



In2p3



LCG-France Tier-1 & Analysis Facility

Summary of Activities in 2008 and Plans for 2009

Fabio Hernandez

Fairouz Malek

Comité de Pilotage LCG-France

Lyon, February 13th 2009



Contents

- Human resources
- Service improvements
- Perspectives
- Questions & Comments
- Acknowledgements
- Annex
 - Contribution LCG-France
 - Resource deployment and utilisation

Ressources Humaines

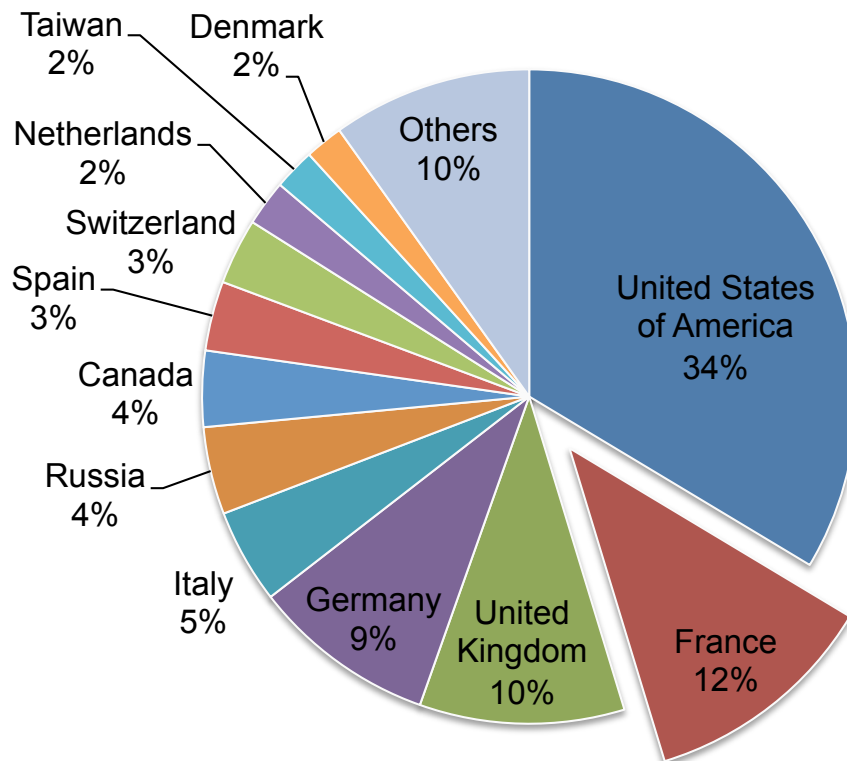
- Mouvement de personnel directement lié aux activités LHC
 - Entrées
 - ◆ Postes permanents: +3 IR, +1 IE
 - ◆ CDDs: +1 chercheur (LHCb), +2 IE, +1 AI
 - Départs
 - ◆ Retraite: -2 IR, -2 TPN
 - ◆ Démission: -1 AI, -2 IR
- Globalement au CC-IN2P3
 - Entrées: +16
 - Départs: -14

Contribution LCG-France

Contribution by LCG-France

WLCG - CPU Contribution per country

All LHC experiments - Jan-Dec 2008



44 countries contributed CPU resources to the LHC experiments in 2008

Resource	Contribution of FR-CCIN2P3 [% resources used at all tier-1s in 2008]
CPU	7%
Disk	15%
Tape	13%

