



LCG-France Tier-1 & AF

Réunion de Coordination

Fabio Hernandez
fabio@in2p3.fr

Lyon, 12 mars 2009



l r f u
cea
saclay

▶ Table des Matières



- Réunions GDB et MB de février et mars 2009
- Disponibilité, fiabilité, efficacité des sites
- Chantiers en cours
- Thème(s) du jour

- Agendas:
 - GDB: <http://indico.cern.ch/categoryDisplay.py?categId=31181>
 - MB: <http://indico.cern.ch/categoryDisplay.py?categId=666>
- Principaux sujets traités
 - Mises à jour middleware
 - CREAM
 - Tape staging

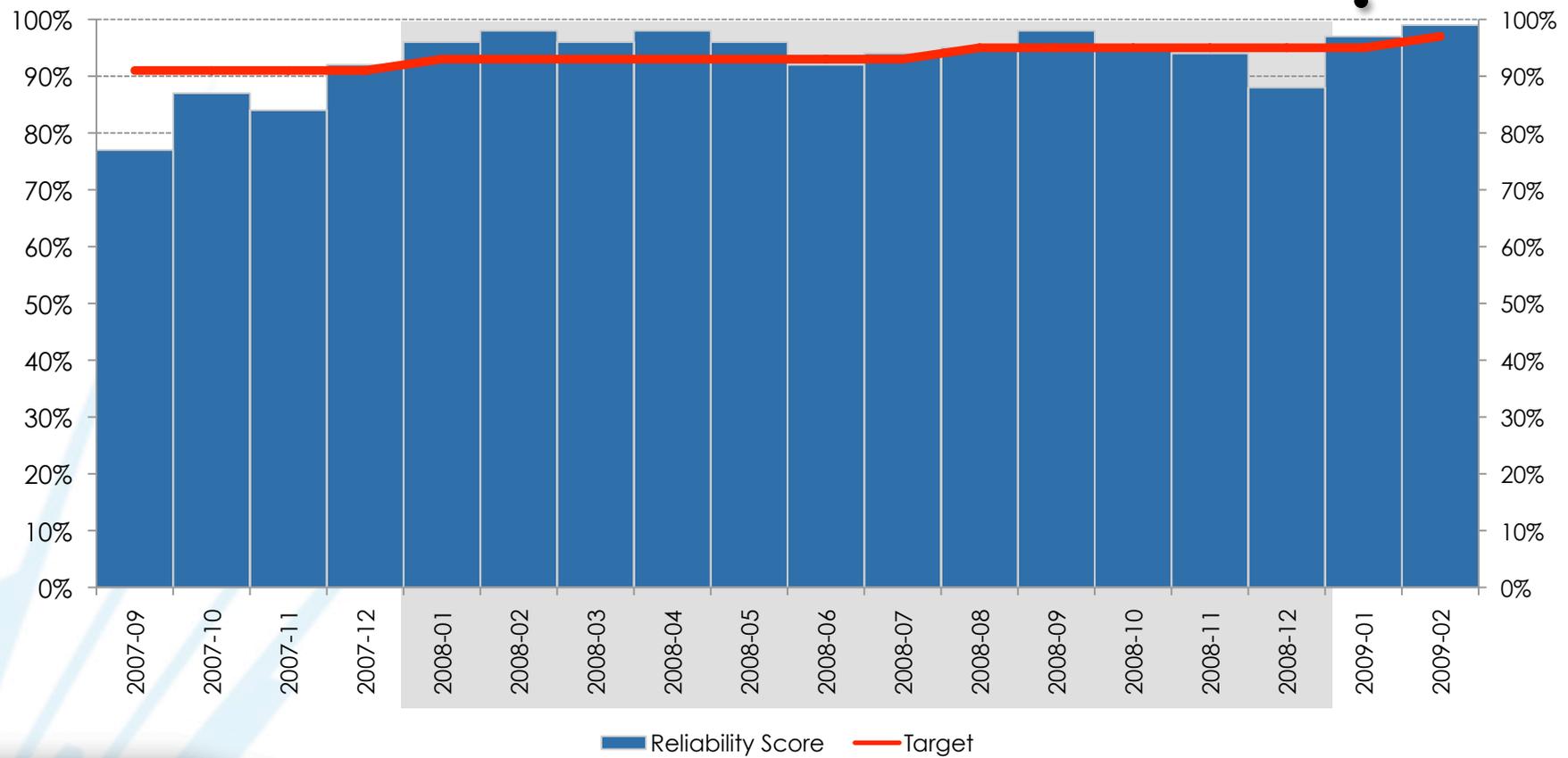
▶ Fiabilité du tier-1



Fiabilité:
Jan'09: 97%
Fév'09: 99%

LCG-France tier-1 - Reliability

(VO OPS, last 18 months)

