



LCG-France Tier-1 & AF

Réunion de Coordination

Fabio Hernandez

fabio@in2p3.fr

Lyon, 11 septembre 2008

dapnia
cea
saclay

CNRS
CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE

► Table des Matières



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

► Démarrage du LHC



- LHC first beam

- <http://lhc-first-beam.web.cern.ch/lhc-first-beam/>

 **LHC First Beam** Contact | CERN Home

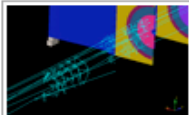


News on LHC start-up



**LHC First Beam on
10 September 2008**



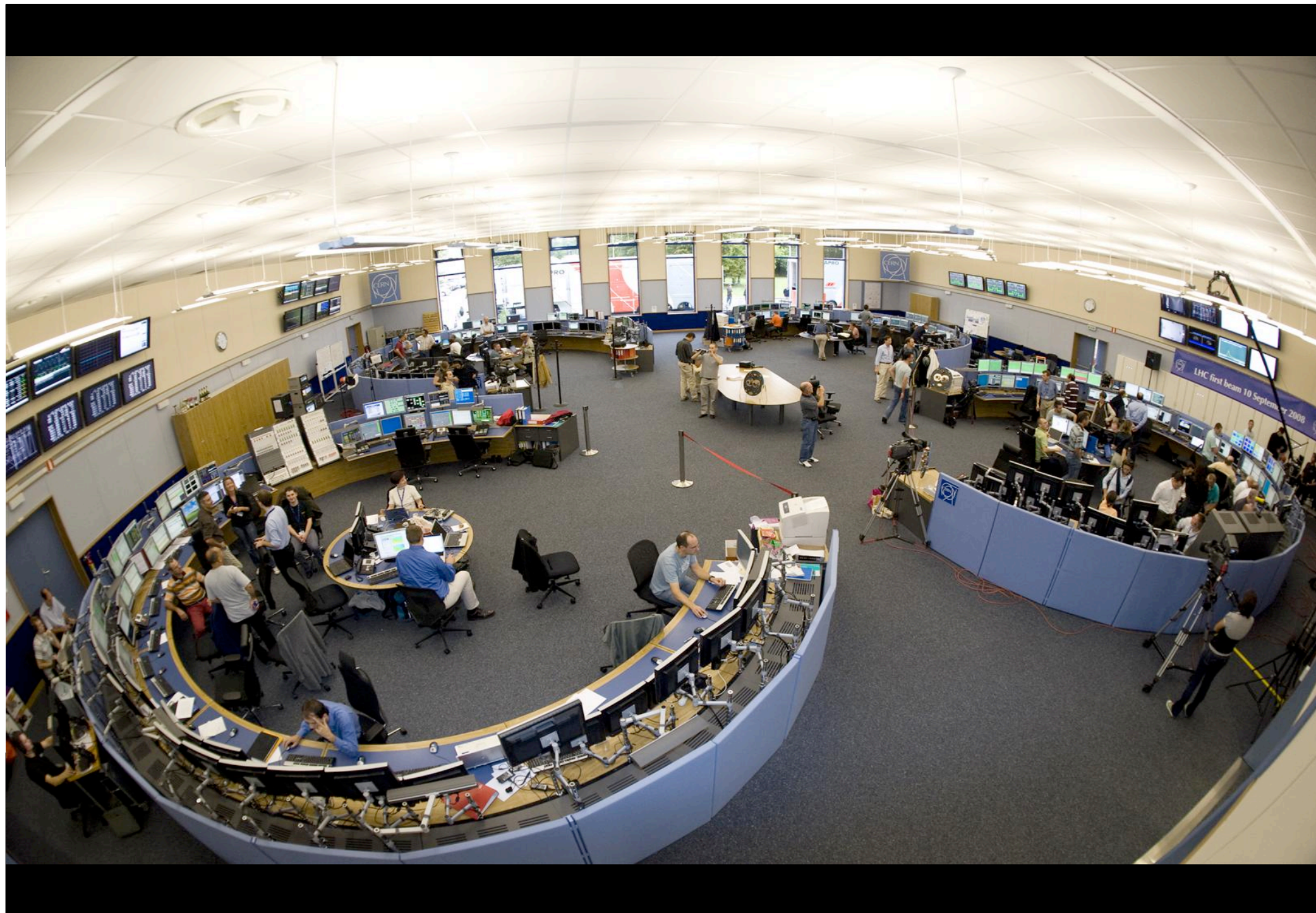
First beam in the LHC - accelerating science 10.09.2008
Geneva, 10 September 2008. The first beam in the Large Hadron Collider at CERN was successfully steered around the full 27 kilometres of the world's most powerful particle accelerator at 10h28 this morning. >>


Final LHC Synchronization Test a Success 25.08.2008
CERN has today announced the success of the second and final test of the Large Hadron Collider's beam synchronization systems which will allow the LHC operations team to inject the first beam into the LHC. >>

General information

**Accreditations closed.
No accreditations will be granted on 10 September**

- Planning of the day [.pdf]
- Live Satellite Broadcast info
- Live webcast
- Live on Evo
- Live on Groovy Gecko
- Live on Livestation
- Live on tsr.ch
- Live on LCI.fr
- What will happen on the day?
- Technical information for media
- Practical information





► Autres événements



- Grid Fest au CERN
 - 3 octobre 2008
- Évènement pour faire parler de la grille pour le LHC

Source: <http://lcg.web.cern.ch/LCG/lhcgridfest>

A banner for the LHC Grid Fest. On the left, there is a sidebar with the LHC Grid Fest logo and navigation links under 'The Event' and 'Press+Media'. The main part of the banner features a large, glowing orange ring of sparks. Below the ring, the text 'LHC Grid Fest' is displayed. At the bottom, there are two paragraphs of text and a small LCG logo.

LHC Grid Fest

When the Large Hadron Collider comes into operation in September 2008, it will begin to produce an expected 15 million gigabytes of data every year, enough information to create a 21-kilometre-high stack of CDs annually.

On 3rd October, the Worldwide Large Hadron Collider Computing Grid consortium announce the readiness of the Worldwide LHC Computing Grid (WLCG), an e-infrastructure conceived and designed to support this data challenge, and with it the research of more than 9000 physicists around the globe.

"The Worldwide LHC Computing Grid is generating the technology for tomorrow's science needs. We are witnessing a unique collaboration on an international scale, with vast potential for accelerating discoveries in physics and other fields of science." Ian Bird, WLCG project leader.

The LCG logo, consisting of a grid of colored squares (yellow, blue, red, green) with the letters 'LCG' in the center.



- Agendas:
 - GDB: <http://indico.cern.ch/categoryDisplay.py?categId=31181>
 - MB: <http://indico.cern.ch/categoryDisplay.py?categId=666>
- Principaux sujets traités
 - Instabilités des systèmes de stockage
 - Service continu
 - Middleware
 - Tickets d'alerte

Stockage



LHCb

■ Site configuration and stability

- In general difficult. More problems with dCache than with CASTOR.