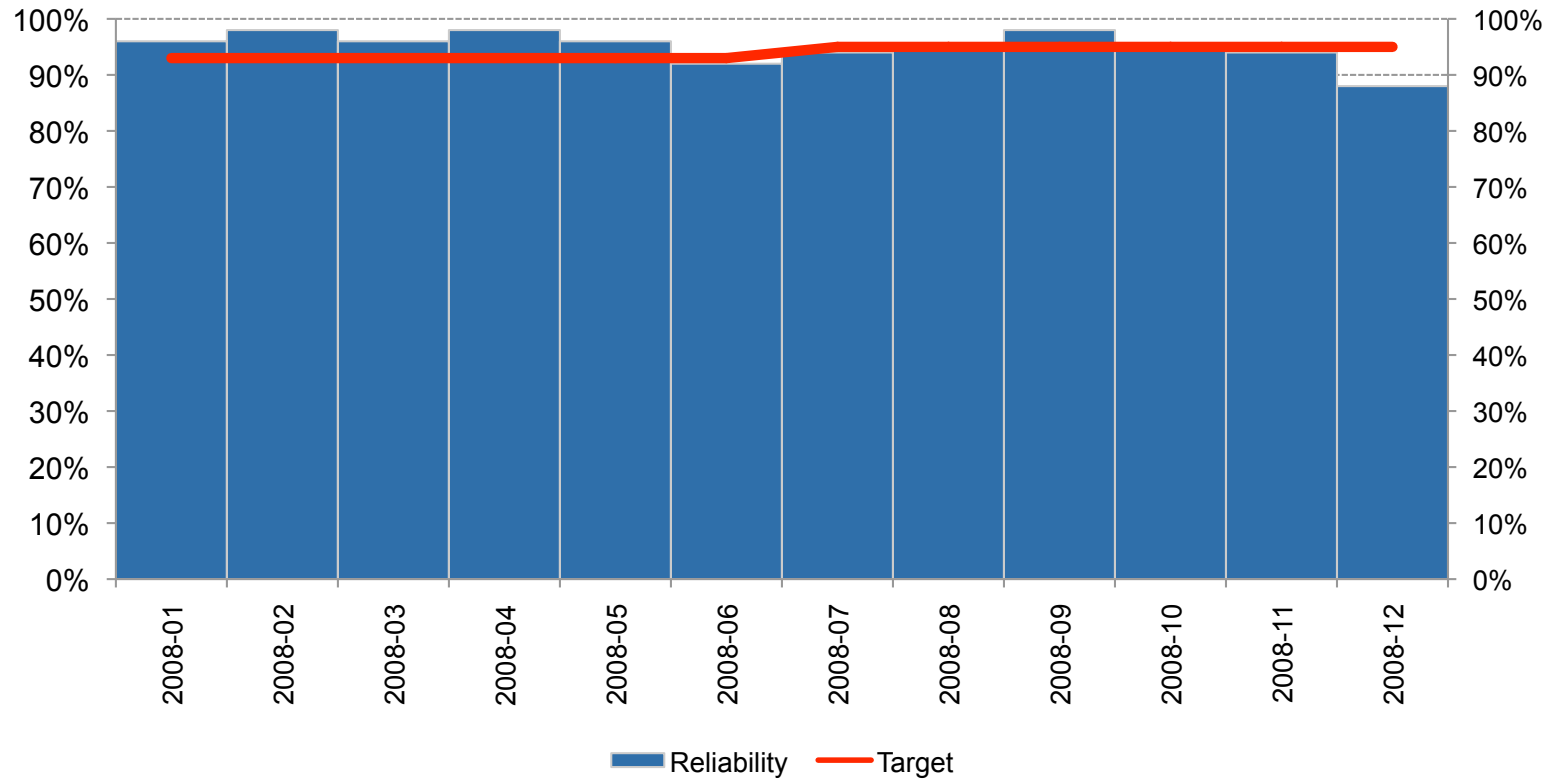
See details in the Annex

Source: [EGEE Accounting Portal](#)

Service Improvements

Reliability & Availability

LCG-France tier-1 - Reliability (VO OPS -- 2008)



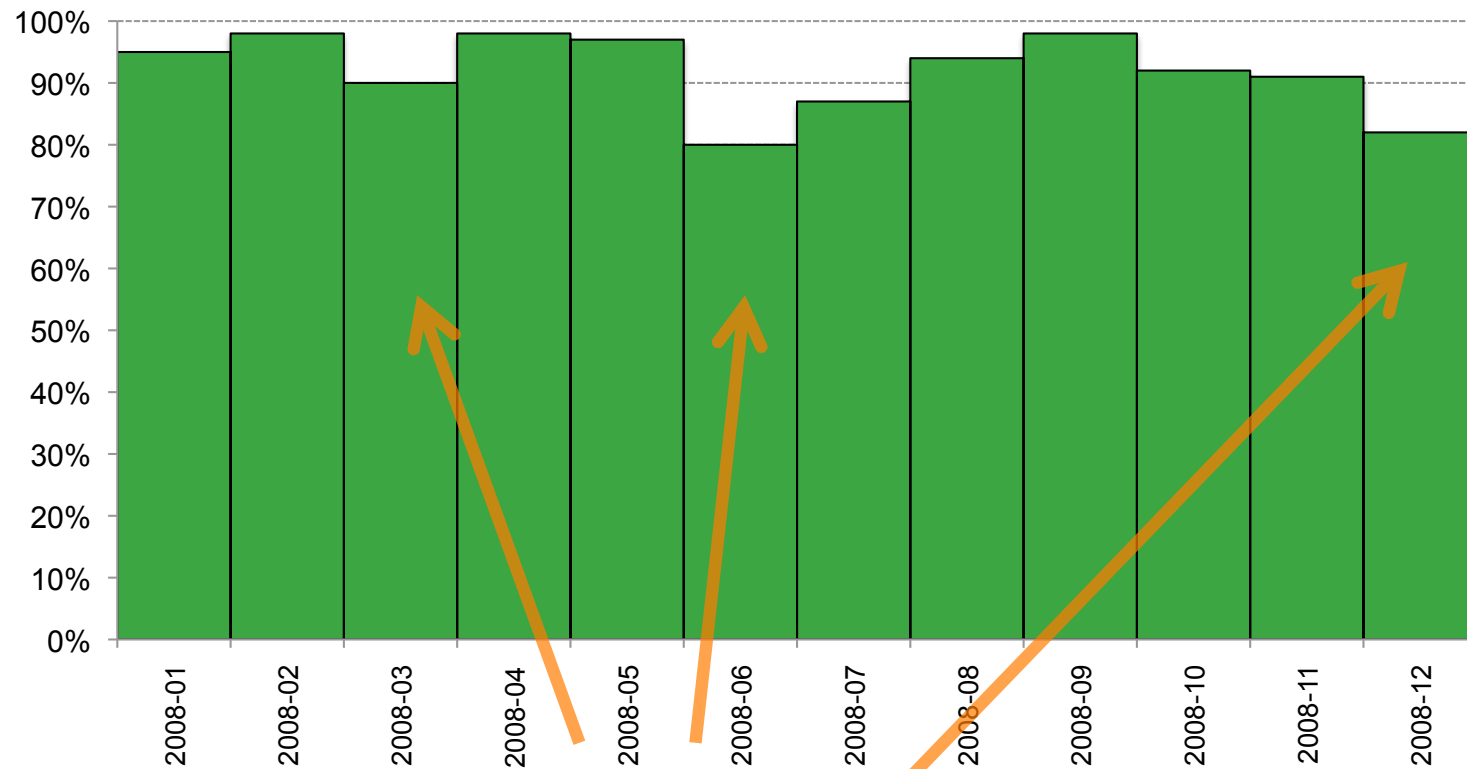
95%

Source: [Reliability of WLCG Tier-1 Sites + CERN](#)



Reliability & Availability (cont.)