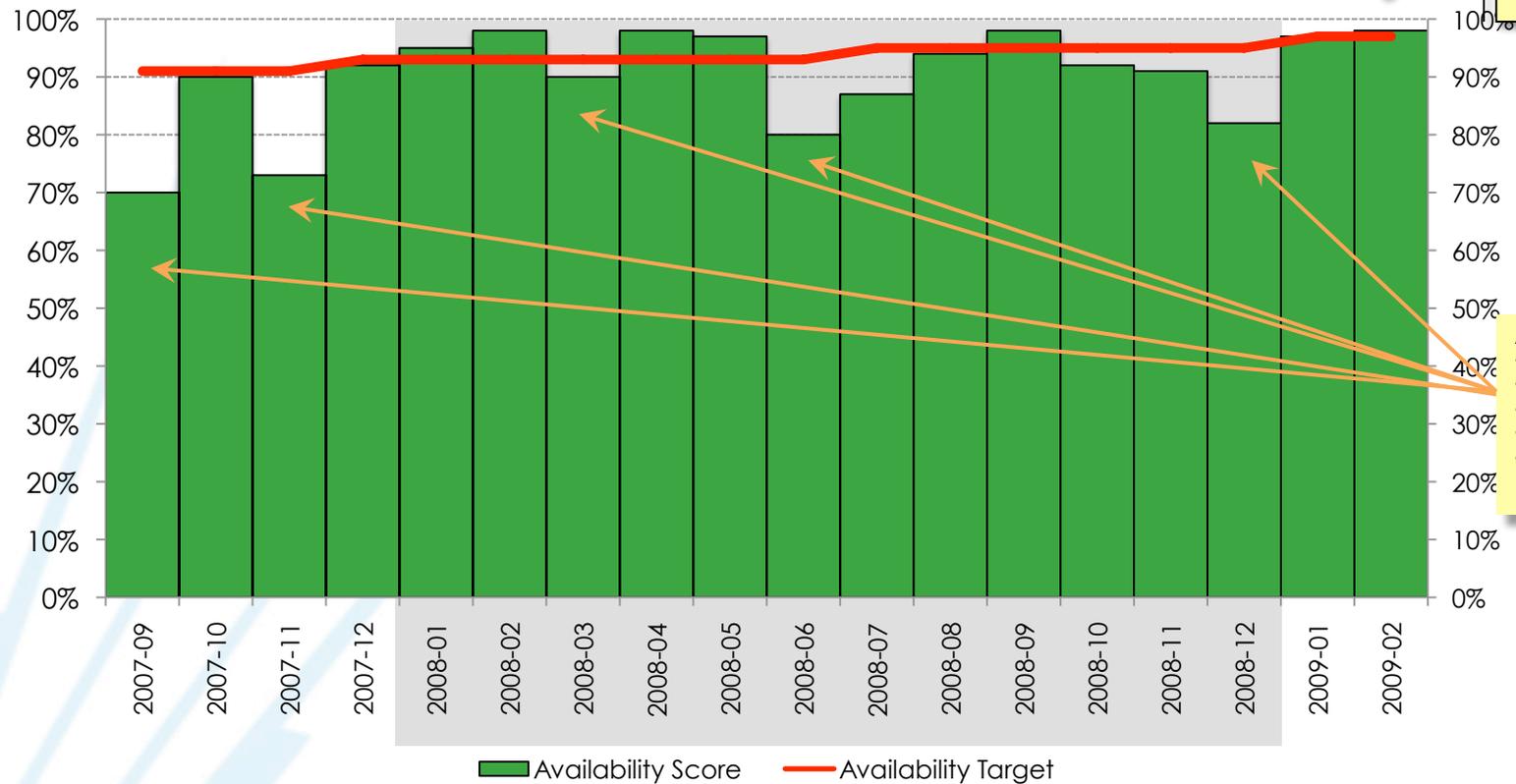


Source: [WLCG Reliability Reports](#)

▶ Disponibilité du tier-1



LCG-France tier-1 - Monthly availability score
(VO OPS, last 18 months)

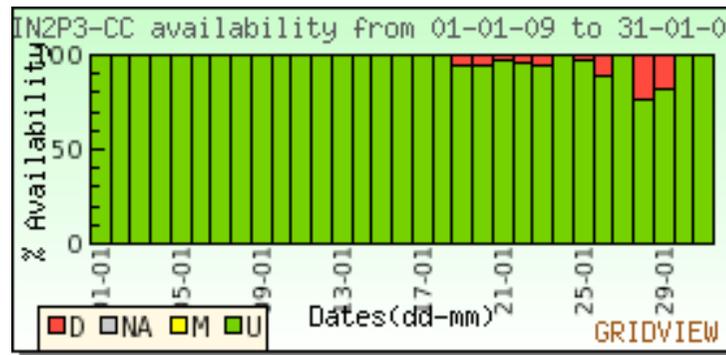


Disponibilité
jan'09: 97%
fév'09: 98%

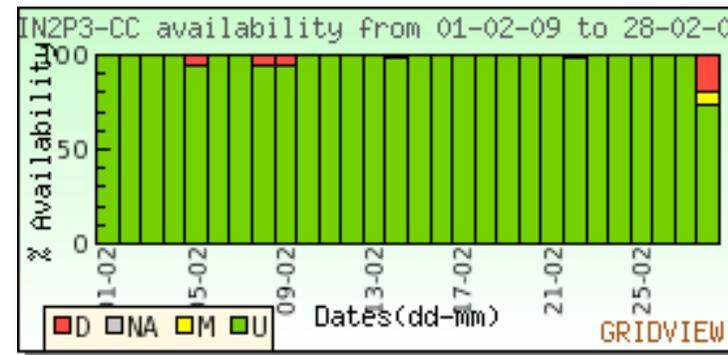
Arrêts programmés:
• 18/09/2007
• 03/11/2007
• 11/03/2008
• 03/06/2008
• 02/12/2008

