Juin 2008

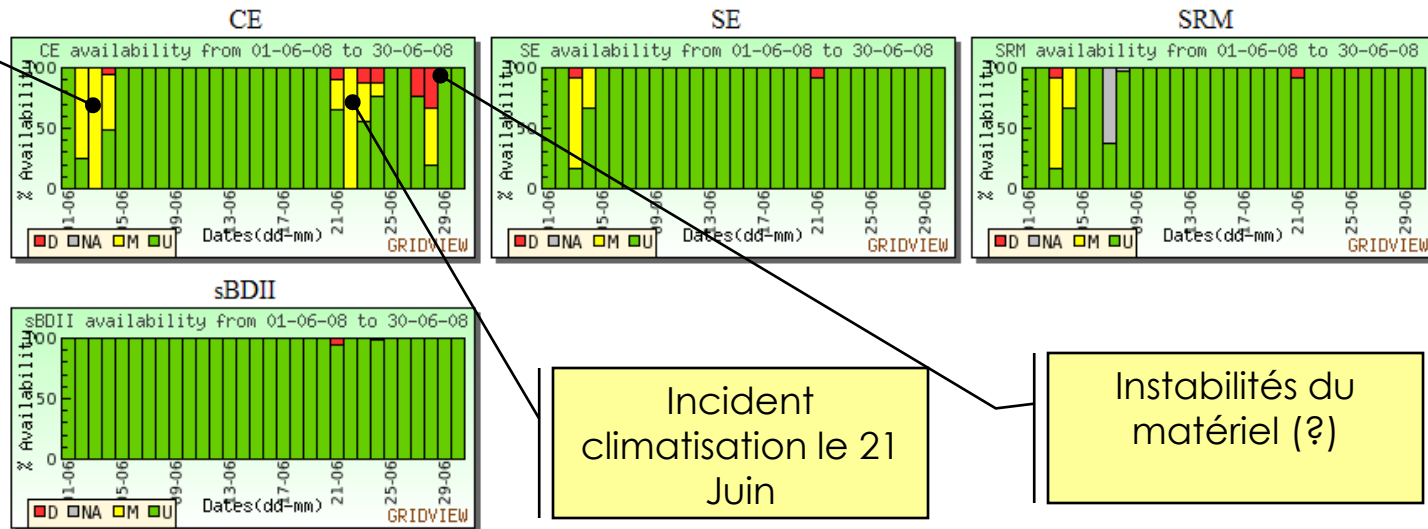
Score: 80%

Overall Service Availability for Site:IN2P3-CC VO:OPS (Daily Report)



Arrêt programmé
le 3 Juin

Individual Service Availability for site:IN2P3-CC VO:OPS (Daily Report)



Incident
climatisation le 21
Juin

Instabilités du
matériel (?)