LCG-France tier-1 - Availability
(VO OPS -- 2008)



92%

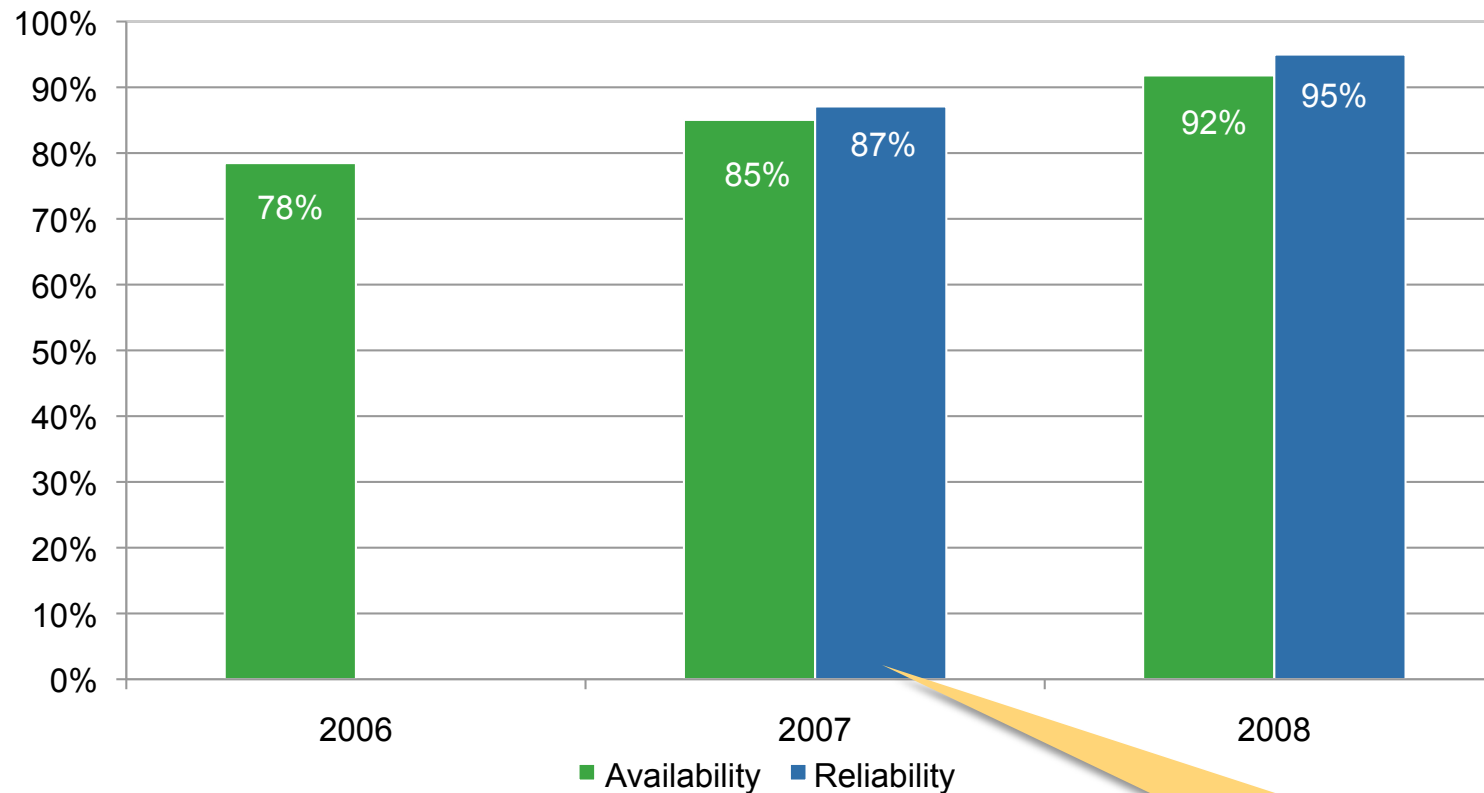
Scheduled shutdown:

- March 11th
- June 3rd
- December 2nd

Source: [Reliability of WLCG Tier-1 Sites + CERN](#)

Reliability & Availability (cont.)

LCG-France tier-1 -- Evolution of availability and reliability (VO OPS)



Reliability measurements in 2007 cover only the period August-December.

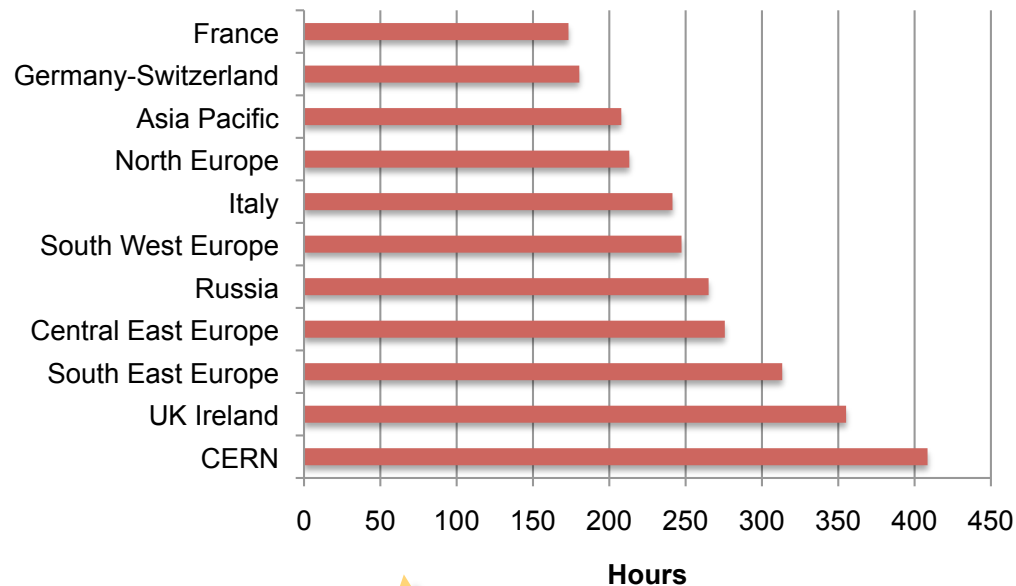
Source: [Reliability of WLCG Tier-1 Sites + CERN](#)

Alarm Tickets

Creation Time	First modification time	Time to formally acknowledge the ticket
24/07/2008 19:16	24/07/2008 19:27	11 min
10/09/2008 05:03	10/09/2008 09:37	274 min
08/10/2008 09:50	08/10/2008 10:11	21 min
24/10/2008 20:36	24/10/2008 20:56	20 min
13/11/2008 09:10	13/11/2008 09:17	7 min

Time to resolve tickets

Average Time To Solve - GGUS Tickets
(EGEE Federations, Mar-Dec 2008)