- CASTOR at CNAF: Failed to perform gfal_ls: [SE][Ls] http://srm-v2.cr.cnaf.infn.it:8443/srm/managerv2: Too many threads busy with Castor at the moment Unknown error 18446744073709551615
- CASTOR at RAL: ORACLE problem.
- NIKHEF: we had so many issue that are all documented by several GGUS tickets.
 - 1. PinManager issue https://gus.fzk.de/pages/ticket_details.php?ticket=39549,
 - 2. SRM dead: https://gus.fzk.de/pages/ticket_details.php?ticket=39996
 - 3. (POLLIN) Control line creation failure https://gus.fzk.de/ws/ticket_info.php?ticket=39953
 - 4. log-gt timing out (because LCG_GFAL_INFOSYS was pointing to local dismissed top bdii kraal.nikhef.nl:2170)
 - 5. File ONLINE with not tURLS: https://gus.fzk.de/pages/ticket_details.php?ticket=40001
 - 6. gfal_turlfromsurls: [SE][StatusOfGetRequest] http://srm.grid.sara.nl:8443/srm/managerv2: org.xml.sax.SAXParseException: An invalid XML character (Unicode: 0x10) was found in the element content of the document. Unknown error 4294967295 **ALARM** ticket https://gus.fzk.de/pages/ticket_details.php?ticket=40170

■ IN2P3:

- 1. Disk files NEARLINE at IN2P3: files supposed to be permanent on disk at IN2P3 are only NEARLINE on the storage. (dcache server down)
- 2. gsidcap fails using shipped version of dcache clients not containing the fix for the concurrent gsidcap connections
- 3. Jobs fail to store files in the remote storage at IN2P3 .

WLCG Grid Deployment Board, CERN 9 September 2008

6

■ IN2P3:

- 4. https://gus.fzk.de/pages/ticket_details.php?ticket=36999 , CGI-gSOAP:Error reading token data: Success ***

- GRIDKA :SAM tests failing https://gus.fzk.de/pages/ticket_details.php?ticket=39551
gfal_ls:Error reading tokendata: Success issue:
https://gus.fzk.de/pages/ticket_details.php?ticket=39995

- PIC: gfal_ls:Error reading tokendata: Success issue:
<http://blogbook.cern.ch/Operations/416> (no GGUS open because like for GridKA this is not reproducible by changing the client versions)



▶ Service continu (a.k.a. 24x7)



Issues / discussion

- How do Tier 1s provide after hours coverage (“on-call”)?
- Has this worked in the last few months – if not how will this be addressed?
- How do (and should) we measure response to issues?
 - This is an MoU parameter
 - If sites reliable and experiments happy – no issue But ...
- How will we manage scheduled downtimes?
 - Do we need regular scheduled maintenance?
 - Some sites propose this, others do not always need it
 - Can’t have several sites scheduled down on same day



Alertes



Enabling Grids for E-science

ALARM TICKETS

- **Ticket processing and routing**

- Tier-1 site notification

- The site is notified directly by sending an ALARM mail the site alarm mail address

- Alarm mail addresses are documented at <https://twiki.cern.ch/twiki/bin/view/LCG/OperationsAlarmsPage>

- As the site alarm mail address is now implemented in GOC DB please add the alarm mail address of your site in GOC DB!

- ALARM mails are signed with the GGUS certificate

- ALARM mails sent to tier-1 sites contain a link to the GGUS Alarm ticket

- Ticket routing

- Alarm tickets are routed to the ROC directly, bypassing the TPM

- The routing is based on the tier-1 site name specified in the ticket

Mécanisme de réception et routage en interne au site mis en place par Marc Haussard

Alertes (suite)



Enabling Grids for E-science

DIFFERENCES BETWEEN ALARM AND TEAM TICKETS

	ALARM Tickets	TEAM Tickets
Ticket processing	Alarming tier-1 site admins at any time (365*24*7)	Notifying tier-1 site admins; actions will only be taken during office hours
Possible submitters	„Alarmers“; these are 3-4 experts of each LHC VO	„Team“ members; these are a number of VO members with expertise
GGUS mail	ALARM mails are signed with the GGUS certificate	Notification mails are not signed

EGEE-III INFSO-RI-222667

Guenter Grein

11

Middleware



WLCG Middleware Baseline

- How should this list be interpreted?
 - As a set of recommended versions?
 - As a set of minimum versions?
- How is this list maintained?
 - WLCG Daily ops meetings
 - This is basically still at the CCRC08 May level
 - Except CASTOR and dCache
- How should 'compliance' be monitored?
 - Service version info providers will help

Source: Oliver Keeble, [GDB, 09/09/2008](#)

Base versions of services and client tools for WLCG operations

The following list is intended to be continually monitored and updated in the daily WLCG Operations Meetings. It does **not** necessarily reflect the latest versions of packages available in the gLite repository.

This list has not yet been validated

Service	Version	Availability	Notes
FTS	2.0	gLite 3.0 / SL3	
FTM	2.0	gLite 3.1 / SL4	Distributed as glite-FTM-3.1.5-0
dCache	1.8.0-15-p11	DESY	
CASTOR Core Services/Stager	2.1.7-16		
CASTOR SRM 2.2 Server	1.3-28 on SLC3 and 2.7-3 on SLC4		
DPM server	1.6.11	gLite 3.1 / SL4	Distributed as glite-SE_dpm_mysql-3.1.15-0
DPM disk	1.6.11	gLite 3.1 / SL4	Distributed as glite-SE_dpm_disk-3.1.15-0
LFC/mysql	1.6.8-1	gLite 3.1 / SL4	Distributed as glite-LFC_mysql-3.1.9-0
LFC/oracle	1.6.8-1	gLite 3.1 / SL4	Distributed as glite-LFC_oracle-3.1.9-0
STORM	1.3.20	release plan	
lcg-CE	lcg-CE- 3.1.16-0	gLite 3.1 / SL4	rpm list
WN	glite-WN-3.1.14-0	gLite 3.1 / SL4	rpm list
UI	glite-UI-3.1.14-0	gLite 3.1 / SL4	rpm list
WMS	glite-WMS-3.1.2-0	gLite 3.1 / SL4	rpm list
LB	glite-LB-3.1.1-1	gLite 3.1 / SL4	rpm list
RB	3.0.19-1	gLite 3.0 / SL3	
ARC server	0.6		

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

► Middleware (suite)



Middleware roadmap

- **Clients on new 'platforms'**
 - SL5 WN 32/64, SL5 UI 32
 - SL4/SL5 Python 2.5
 - Debian4/32 WN
 - SL4/SL5 + changed compiler ???
 - which one, 3 aren't an answer.
 - gcc4.3?
- **Imminent or available updates**
 - FTS/SL4 (available)
 - Globus bugfixes (available)
 - lcg-CE - further performance improvements
 - Results of WN working group - cluster publishing