Source: <http://gridview.cern.ch>

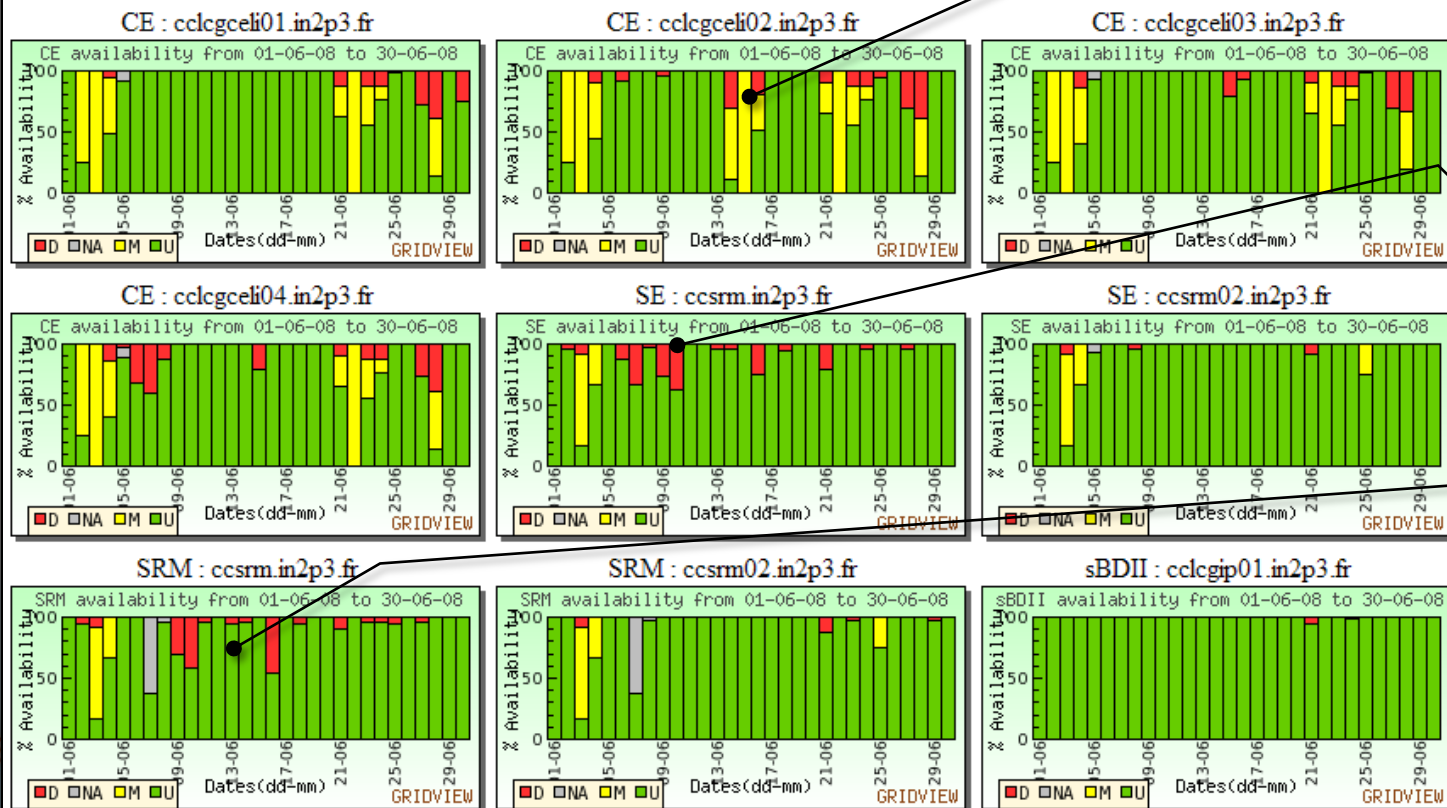
Disponibilité des sites: tier-1 (2/4)



- Juin 2008 (suite)

Service Instance Availability for site:IN2P3-CC VO:OPS (Daily Report)

Site Services



Incidents
avec le CE?

???

Indisponibilité
dCache?

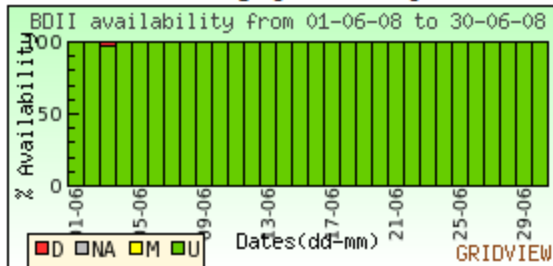
Disponibilité des sites: tier-1 (3/4)



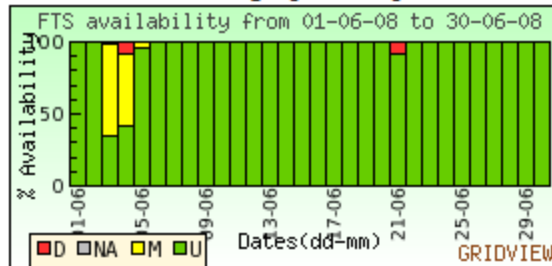
- Juin 2008 (suite)

Central Services

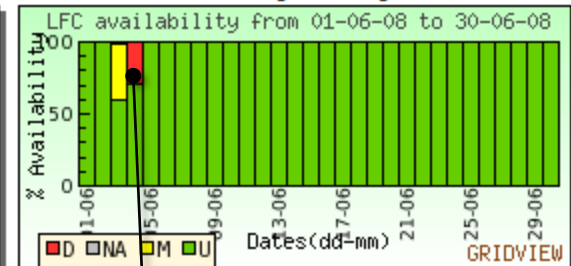
BDII : cecgtopbdii01.in2p3.fr



FTS : cecgftsprod.in2p3.fr



LFC : lfc-prod.in2p3.fr



Indisponibilité
LFC?