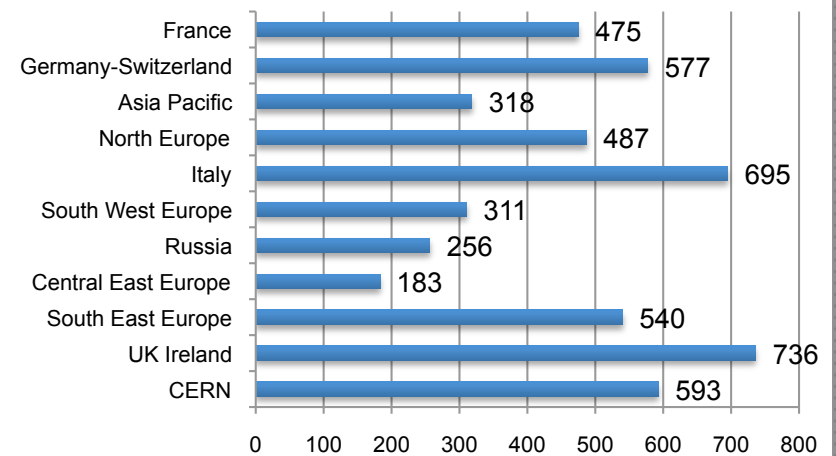


Include but not limited to tickets issued by or related to LHC experiments

Shorter is better

Source: [GGUS Metrics Reports](#)

Number of GGUS Tickets Assigned
(EGEE Federations, Mar-Dec 2008)



Tools for Operations



Bienvenue sur les logbooks du CC-IN2P3

Voici la liste des logbooks disponibles :

Registre	Entrées	Dernière entrée
Support		
support Logbook du support	40	mar 27 jan 2009 18
support-atlas Logbook du support Atlas	36	mer 14 jan 2009 19
support-cms Logbook du support CMS	22	ven 30 jan 2009 15
Exploitation		
exploitation Logbook de l'exploitation	998	lun 09 fév 2009 14
astreinte Logbook de l'astreinte	140	ven 06 fév 2009 21
evenements Evénements generaux liés à l'Opération	33	ven 06 fév 2009 09
problemes-VO Logbook des problemes recurrents des VO's LHC	37	lun 09 fév 2009 11
Services		
HPSS Logbook d'HPSS	159	lun 09 fév 2009 08
dCache Logbook de dCache	40	lun 09 fév 2009 11
logon Logbook de logon@cc	2112	ven 06 fév 2009 17
Suivi des opérations		
operations-HPSS Logbook des opérations sur HPSS	8	ven 30 jan 2009 12
operations-dCache Logbook des opérations sur dCache	24	ven 06 fév 2009 09
operations-LFC Logbook des opérations sur LFC	3	mer 04 fév 2009 14
operations-FTS Logbook des opérations sur FTS	13	ven 06 fév 2009 17
operations-CE Logbook des opérations sur les CE's	6	jeu 05 fév 2009 17
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	-

Nagios®

General

- Home
- Documentation

Monitoring

- Tactical Overview
- Service Detail
- Host Detail
- Hostgroup Overview
- Hostgroup Summary
- Hostgroup Grid
- Servicegroup Overview
- Servicegroup Summary
- Servicegroup Grid
- Status Map
- 3-D Status Map

Service Problems

- Unhandled
- Host Problems
- Unhandled
- Network Outages

Show Host:

Comments

Downtime

Process Info

Performance Info

Scheduling Queue

Reporting

Trends

Availability

Alert History

Alert Summary

Notifications

Event Log

Configuration

View Config

Host	Service	Status	Output	Current State	Next Check	State Change	State Duration	State Type	State Info
ccxrdsn009	Checking port 1998	OK	2009-02-09 14:04:22 17d 22h 25m 24s	1/1	Check ccxrdsn009 on port 1998 is successful				
	Checking port 1999	OK	2009-02-09 14:19:30 17d 22h 10m 16s	1/1	Check ccxrdsn009 on port 1999 is successful				
	machine-status	OK	2009-02-09 14:20:35 6d 0h 9m 11s	1/2	SUCCESS: CC machine_status is: Up				
ccxrdsn010	Checking port 1998	OK	2009-02-09 14:04:38 17d 22h 25m 8s	1/1	Check ccxrdsn010 on port 1998 is successful				
	Checking port 1999	OK	2009-02-09 14:17:20 17d 22h 28m 18s	1/1	Check ccxrdsn010 on port 1999 is successful				
	machine-status	OK	2009-02-09 14:28:27 6d 0h 1m 19s	1/2	SUCCESS: CC machine_status is: Up				
ccxrdsn011	Checking port 1998	OK	2009-02-09 14:28:37 17d 22h 24m 51s	1/1	Check ccxrdsn011 on port 1998 is successful				
	Checking port 1999	OK	2009-02-09 14:20:03 17d 22h 9m 43s	1/1	Check ccxrdsn011 on port 1999 is successful				
	machine-status	OK	2009-02-09 14:26:19 5d 23h 53m 27s	1/2	SUCCESS: CC machine_status is: Up				
ccxrdsn012	Checking port 1998	OK	2009-02-09 14:05:11 17d 22h 24m 35s	1/1	Check ccxrdsn012 on port 1998 is successful				
	Checking port 1999	OK	2009-02-09 14:20:20 17d 22h 9m 26s	1/1	Check ccxrdsn012 on port 1999 is successful				
	machine-status	OK	2009-02-09 14:20:41 6d 0h 9m 5s	1/2	SUCCESS: CC machine_status is: Up				
ccxrdsn013	Checking port 1998	OK	2009-02-09 14:05:28 17d 22h 24m 18s	1/1	Check ccxrdsn013 on port 1998 is successful				
	Checking port 1999	OK	2009-02-09 14:20:36 17d 22h 9m 10s	1/1	Check ccxrdsn013 on port 1999 is successful				
	machine-status	OK	2009-02-09 14:28:34 6d 0h 1m 12s	1/2	SUCCESS: CC machine_status is: Up				
ccxrdsn014	Checking port 1998	OK	2009-02-09 14:05:44 17d 22h 24m 2s	1/1	Check ccxrdsn014 on port 1998 is successful				
	Checking port 1999	OK	2009-02-09 14:20:53 17d 22h 8m 53s	1/1	Check ccxrdsn014 on port 1999 is successful				
	machine-status	OK	2009-02-09 14:26:26 5d 23h 53m 20s	1/2	SUCCESS: CC machine_status is: Up				
ccxrdsn015	Checking port 1998	OK	2009-02-09 14:06:01 17d 22h 23m 45s	1/1	Check ccxrdsn015 on port 1998 is successful				
	Checking port 1999	OK	2009-02-09 14:21:09 17d 22h 8m 37s	1/1	Check ccxrdsn015 on port 1999 is successful				
	machine-status	OK	2009-02-09 14:20:48 6d 0h 8m 58s	1/2	SUCCESS: CC machine_status is: Up				
ccxrdsn016	Checking port 1998	OK	2009-02-09 14:06:17 2d 20h 53m 29s	1/1	Check ccxrdsn016 on port 1998 is successful				
	Checking port 1999	OK	2009-02-09 14:21:26 2d 20h 38m 20s	1/1	Check ccxrdsn016 on port 1999 is successful				
	machine-status	WARNING	2009-02-09 14:23:40 2d 22h 31m 6s	2/2	WARNING: CC machine_status is: Test				
ccxrdsn017	Checking port 1998	OK	2009-02-09 14:06:34 17d 22h 23m 12s	1/1	Check ccxrdsn017 on port 1998 is successful				
	Checking port 1999	OK	2009-02-09 14:21:42 17d 22h 8m 4s	1/1	Check ccxrdsn017 on port 1999 is successful				
	machine-status	OK	2009-02-09 14:26:33 5d 23h 53m 13s	1/2	SUCCESS: CC machine_status is: Up				
ccxrdsn018	Checking port 1998	OK	2009-02-09 14:06:51 17d 22h 22m 55s	1/1	Check ccxrdsn018 on port 1998 is successful				
	Checking port 1999	OK	2009-02-09 14:21:59 17d 22h 7m 47s	1/1	Check ccxrdsn018 on port 1999 is successful				
	machine-status	OK	2009-02-09 14:25:55 5d 4h 53m 51s	1/2	SUCCESS: CC machine_status is: Up				
ccxrdsn019	Checking port 1998	OK	2009-02-09 14:07:07 17d 22h 22m 39s	1/1	Check ccxrdsn019 on port 1998 is successful				
	Checking port 1999	OK	2009-02-09 14:22:15 17d 22h 7m 31s	1/1	Check ccxrdsn019 on port 1999 is successful				
	machine-status	OK	2009-02-09 14:28:47 6d 0h 0m 59s	1/2	SUCCESS: CC machine_status is: Up				
ccxroot	Checking port 1998	OK	2009-02-09 14:27:40 16d 21h 32m 6s	1/1	Check ccxroot on port 1998 is successful				
	Checking port 1999	OK	2009-02-09 14:12:48 16d 21h 16m 58s	1/1	Check ccxroot on port 1999 is successful				
dCache-virtual-host	Searching for OFFLINE pool in dCache EGEE	OK	2009-02-09 14:26:40 3d 3h 59m 21s	1/1	No offline pool in egee				
	Searching for OFFLINE pool in dCache LCG	WARNING	2009-02-09 14:27:07 0d 1h 33m 39s	1/1	4 NON-critical pool(s) OFFLINE in lcg				
diva-virtual-host	Check on DIVA database	OK	2009-02-09 14:26:16 20d 4h 55m 28s	1/1	DMS up				
	Checks for offline drives	OK	2009-02-09 14:25:31 10d 16h 58m 15s	1/1	No offline drive detected				
	Queries DIVA Master	OK	2009-02-09 14:26:44 20d 4h 55m 2s	1/1	Diva master responding				
	Searches for pending requests in DIVA	OK	2009-02-09 14:25:57 56d 18h 35m 51s	1/1	No pending request				
lco-farm	Check DataBase status	OK	2009-02-09 14:23:26 13d 1h 57m 20s	1/1	Check on lcg database is successful				
	Check Master status	OK	2009-02-09 14:23:26 13d 1h 43m 20s	1/1	Check farm lcg successful				
nagios-virtual-host	Check mail grouper system	OK	2009-02-09 14:22:10 47d 0h 53m 47s	1/3	OK: No problem detected				
oracle-virtual-host	Checks DB AMI	OK	2009-02-09 14:26:23 73d 7h 4m 27s	1/1	dbami was contacted successfully.				
	Checks DB ATL	OK	2009-02-09 14:27:36 88d 5h 43m 10s	1/1	dbatl was contacted successfully.				
	Checks DB LHCb	OK	2009-02-09 14:26:49 73d 11h 16m 1s	1/1	dbhcb was contacted successfully.				



Tools for Operations (cont.)

Planification

[[arret:arret02122008]]

Afficher le texte source Anciennes révisions

Vous êtes ici: start » arrets » arret02122008

Arrêt du 02 décembre 2008

Tâches de planification

Tâches

- Envoi du premier mail de recensement (mod)
- Relance pour le recensement des interventions
- Pré-planning établi
- Réunion de préparation de l'arrêt (principale)
- Planning validé par centre

Interventions prévues

Equipe Infrastructure

Groupe Sysunix

- Racks DDN : MAJ hard (2h) + firmwa
 - Durée : 4 heures (en partie en te
 - Responsable : PYF
 - Implications :
 - impacts sur :
 - HPSS : arrêt programmé po
 - GPFS : arrêt programmé po
- Workers + Interactives : MAJ sécurité
 - Durée : 2 heures
 - Responsable : YP
 - Implications : workers + interacti
- AFS : quelques manips
 - Durée : 2 heures
 - Responsable : XC
 - Implications : quelques perturbat
- MAJ X3550 :
 - Durée : 2 reboots (30 minutes)
 - Responsable : MP
 - Implications :
 - certaines services seront indis