New Services



- **glexec/SCAS**
 - glexec is being verified for use with experiment frameworks in the PPS (setuid mode, **without SCAS**)
 - SCAS still in 'developer testing'
- **CREAM**
 - First release to production imminent
 - NOT as a replacement for lcg-CE
 - Issues with proxy renewal
- **WMS/ICE**
 - Patch under construction
- **Glue2**
 - OGF Public Comments are now over
 - Glue WG will incorporate these during October
 - Will be deployed in parallel as it is non backward compatible

Source: Oliver Keeble, [GDB, 09/09/2008](#)

► Middleware distribution



Alternative #1



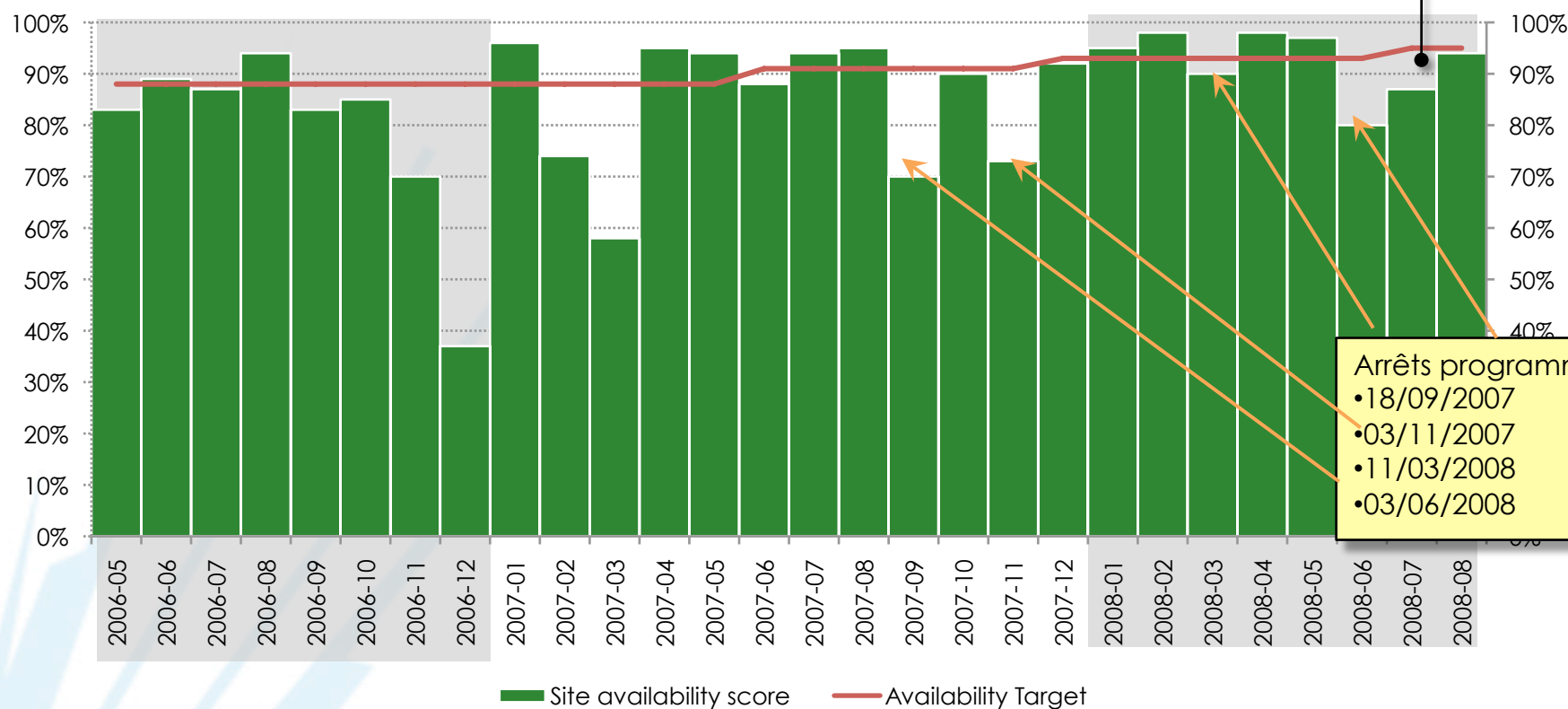
- Can this be done in 'experiment space'?
- We provide versioned tarballs via the Application Area, and these are integrated and distributed by the experiments
- Advantages
 - Uses an accepted distribution model
 - No need to use a separate VO
- Disadvantages
 - Extra effort for the experiments
 - Duplication of installations
 - Not clear that reactivity would be improved
 - The would still need the mechanism for choosing the environment

Approche retenue:
installation(facultative)
du middleware WN
dans la zone logicielle
de chaque expérience

► Disponibilité du tier-1



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



Disponibilité:
juillet 87%
Août 94%

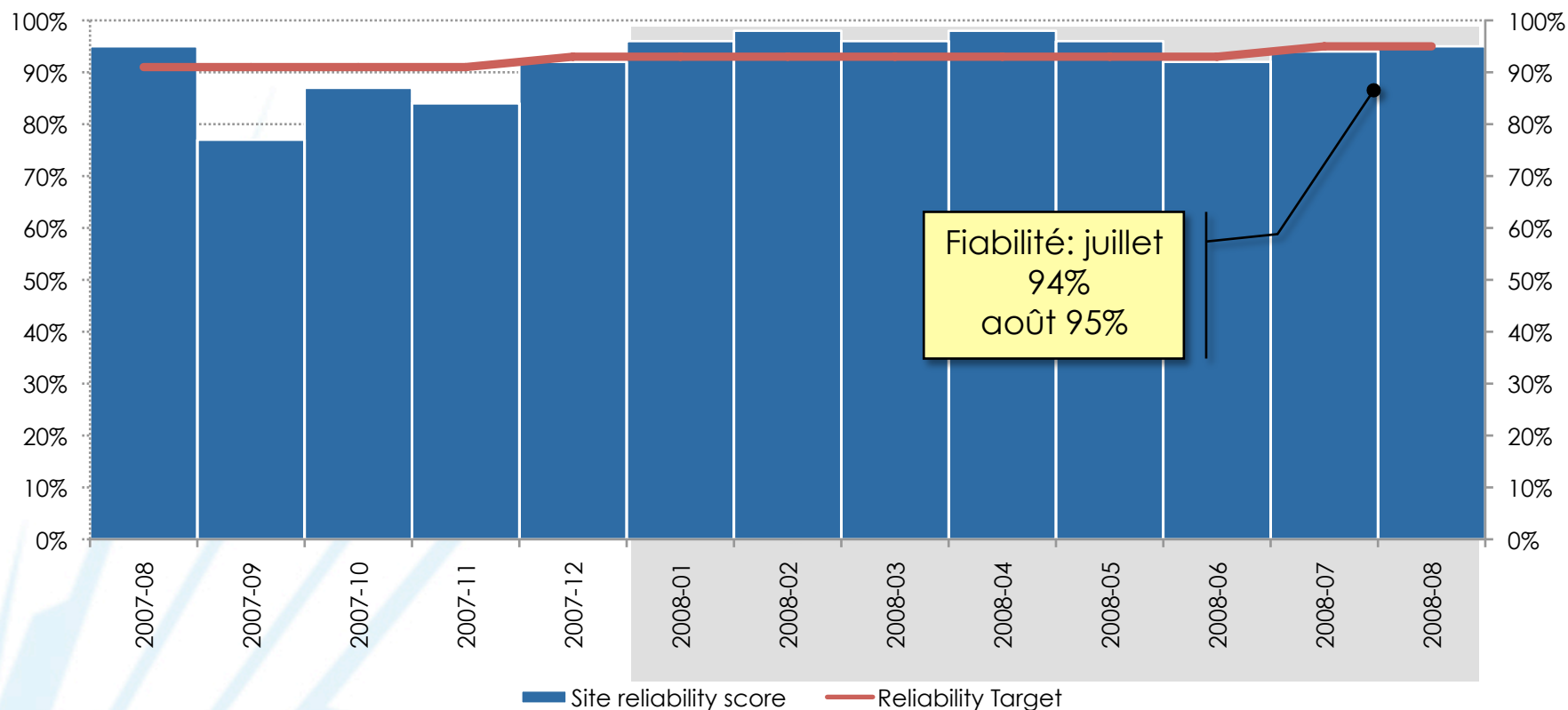
Arrêts programmés:
 • 18/09/2007
 • 03/11/2007
 • 11/03/2008
 • 03/06/2008