Source: [WLCG Reliability Reports](#)

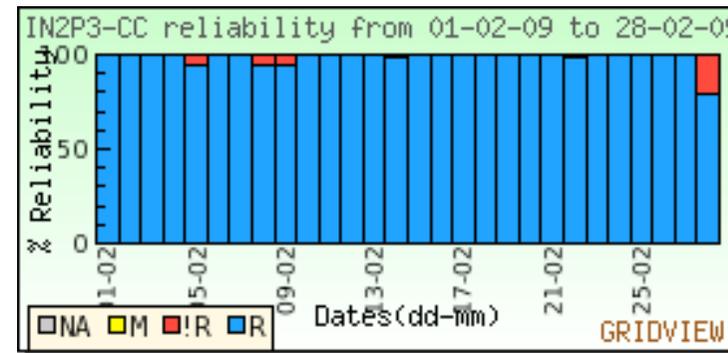
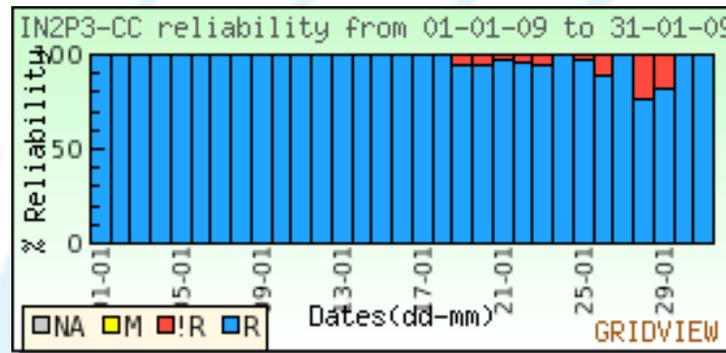
▶ Disponibilité & Fiabilité tier-1



Jan 2009



Fév 2009

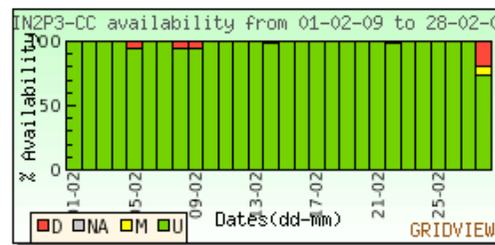


▶ Disponibilité tier-1



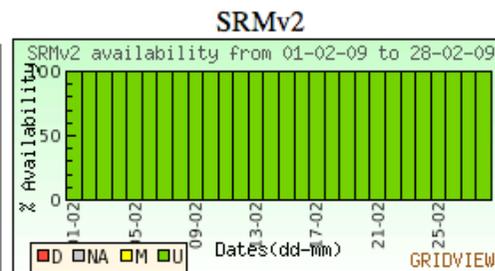
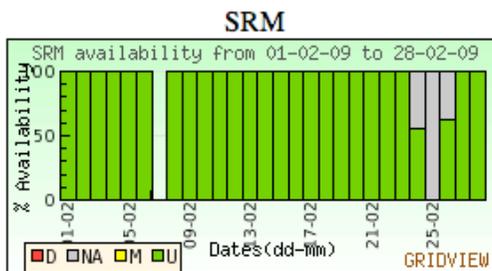
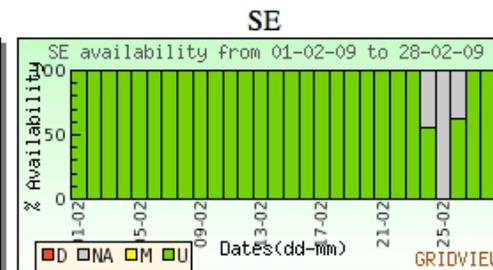
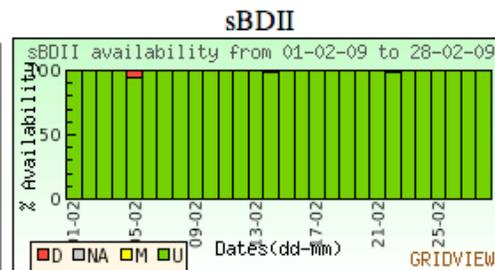
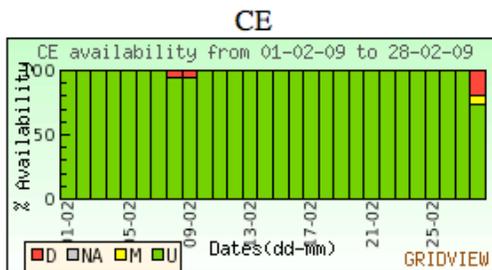
- Février 2009