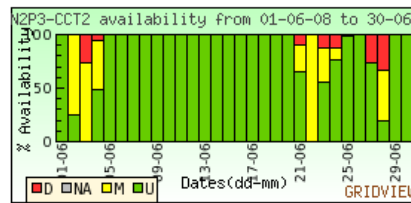
Disponibilité des sites: tier-2 (4/4)



Disponibilité: 82%
(Fiabilité: 95%)

• Juin 2008

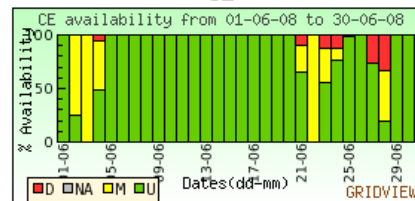
Overall Service Availability for Site:IN2P3-CCT2 VO:OPS (Daily Report)



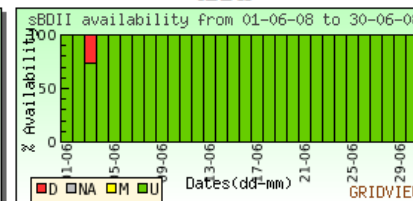
Individual Service Availability for site:IN2P3-CCT2 VO:OPS (Daily Report)

Instabilités du matériel (?)

CE



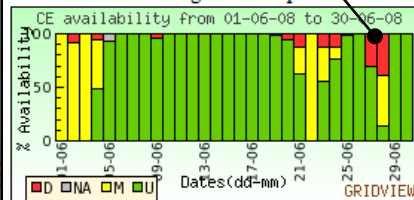
sBDII



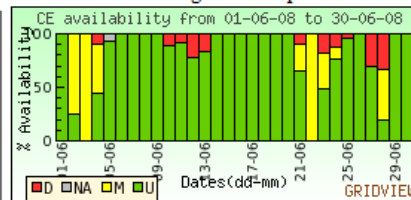
Service Instance Availability for site:IN2P3-CCT2 VO:OPS (Daily Report)

Site Services

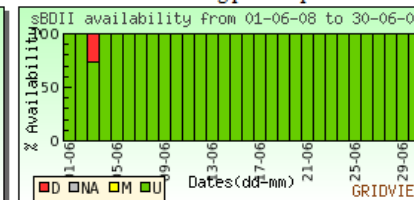
CE : cclcgceli05.in2p3.fr



CE : cclcgceli06.in2p3.fr



sBDII : cclcgip03.in2p3.fr



Source: <http://gridview.cern.ch>

Disponibilité des sites EGEE-FR



• Juin 2008

France (France)

AUVERGRID	96 %	96 %
CGG-LCG2	99 %	99 %
GRIF	100 %	100 %
IBCP-GBIO	85 %	85 %
IN2P3-CC	82 %	96 %
IN2P3-CC-T2	82 %	95 %
IN2P3-CPPM	90 %	92 %
IN2P3-IPNL	99 %	99 %
IN2P3-IRES	99 %	98 %
IN2P3-LAPP	99 %	99 %
IN2P3-LPC	99 %	100 %
IN2P3-LPSC	98 %	99 %
IN2P3-SUBATECH	91 %	91 %
IPSL-IPGP-LCG2	95 %	95 %

Region	Avail- ability	Reli- ability
France	94 %	96 %
UKI	84 %	93 %
NorthernEurope	81 %	84 %
AsiaPacific	80 %	83 %
CERN	78 %	81 %
Russia	77 %	79 %
SouthEasternEurope	77 %	83 %
SouthWesternEurope	75 %	83 %
GermanySwitzerland	74 %	76 %
CentralEurope	74 %	84 %
Italy	65 %	72 %