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

► Fiabilité du tier-1



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



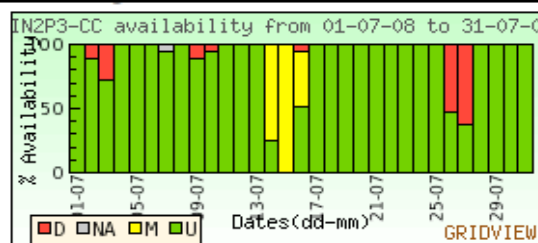
Source: [WLCG T0 & T1 Site Reliability Reports](#)

► Disponibilité des sites: tier-1 (1/8)



- Juillet 2008

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

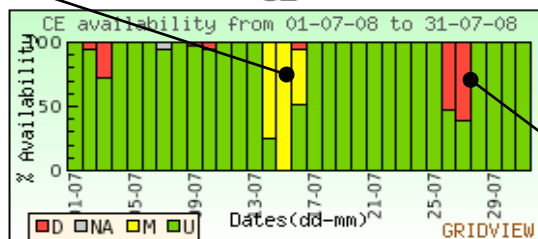


Score: 87%

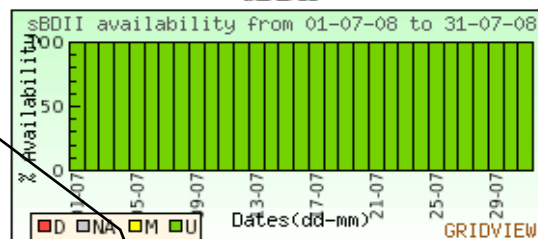
Maintenance?

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

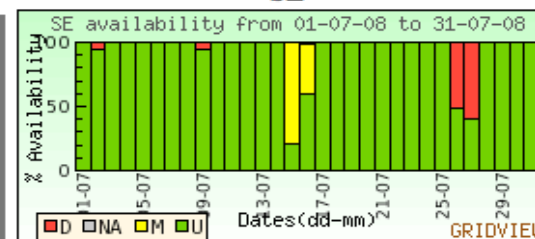
CE



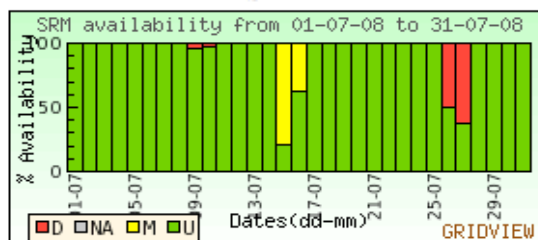
sBDII



SE



SRM



Problème de publication des informations SRM sur le site, pendant tout un week-end

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

► Disponibilité des sites: tier-1 (2/8)



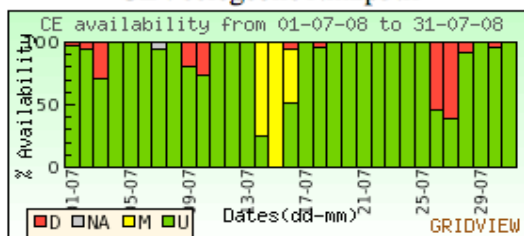
- Juillet 2008 (suite)

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

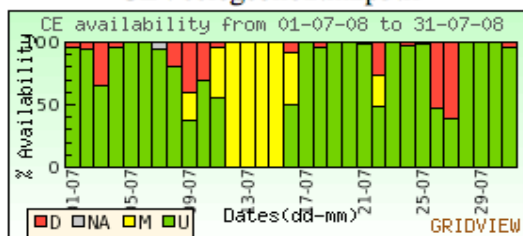
Site Services

Des instabilités sur les CEs.
Causes?