Déroulement

Planning

30 novembre 2008

Début prévu	Fin prévu	Inte
04H00	H + 50	Dim

01 décembre 2008

Début prévu	Fin prévu	Inte
08H00		Ferr

02 décembre 2008

Début prévu	Fin prévu	Intervention
08H00	08H30	Vider les wor
08H00	09H00	Arrêt du serv
08H00	08H30	Arrêt du serv
		Arrêt des cro
08H30	17H00	MAJ BQS-DB
08H00	11H00	MAJ micrococ
08H15		Bloquer l'accè
08H30	10H30	MAJ Kernel d
08H30	09H30	Changement
		Les thumpers
09H30	10H00	Remise en se
08H30	10H30	Arrêter les s
		la coupure de
09H00	12H00	Maintenance
09H00	12H00	Thumpers ar
09H00	12H00	NGOP arrêté
10H30	11H00	Intervention
		les services
		être arrêtés
11H00	13H00	travaux sur a

WIKI OPERATIONS

Bilan

Bilan technique

Equipe Infrastructure

Groupe Sysunix

- MAJ des BIOS + firmwares (hdd,bmc,controller)
- vérification de la redondance électrique de toutes nos machines de service (rack 1 à 4)
 - un switch du rack service 1 dégage si B est coupée -> on le sait.
- reinstalle des ccwl*, ccali*, ccpl*
- MAJ kernel, openafs, gpfs, logons grille
- MAJ logons grille sur master BQS (bqsmn1 et 2, bqsmli02)
- nouveaux switches (remontée 10 Gbts) sur les 6 racks de "petits" workers Dell, reste que les gros, maintenant
- préparation du départ des IBM-e326m

Bilan matériel : 2 alims dans le rack de service2, 2 dans Semiper00, deux disques de machines de service. Bilan non anodin, si on avait coupé E (deuxième patte électrique des machines de service), ces machines auraient grillé (il s'agit de ccdb01 et ccxrdri02).

- disjonctions de disjoncteurs de racks lors de la reprise du courant : des racks sont restés sur une « alimentation »
- non suivi de procédure de reboot de machines de services :
 - services grilles (BDII, VO Boxes, ...)
 - services locaux qui ne devait pas être arrêtés : xrootd
- La coupure électrique des armoires électriques A et B a engendrée de gros imprévus sur des machines critiques du centre : NIS + DNS + switches
 - interruption services minimum dcache-LCG pendant 1 heure
 - cela aurait pu engendrer de gros retards dans des interventions annexes (BQS par exemple)
 - a provoqué du retard dans l'installation des workers : redémarrage très laborieux
 - de nombreux autres services touchés plus ou moins directement :
 - SL8500 : hors service a cause d'un switch éteint
 - machines diva + tms éteintes
 - pas de backup TSM sur bande possible

Groupe Stockage

- problème matériel lors du redémarrage (pas forcément impliqué par l'arrêt):
 - HPSS : unité de stockage inaccessible
 - baie disque de la DSI qui pose problème

Groupe Telecom

- changement des branchements des alimentations de ccpn-core pour etre sur 2 armoires electriques differentes.
- changement de plusieurs switches "satellites" pour passer des machines en gigabit/s avec remontée 10 Gb/s vers les switches "de coeur de reseau":
 - 1 switch interactives
 - 5 switches workers (il ne reste plus qu'a changer les switches des derniers racks Dell plus grands que les autres).

Equipe Développements

- BQS :
 - La manip avait pour but le passage de MySQL v4 en v5 (5.0.67)
 - Tout s'est bien passée, la manip a duré une bonne partie de la journée comme prévue

Tools for Operations (cont.)

Centre de Calcul IN2P3 - CNRS : page d'accueil

http://cc.in2p3.fr/cc_accueil.php3?lang=en

iPhone Dev Center Apple Yahoo! Google Maps YouTube Wikipedia News (3205) Popular W-LCG

CCIN2P3 CNRS IN2P3 Autres sites CNRS

CCIN2P3 computing systems status

Arrêt partiel du Centre le 30/09/2008

Services status :

Network service : available
AFS Service : available
Interactive access : available
Batch system : available
HPSS service : unavailable until 03:00PM
Dcache service : unavailable until 02:00PM
Xrootd and SRB service : degraded
SPS (NFS + GPFS) space : degraded
TSM service : available
Oracle service : available
MySQL + PostgreSQL : available
Web service : available
CVS/SVN services : available