- Voir résumé dans la présentation faite en séminaire interne le 27/06/2008
 - <http://cc.in2p3.fr/article1604.html>

► Travaux en cours ou terminés



- Problème d'accès aux données LHCb gérées par dCache via gsidcap
 - Ticket GGUS: https://gus.fzk.de/ws/ticket_info.php?ticket=36625
 - Travail intensif de Pierre, Lionel et Jonathan
 - *En interaction avec les experts dCache et LHCb*
 - Résultat: bug identifié du client gsidcap dans la gestion des contextes de sécurité
 - Contournement: utilisation d'une seule porte gsidcap pour LHCb, contre 3 auparavant
 - En attente de correction (?)

▶ Travaux en cours ou terminés (suite)



- Tests d'une nouvelle version de glexec pour déploiement en mode logging-only
- Consolidation des VO boxes CMS
 - Et tests de déploiement de CRAB (outil CMS pour la soumission de jobs d'analyse)
- Problème au redémarrage de FTS après l'arrêt début juin
 - Voir présentation de David aujourd'hui
- Instabilité de LFC
 - Voir présentation de David aujourd'hui

▶ Travaux en cours ou terminés (suite)



- Acquisition de serveurs de disque: campagne 2008
 - Procédure d'appel d'offres finie et candidat sélectionné
 - Livraison prévue en septembre 2008: 3.5 PB (environ 120 machines)
- Nouvelle release de dCache v1.8.0-15p4 installée le 3 juin
- Adresse (et workflow) spécifique pour la réception de messages nécessitant une intervention urgente
 - Un groupe de personnes bien identifiées de chaque VO habilitées à envoyer des alertes (courriels signés électroniquement)
 - lhc-alarm@cc.in2p3.fr
 - En savoir plus: <https://twiki.cern.ch/twiki/bin/view/LCG/OperationsAlarmsPage>
- Analyse post-mortem de l'incident de climatisation du 21/06/2008
- Collection et visualisation de données sur l'activité HPSS générée par dCache
 - Débit observés pour les transferts dCache ↔ HPSS, agrégés et par VO

► Autres événements



- Visite à Alice, salle du contrôle du LHC et centre de calcul du CERN le 01/07/2008
 - Détails: <http://cc.in2p3.fr/article1608.html>
- Arrêt programmé le 15/07/2008 pour intervention sur l'installation électrique et de climatisation
 - Installation des équipements temporaires nous permettant d'augmenter notre infrastructure
 - Détails: <http://cctools2.in2p3.fr/operations/wiki/doku.php?id=arrets:arret15072008>

Revue des chantiers en cours



- Intégration
 - Tests de priorisation des jobs basée sur le rôle VOMS
 - Tests de worker nodes SL4 64bits avec middleware gLite 32bits
 - Développement de l'interface BQS pour CREAM-CE
 - Consolidation des services grid: LFC, FTS, CE, SRM, système d'information
- Exploitation
 - Intégration des VO boxes à l'exploitation standard
 - Procédures d'exploitation services grid
 - Plate-forme de monitoring basée sur NAGIOS (et intégration avec SAM)
 - Séparation de la consommation CPU tier-1 et tier-2 dans la comptabilité interne
 - Système d'alarme basé sur e-mail pour des experts sélectionnés des expériences
 - Rédaction des SLAs pour les VO boxes



Aujourd'hui



- Point sur l'état de LFC et FTS
- Sujets prochaines réunions
 - Point sur les alertes NAGIOS
 - Point sur le niveau de redondance de chacun des services (LFC, FTS, dCache, CEs, VOMS, BDII, ...)
 - Présentation détaillée des outils CMS (Phedex et Squid)
- Prochaine réunion
 - Jeudi 11 septembre 13h30-15h30 Amphi
- Agendas de toutes les réunions
 - <http://indico.in2p3.fr/categoryDisplay.py?categId=102>

Vos impressions

► Questions/Commentaires