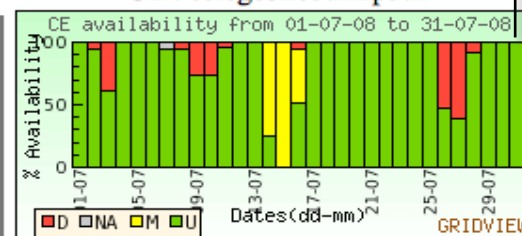
CE : cclcgceli01.in2p3.fr



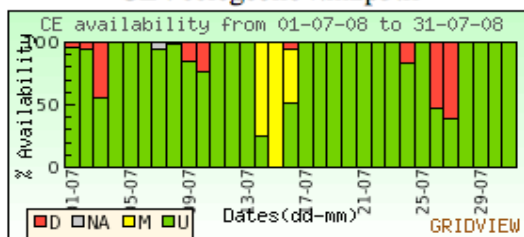
CE : cclcgceli02.in2p3.fr



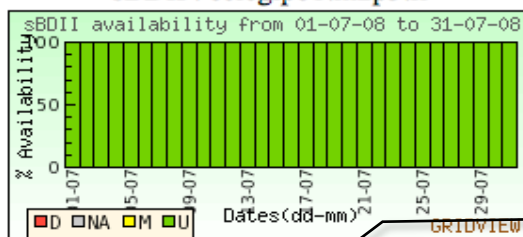
CE : cclcgceli03.in2p3.fr



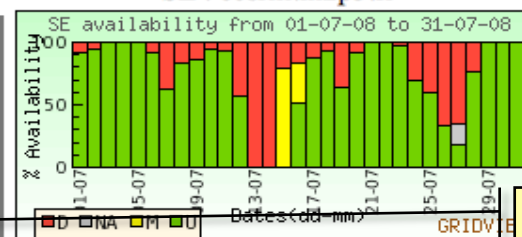
CE : cclcgceli04.in2p3.fr



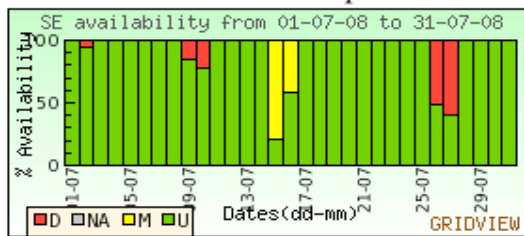
sBDII : cclcgip01.in2p3.fr



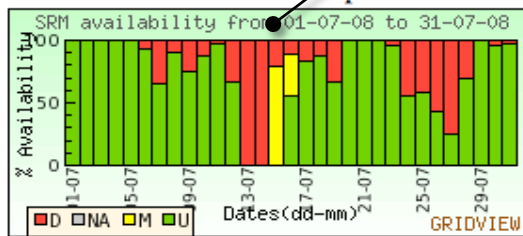
SE : ccsrm.in2p3.fr



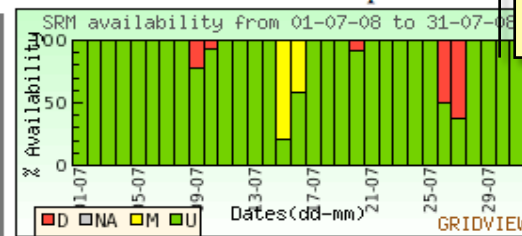
SE : ccsrm02.in2p3.fr



SRM : ccsrm.in2p3.fr



SRM : ccsrm02.in2p3.fr



???

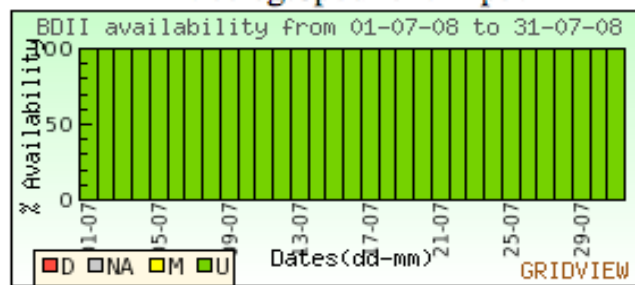
► Disponibilité des sites: tier-1 (3/8)



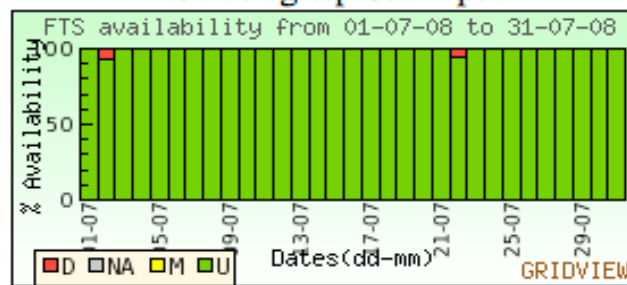
- Juillet 2008 (suite)

Central Services

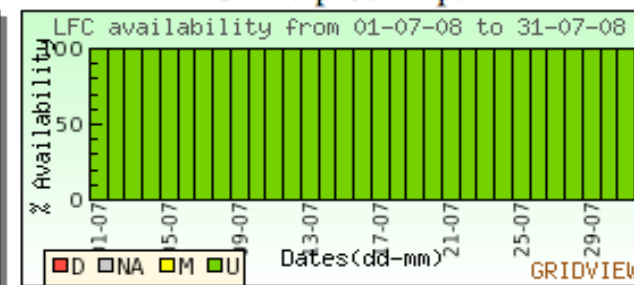
BDII : cclcgtopbdii01.in2p3.fr



FTS : cclcgftsprod.in2p3.fr



LFC : lfc-prod.in2p3.fr

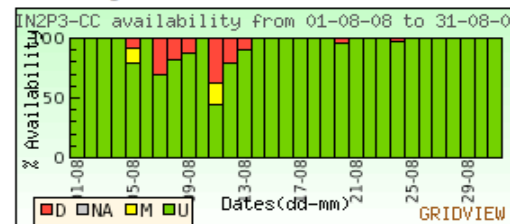


► Disponibilité des sites: tier-1 (4/8)



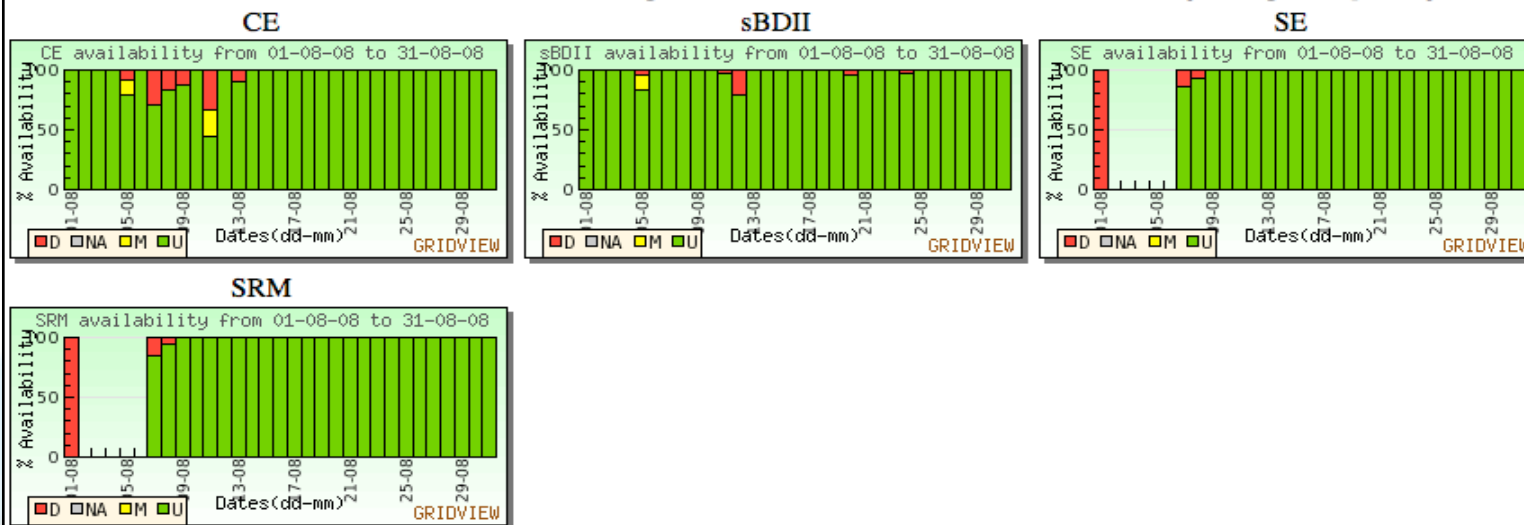
- Août 2008

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



Score: 94%

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



► Disponibilité des sites: tier-1 (5/8)