updated Tuesday 30 September 2008 at 11h26 - CET

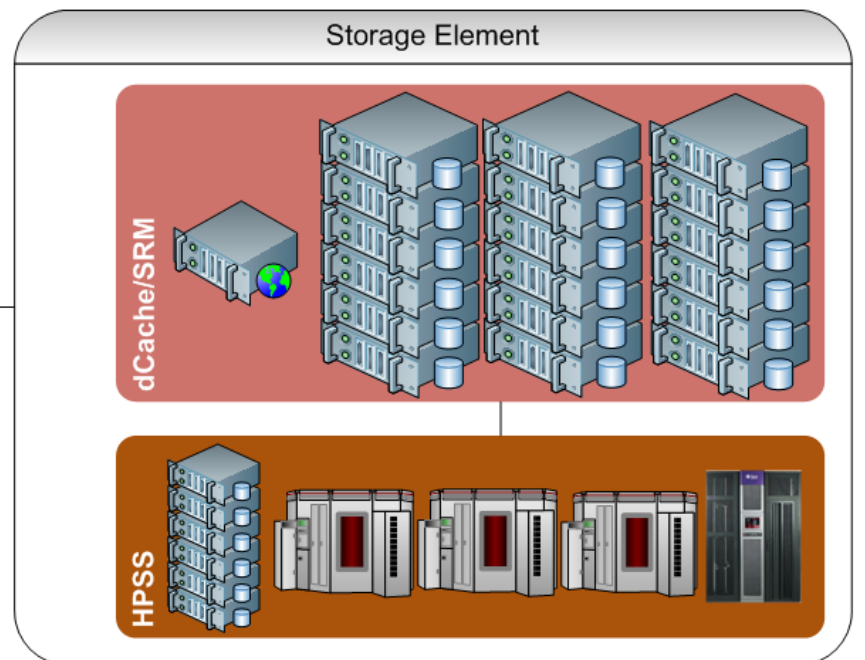
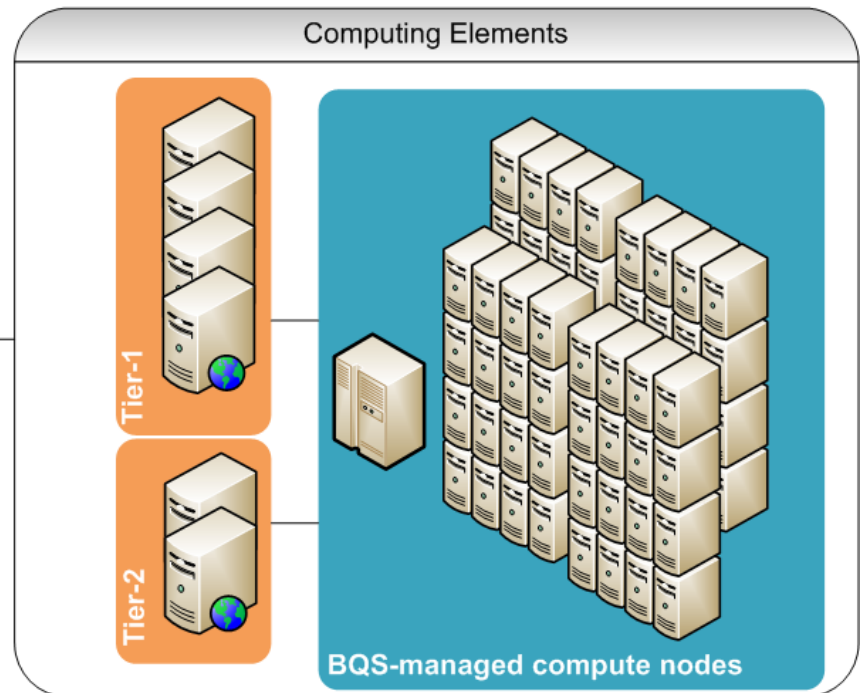
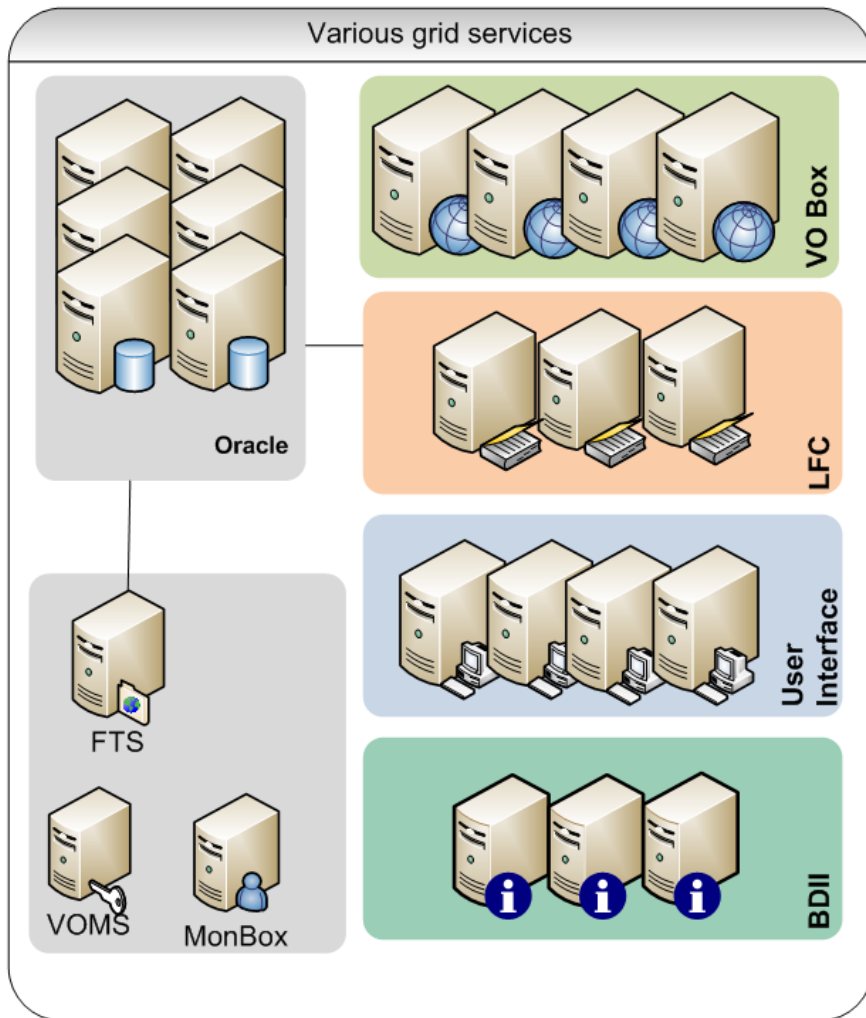
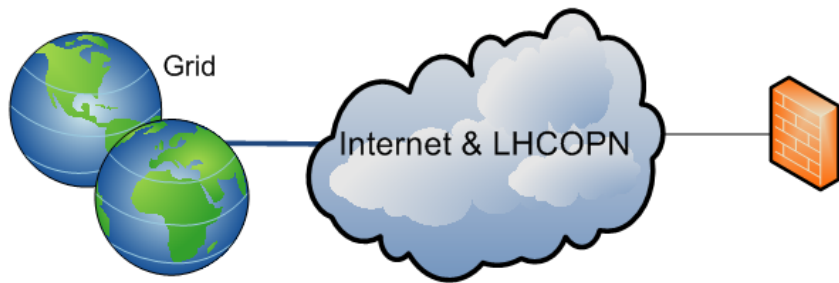
[Close this popup](#)

Site language : [English](#)

research and a solid background in scientific computing. The successful candidate will be based at CC-IN2P3 in Lyon, the French (...)