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



Score: 98%

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



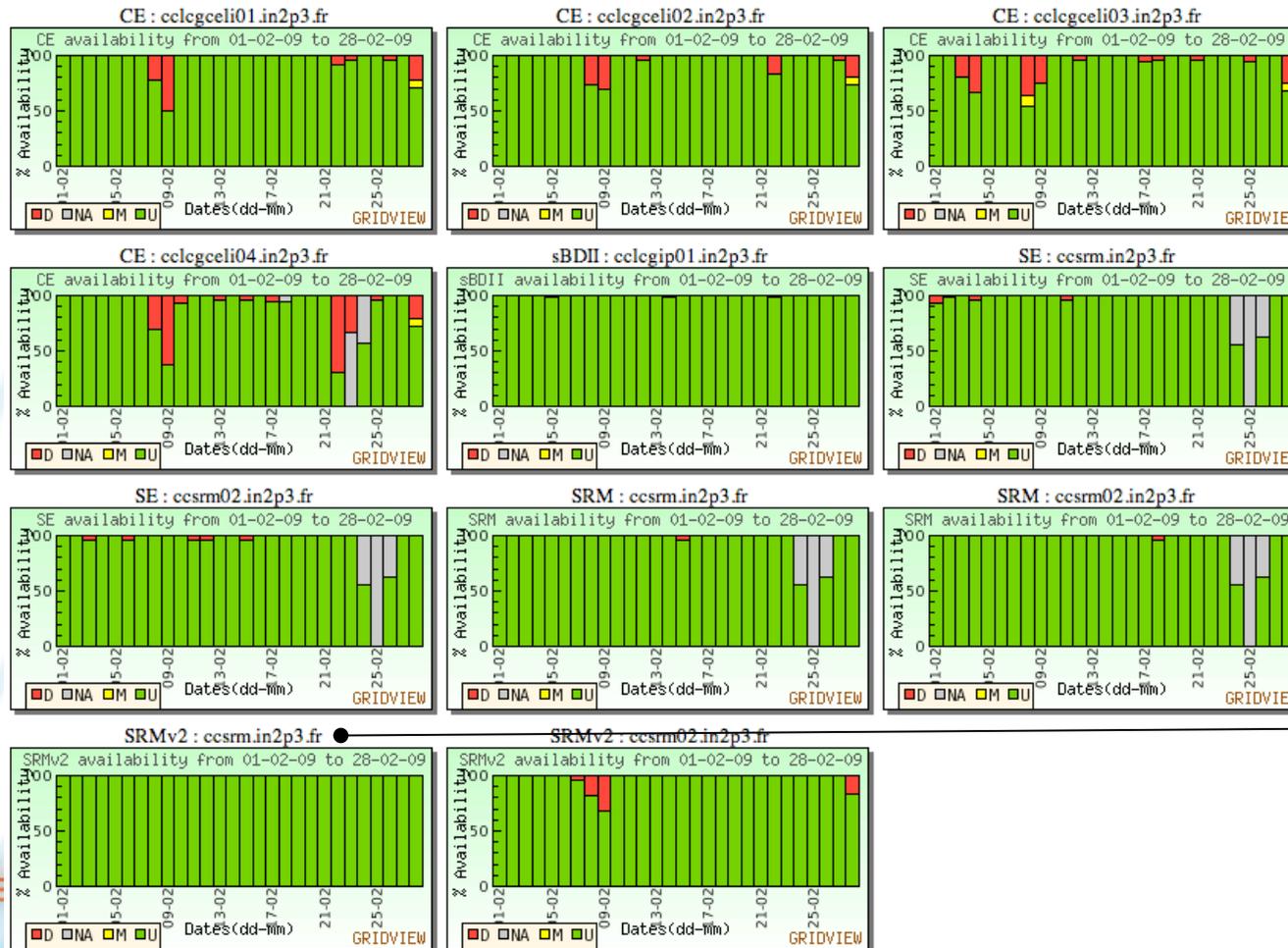
Source: <http://gridview.cern.ch>

▶ Disponibilité tier-1 (suite)



- Février 2009 (suite)

Site Services



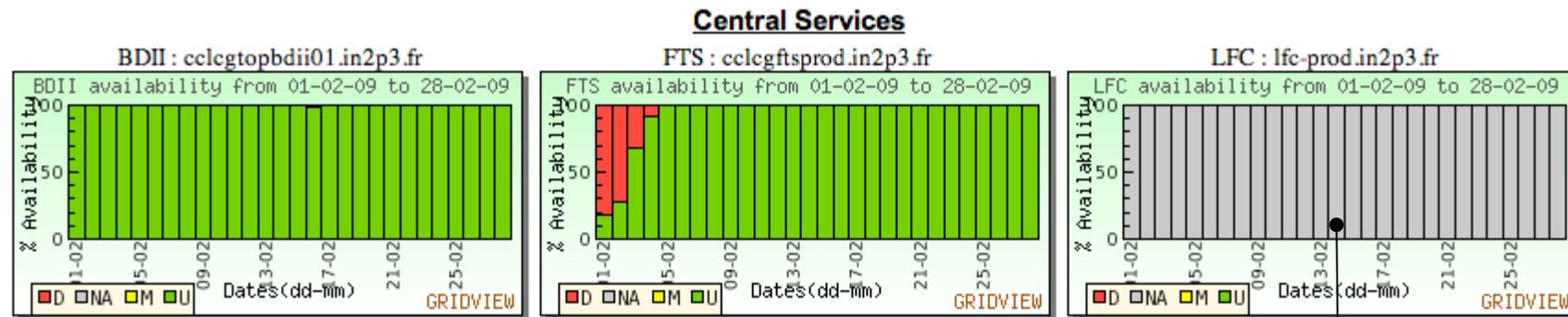
Remarable stabilité de SRM/dCache

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

▶ Disponibilité tier-1 (suite)