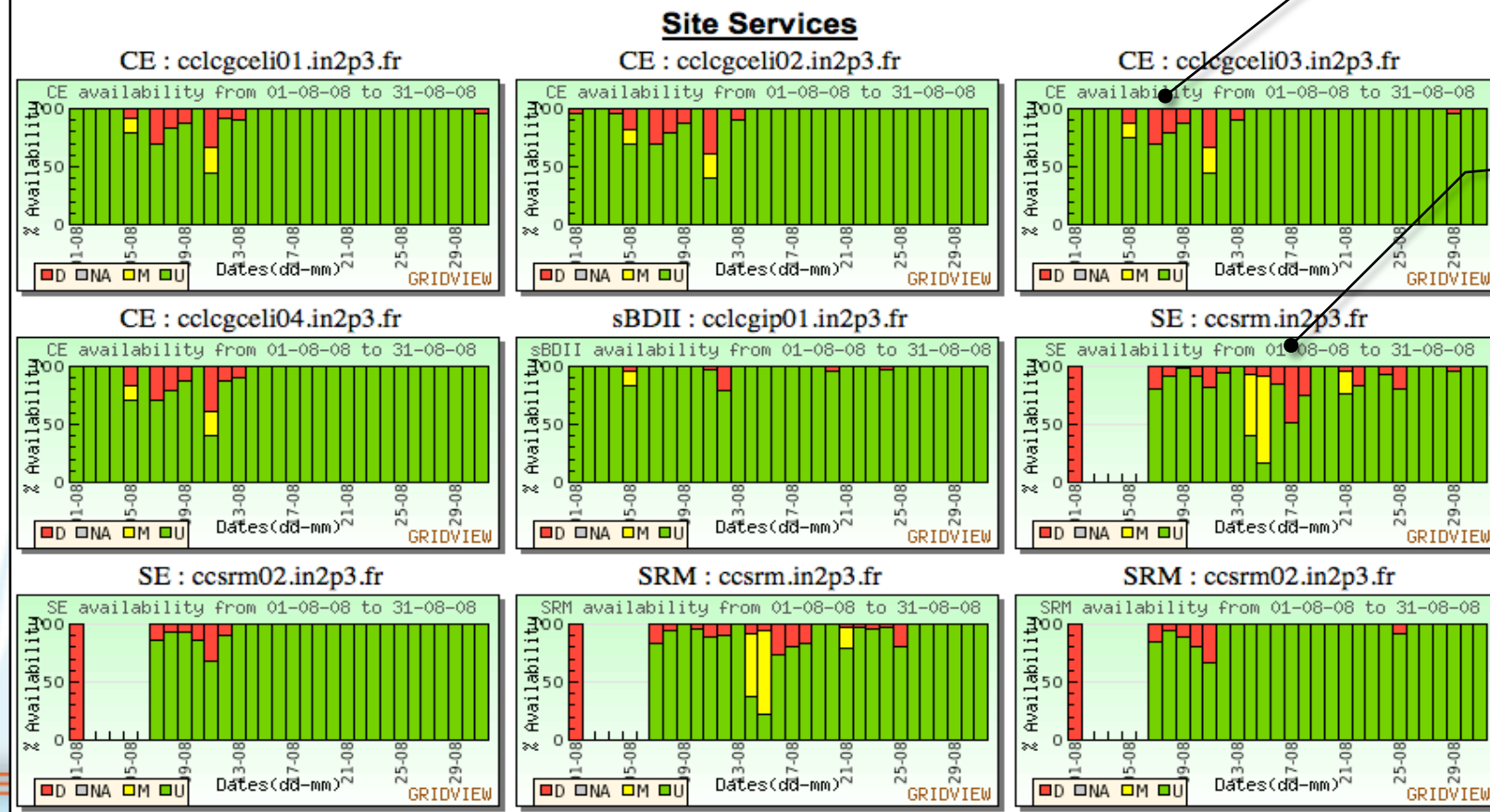


• Août 2008 (suite)

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

Problème simultané sur tous les CEs?

????



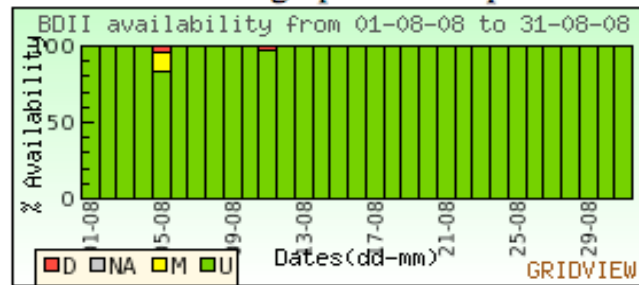
► Disponibilité des sites: tier-1 (6/8)



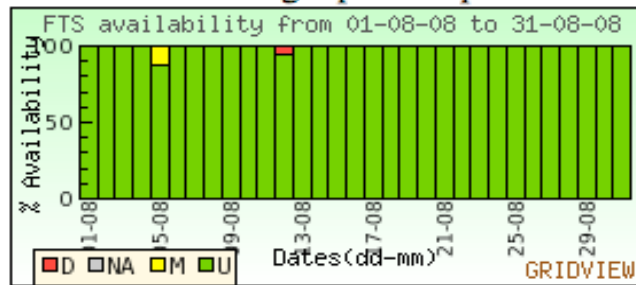
- Août 2008 (suite)

Central Services

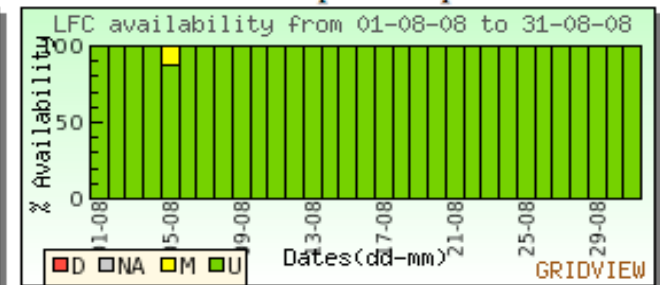
BDII : cclcgtopbdii01.in2p3.fr



FTS : cclcgftsprod.in2p3.fr



LFC : lfc-prod.in2p3.fr





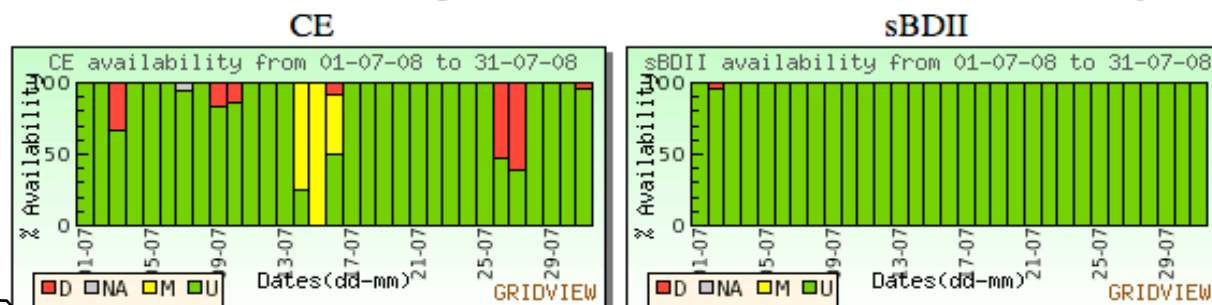
Disponibilité des sites: tier-2 (7/8)



- Juillet 2008

Disponibilité: 87%
(Fiabilité: 93%)

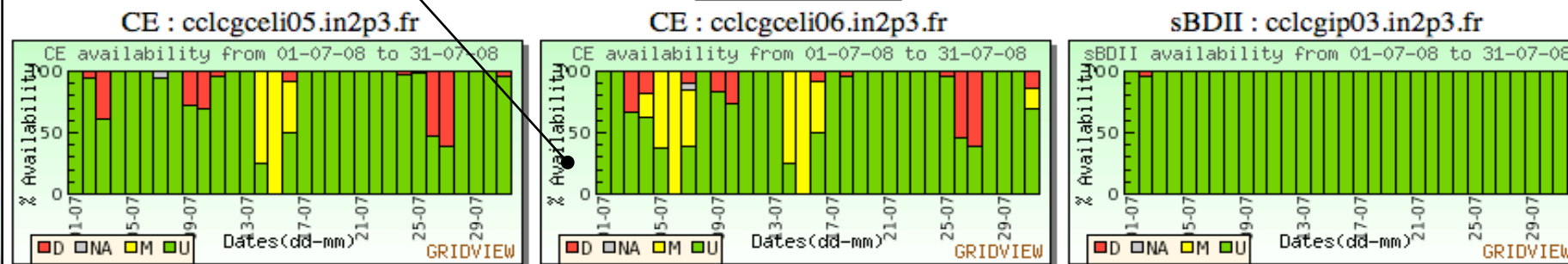
Individual Service Availability for site:IN2P3-CCT2 VO:OPS (Daily Report)



Blah

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

Site Services



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



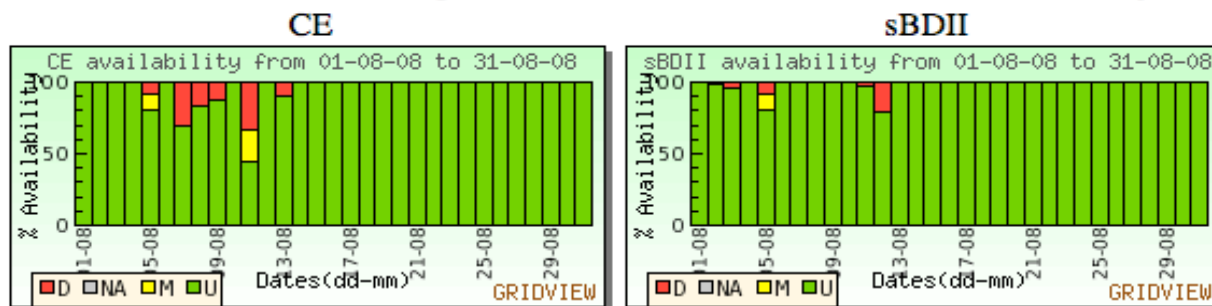
Disponibilité des sites: tier-2 (8/8)



- Août 2008

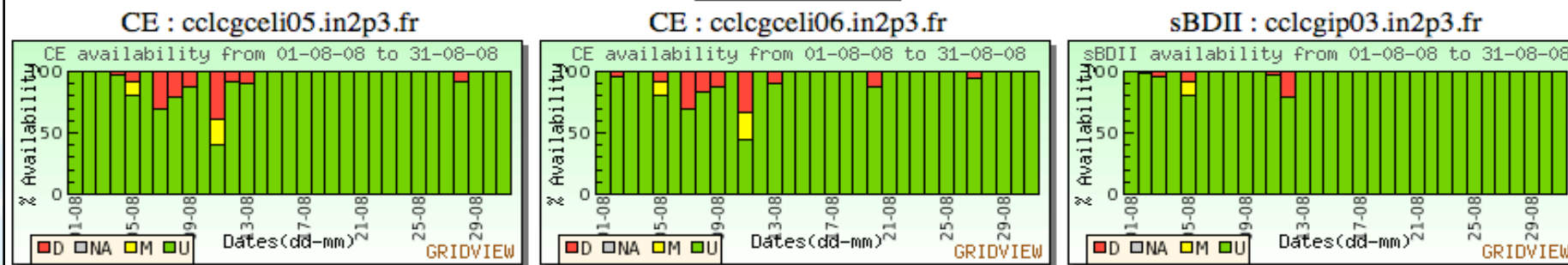
Disponibilité: 94%
(Fiabilité: 95%)