- Février 2009 (suite)



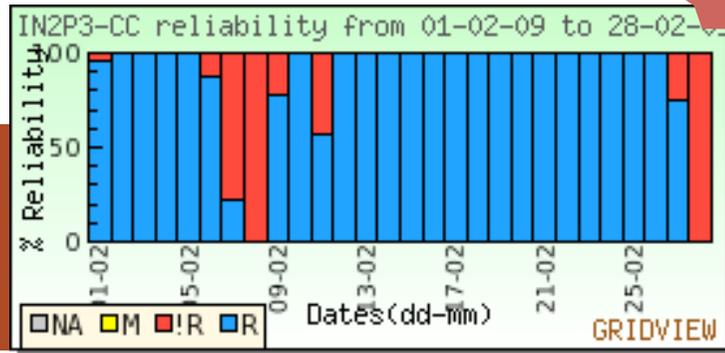
Suite de tests
LFC non
fonctionnelle
depuis
15/12/2008



Fiabilité vue par les VOs: Fév. 2009

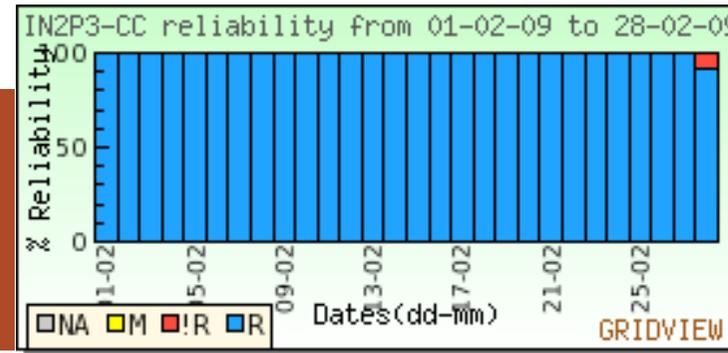


ALICE



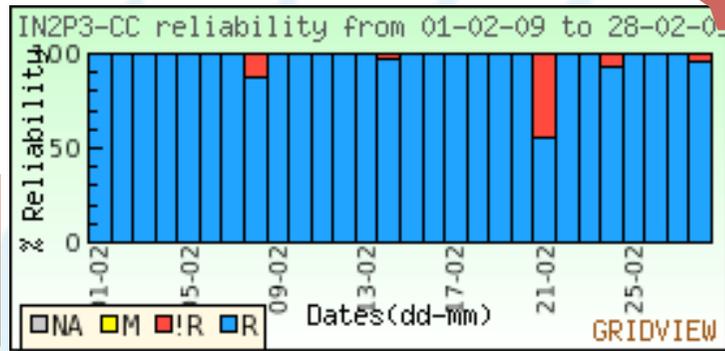
90%

ATLAS



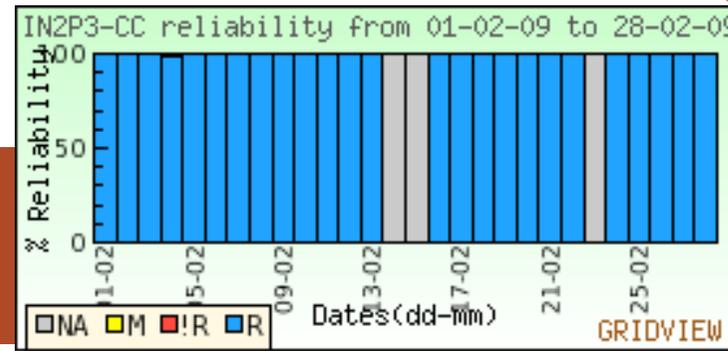
100%

CMS



97%

LHCb



100%

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

EGEE France – disponibilité & fiabilité



Jan 09

Region	Avail-ability	Reli-ability
CERN	99 %	99 %
France	97 %	97 %
UKI	95 %	96 %
Russia	91 %	92 %
AsiaPacific	90 %	92 %
SouthEasternEurope	90 %	93 %
CentralEurope	88 %	97 %
NorthernEurope	87 %	88 %
Italy	83 %	88 %
SouthWesternEurope	82 %	89 %
GermanySwitzerland	80 %	81 %

Fév 09

Region	Avail-ability	Reli-ability
CERN	98 %	99 %
France	97 %	97 %
AsiaPacific	96 %	98 %
UKI	95 %	94 %
Italy	93 %	96 %
SouthEasternEurope	88 %	89 %
NorthernEurope	83 %	84 %
SouthWesternEurope	80 %	94 %
Russia	77 %	77 %
GermanySwitzerland	74 %	80 %
CentralEurope	73 %	95 %

Source: <https://edms.cern.ch/document/963325>

▶ Mass storage performance



Reprocessing Target Rates

Tier-1	ATLAS Share %	Rate to Tape MB/s	Reprocessing Rate MB/s
BNL	25	47	465
IN2P3	15	28	279
SARA	15	28	279
RAL	10	19	186
FZK	10	19	186
PIC	5	9	93
TRIUMF	5	9	93
CNAF	5	9	93
ASGC	5	9	93
NDGF	5	9	93

ATLAS

The Story So Far

Tier-1	Target Rate MB/s	Measured Rate MB/s	Notes
BNL	465	-	
IN2P3	279	-	We are worried
SARA	279	150	Bottleneck understood – waiting for more DMF hardware
RAL	186	360	1 file MIA
FZK	186	40	Suffering from dCache pnfs overload
PIC	93	200	Some 10s of files MIA
TRIUMF	93	300	
CNAF	93	4	Much worse than last year – experts investigating
ASGC	93	570	
NDGF	93	40	Have applied a firmware fix to one tape library – should retest

Source: Graeme Stewart, [GDB, 11/02/2009](#)

▶ SCAS & glexec stress tests



- Latests results (10/03/2009)
 - Test duration: 4 days
 - Conditions: 1 SCAS server, 10 worker nodes (all in virtual machines), set of 100 proxies
 - Rate achieved: 10 Hz (!)
 - Total requests: 3.3M
 - Total errors: 1
- Details:
 - <https://twiki.cern.ch/twiki/bin/view/EGEE/SCASstressresults>

 **Deployment**
Enabling Grids for E-science

- **FZK (ready since 19th-Feb)**
 - Dedicated CE test-mw-2-fzk.gridka.de accounted for FZK-LCG2
 - CE-Status is set to SCASPilot
 - currently published in pre-production. SITE_NAME=FZK-PPS
 - can be used by
 - lhcb:/lhcb/Role=pilot (queue lhcbXXL)
 - atlas:/atlas/Role=production + atlas:/atlas/usatlas/Role=pilot (queue atlasXXL)
 - cms:/cms/Role=production (queue cmsXL)
 - alice:/alice/Role=pilot (queue aliceXL)
- **Nikhef EL-Prod (ready since 27th-Feb)**
 - All the WNs are gLExec-enabled and accessible by all prod CEs
 - multiple SCAS endpoints for fault tolerance. All production CEs are configured to use WNs with gLExec.
 - can be used by:
 - lhcb:/lhcb/Role=pilot
 - atlas:/atlas/Role=production
- **Both sites run (#2829)**
- **IN2P3 ready to step-in after the 15th Mar**

EGEE-III INFSO-RI-222667 GDB - 1 Mar 09 - CERN 13

Source: Antonio Retico, [GDB, 11/03/2009](#)

▶ Reporting installed capacity



Outline

- **The Installed Capacity document (v. 1.8-7) has been approved by the WLCG Management Board on the 3rd of February 2009.**
 - https://twiki.cern.ch/twiki/pub/LCG/WLCGCommonComputingReadinessChallenges/WLCG_GlueSchemaUsage-1.8.pdf
 - **V 2.0 will be published soon with some more operational clarifications and OSG examples**
- **We are preparing a detailed operational plan that includes EGEE and OSG sites.**
 - **To tentatively start deployment during March 2009.**
- **In what follows we give a short summary of the deployment plan for both EGEE and OSG sites and some operational details concerning EGEE sites.**



WLCG GDB, CERN 11 February 2009

2

Vérification nécessaire pour nous assurer que l'information que nous publions dans le système d'information corresponde aux conventions et est cohérente par rapport à la réalité

Goals

- **Publishing information in order to provide the WLCG management with a view of the total installed capacity and resource usage by VO at sites. Focus on:**
 - **Static information that produces monthly reports**
 - Publishing information to provide the WLCG management with a view of the *total installed capacity at the sites*.
 - Publishing information to provide the WLCG and VO management with a view of the *resource assignment per VO at the sites*.
 - **Dynamic/ongoing information**
 - Publishing information to allow VO operators to *monitor the VO usage of the resources at a site*
 - Allowing clients (consumers of the information published) to *select the correct resources and their characteristics*.