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



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

Site Services

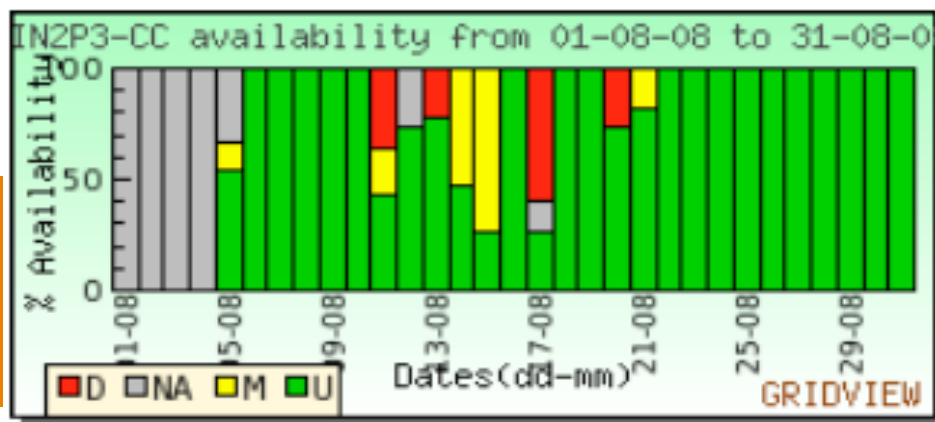


Source: <http://gridview.cern.ch>

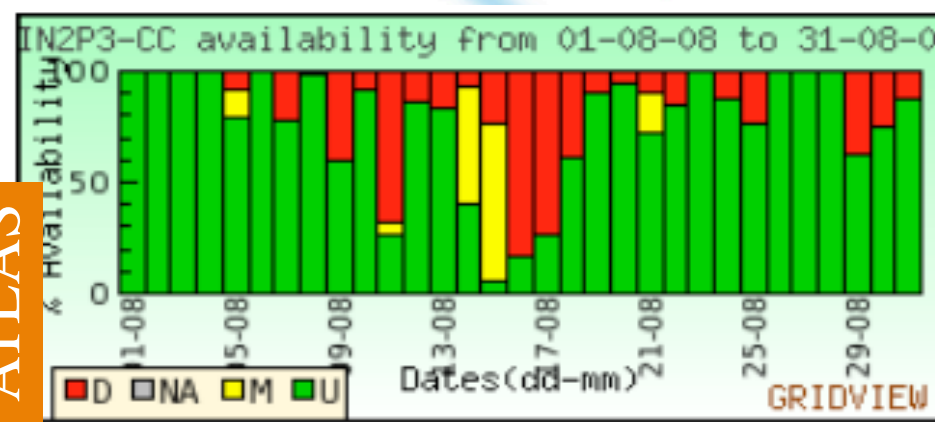
► Disponibilité vue par les VOs



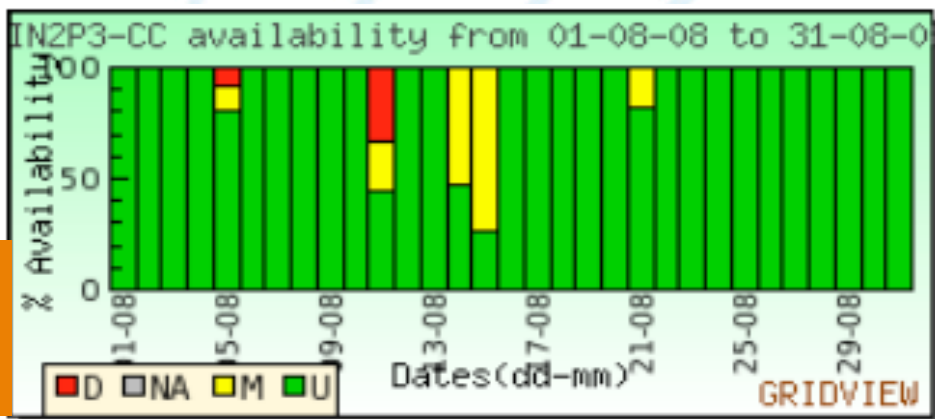
ALICE



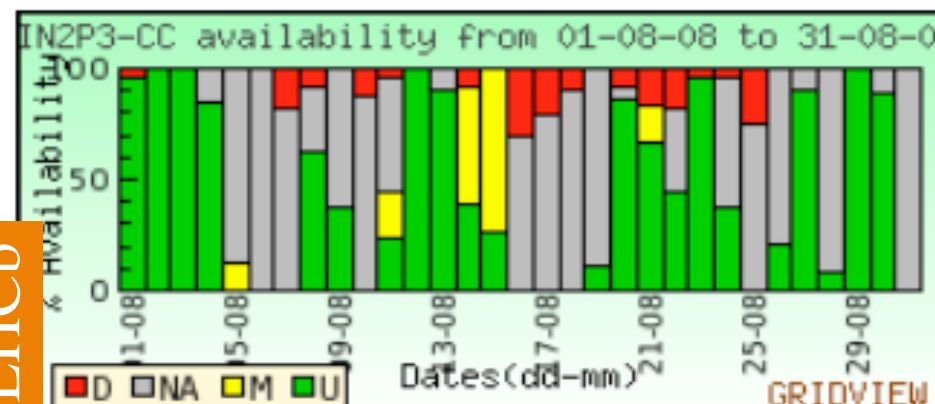
ATLAS



CMS



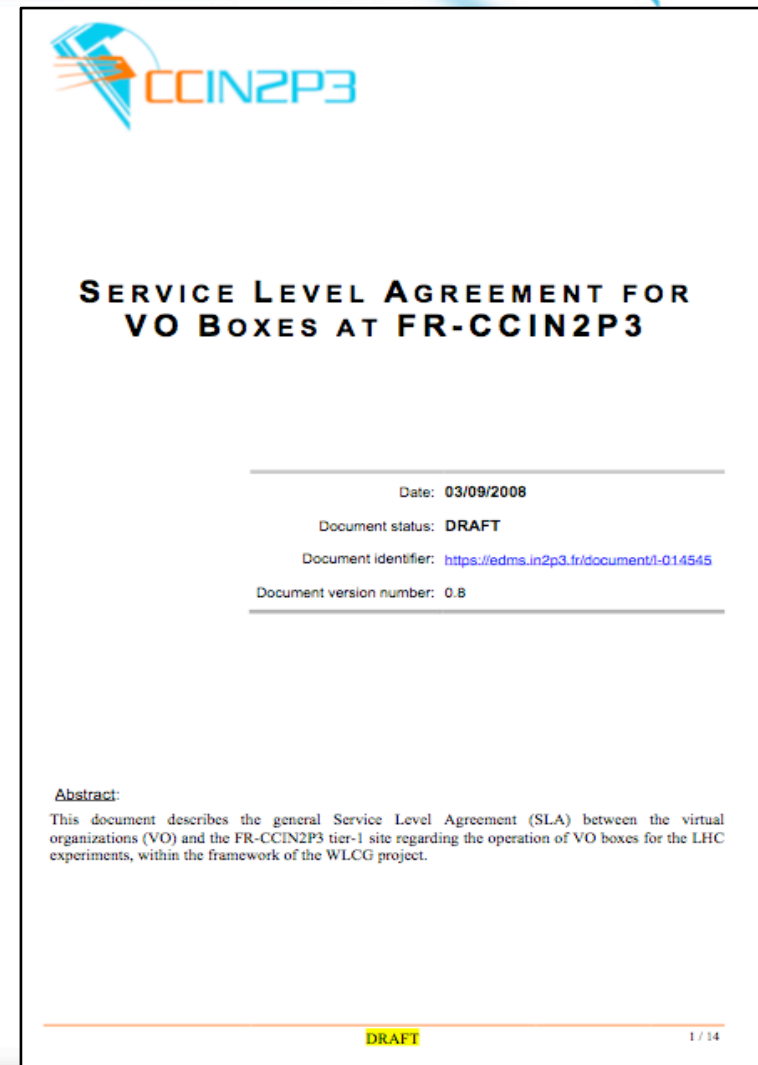
LHCb



▶ Travaux en cours



- En phase finale de la rédaction d'un *Service Level Agreement* concernant les VO boxes pour les expériences LHC
- Etape suivante
 - acceptation formelle par des représentants de chacune des expériences



Source: <https://edms.in2p3.fr/document/I-014545>

► Equipement



- Robotique
 - 1 librairie SL8500 (10.000 slots, 30 dérouleurs) en cours d'installation
 - +10 PB de capacité supplémentaire
 - 1 librairie supplémentaire prévue pour livraison en décembre 2008
- Stockage disque
 - Livraison cette semaine-ci de 108 Thumpers à 40 TB utiles chacun
- CPU
 - Préparation en cours d'un marché pour livraison vers la fin de l'année

Travaux en cours ou terminés



- 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)

► 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

- Page d'accueil de Google le 10/09/2008



► Aujourd'hui



- Outils CMS: Squid & Phedex
- Sujets prochaines réunions
 - Surveillance et alertes NAGIOS
 - Point sur le niveau de redondance de chacun des services (LFC, FTS, dCache, CEs, VOMS, BDII, ...)
- Prochaines réunions
 - <http://indico.in2p3.fr/categoryDisplay.py?categId=102>



Questions/Commentaires