WLCG GDB, CERN 11 March 2009

3

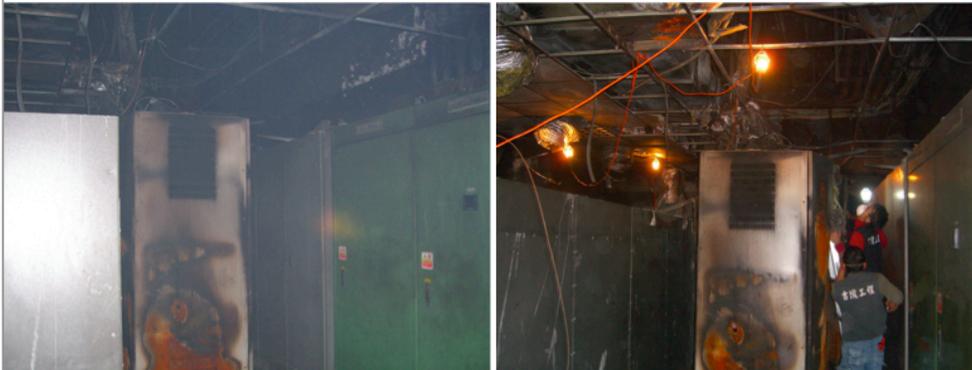
Source: Steve Traylen, [GDB, 11/03/2009](#)

▶ Fire incident at ASGC tier-1



Event Scene Investigation

- 16:53 on 25 February (Taipei Time), fire lasted less than 3 hours:
 - SMS sent out "UPS power lost" and power generator did not start
 - Smell and visible of smoke, and then fire at 2nd column of battery rack. Electric arc also viewed in the middle battery racks.
 - Fire spot (V shape) on the far-end wall, and near-end battery rack was also burned out → fire spread to left and right within the middle battery racks, and up to the ceiling
 - But no fire burning evidence at racks in the other two ends



Burnt-out UPS Battery



Source: Simon Lin, [GDB, 11/03/2009](#)

CREAM deployment



Current Site Situation (10/03/09)

Site	queues	Status of the queues	2 nd VOBOX	VOBOX with clients	General Status
FZK	4	OK	YES	YES	IN PRODUCTION
KOLKATA	2	OK	YES	YES	IN PRODUCTION
ATHENS	1	OK	NO	NO	NOT READY
KISTI	1	OK	YES	YES	IN PRODUCTION
GSI	1	OK	NO	YES	IN * PRODUCTION
IHEP	1	NOT OK	NO	NO	NOT READY
RAL	1	OK	NO	YES	IN PRODUCTION
CNAF	1	NOT OK	YES	YES	NOT READY
CERN	2	Provided yesterday	YES	YES	READY

* Performing special production

11/03/09 ALICE Experiences with CREAM

3

Source: Patricia Mendez, [GDB, 11/03/2009](#)

Status of criteria to be met before transition can start

Note:

This table deals with the criteria for *starting the transitioning* from the LCG CE to the CREAM CE. It does **NOT** deal with the initial deployment of the CREAM CE into the production service: the CREAM CE is *already* available in the production middleware repositories and all sites are *strongly* encouraged to deploy a CREAM CE in parallel to their LCG CE.

You can also [click here](#) to see the status of those features that weren't considered showstoppers for the transition, but which are nevertheless still considered as **very** important for the CREAM CE.

	CRITERIA	COMMENTS	STATUS
A	The CREAM CE should provide at least equivalent functionality and performance as the LCG CE (excluding the ability for users to fork processes directly on the CE)	The PPS Pilot of the CREAM CE has not shown any problems with missing functionality so far. The performance of the CREAM CE is still being tested. <i>NOTE: It is not possible for users to directly fork processes on the CREAM CE.</i>	
B	Condor-G submission to CREAM must be available in production with no significant bugs.	The basic functionality has been tested by US-CMS. They are due to start more in-depth testing now. Full details available here: http://hepuser.ucsd.edu/twiki2/bin/view/HEPProjects/CMS-Cream This functionality will be released in Condor => 7.3.0. <i>Waiting for a reply to find out when Condor 7.3.0 will be available.</i>	

Source: <https://twiki.cern.ch/twiki/bin/view/LCG/LCGCEtoCREAMCETransition>

I

The process for integrating other batch systems must be fully documented.

The Blah Guide has been checked over by the BQS experts at IN2P3 and is considered to be "very useful".



▶ Chantiers récents et en cours



- Amélioration du BQS Job Manager
 - Meilleure gestion des erreurs
 - Bénéfice observé en particulier pour les jobs ATLAS soumis via PANDA
 - Déploiement finalisé sur tous les CEs (T1 & T2)?
- Mise à jour du middleware des CEs
- Participation aux tests ATLAS d'accès aux données de conditions gérées localement par Oracle
- Campagne de remplacement matériel et de mise à jour logicielle des VO boxes (toutes les VOs LHC)
- Security challenge (26/02/2009)
- Début de migration vers SL5 64 bits (avec compatibilité 32 bits) pour les worker nodes
 - Philippe Olivero coordonne l'opération

▶ Chantiers récents et en cours (cont.)



Exploitation		
exploitation Logbook de l'exploitation	1219	mer 11 mar 2009 09:21:42 CET
astreinte Logbook de l'astreinte	177	dim 08 mar 2009 13:44:06 CET
evenements Evénements generaux liés à l'Opération	38	mer 11 mar 2009 09:20:42 CET
problemes-VO Logbook des problemes recurrents des VOs LHC	48	mar 10 mar 2009 16:58:11 CET
Services		
HPSS Logbook d'HPSS	181	mer 11 mar 2009 10:22:48 CET
dCache Logbook de dCache	64	mer 11 mar 2009 09:37:11 CET
logon Logbook de logon@cc	2217	mer 11 mar 2009 07:45:59 CET
Suivi des opérations		
operations-HPSS Logbook des opérations sur HPSS	9	sam 28 fév 2009 22:04:27 CET
operations-dCache Logbook des opérations sur dCache	43	mer 11 mar 2009 10:58:27 CET
operations-LFC Logbook des opérations sur LFC	3	mer 04 fév 2009 14:44:09 CET
operations-FTS Logbook des opérations sur FTS	17	mer 11 mar 2009 10:29:03 CET
operations-CE Logbook des opérations sur les CEs	8	mar 24 fév 2009 18:40:30 CET
operations-SI Logbook des opérations sur le SI	0	-
operations-VOBOX Logbook des opérations sur les VOBox	0	-
operations-NAGIOS Logbook des opérations sur NAGIOS	0	-



Logbook de suivi des opérations VO boxes, CEs, HPSS semblent ne pas être tout à fait à jour

Source: CC-IN2P3 logbooks, snapshot taken on 11/03/2009 12h, <http://cctools2.in2p3.fr/elog/>

▶ Chantiers récents et en cours (cont.)



- dCache: vérification de la cohérence du catalogue interne (semaine du 26/01/2009)
 - 1.3% de fichiers ATLAS inexistant enregistrés dans le catalogue PNFS (5.6 M de fichiers)
 - Opération faite sans perturber la production dans un temps relativement court (quelques heures)
 - Nous avons convenu que cette vérification sera faite régulièrement (~1 fois par mois) et pour les fichiers des 4 expériences LHC
- dCache: mise à jour matérielle des serveurs centraux et répartition des services sur 4 machines
- Tape staging: réunion interne dédiée à ce sujet le 30/01/2009
 - <http://indico.in2p3.fr/conferenceDisplay.py?confId=1676>
 - Voir présentation de Jonathan aujourd'hui

▶ Réunions de coordination avec les expériences



- CMS (03/02/2009)
 - <http://indico.in2p3.fr/conferenceDisplay.py?confId=1694>
 - Séparation des portes d'accès aux zones de stockage tier-1 et tier-2 dans dCache, tout en évitant la duplication des données
- ATLAS
 - Partie de la réunion mensuelle Computing Atlas France (CAF) consacrée aux sujets liés au CC-IN2P3
 - An plus des expertes ATLAS computing, participation des experts grid, opérations, stockage
 - Agendas: <http://indico.cern.ch/categoryDisplay.py?categId=5131>



Réunions de coordination avec les expériences (suite)



- Alice (04/03/2009)
 - <http://indico.in2p3.fr/conferenceDisplay.py?confId=1773>
 - Possibilités d'utilisation de serveurs xrootd en tant que storage elements
 - Déploiement et tests à programmer

▶ Axes de travail prioritaires



- Tier-1
 - Améliorer l'efficacité de la copie bande → disque, via l'ordonnancement des requêtes passées à HPSS par dCache
- Séparation claire des données tier-1 et tier-2
 - Portes d'accès différentes pour chacun des sites, sans pour autant dupliquer les données
- Déploiement d'un prototype de ferme d'analyse interactif basée sur PROOF
- Finaliser le mécanisme d'installation du logiciel des expériences dans la zone AFS
 - Utilisation de la réplication de volumes AFS pour l'équilibrage de charge
 - Documentation et généralisation pour les 4 expériences LHC
- Monitoring !!!

▶ Ressources LHC pour 2009



Expérience	CPU [K SI2000]	Disk [TB]	MSS [TB]
ALICE	2593	824	992
ATLAS	6141	2427	2204
CMS	3229	1352	1650
LHCb	3049	820	829

Source: [LCG-France Revised Resource Planning 2007-2013](#)

Ces engagements de ressources comprennent ceux au niveau du tier-1, tier-2 et tier-3, qu'elles soient déclarées ou non à WLCG.

▶ Evénements passés et à venir



- Comité de Pilotage LCG-France
 - Lyon, 13/02/2009
 - <http://indico.in2p3.fr/conferenceDisplay.py?confId=1672>
 - Voir ma présentation attachée à l'agenda de la réunion de ce jour
- WLCG Collaboration Workshop
 - Prague, 21-22/03/2009
 - <http://indico.cern.ch/conferenceDisplay.py?confId=16861>
- CHEP 2009
 - Prague, 23-27/03/2009
 - <http://www.particle.cz/conferences/chep2009/>
- Hepix Spring 2009
 - Umeå (Suède), 25-29/05/2009

▶ Aujourd'hui et à venir



- Aujourd'hui
 - Synthèse habituelle des opérations LHC
 - Avancement du déploiement de la surveillance par NAGIOS
 - Avancement du développement de l'ordonnanceur de requêtes de tape staging
 - Résultats préliminaire de l'analyse des résultats des tests SAM sur SRM/dCache
- Prochaine réunion
 - 16 avril, 13h30-15h30, salle 202
 - NOTE: changement de date, initialement prévue le 9 avril
- Toutes les réunions à venir
 - <http://indico.in2p3.fr/categoryDisplay.py?categId=102>



Questions/Commentaires