Information des utilisateurs de l'évolution des opérations et de l'état des services pendant l'arrêt

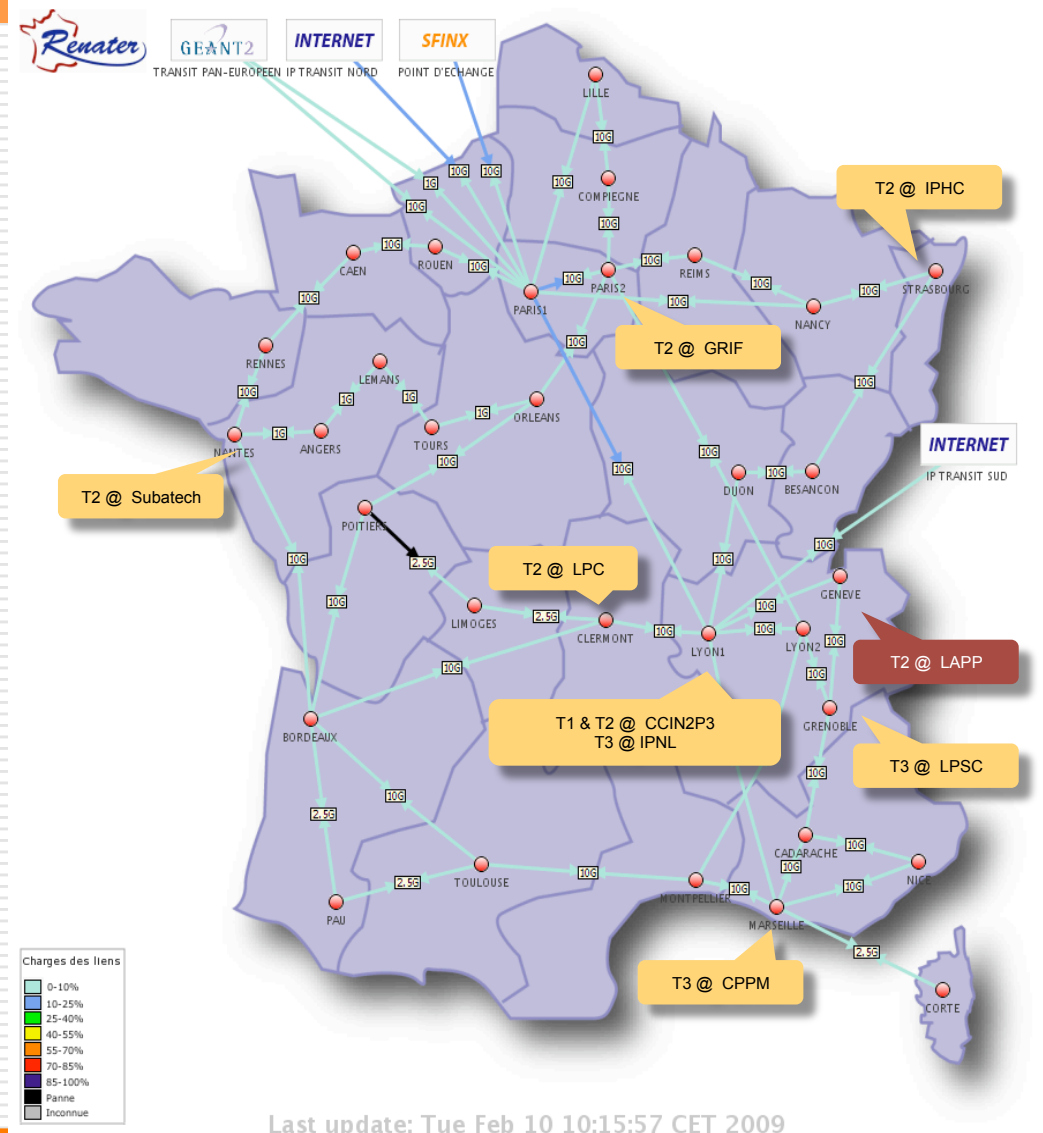
Site Overview



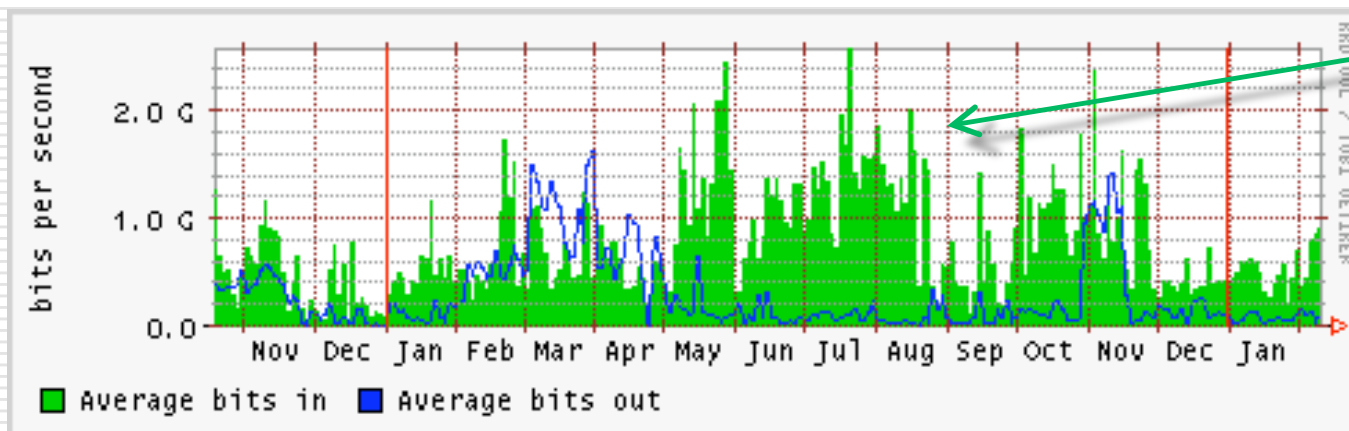
Connectivity

Wide Area Network

- Tier-1 well connected to tier-0, other tier-1s and associated tier-2s
 - National T2s and T3s connected at 1 Gbps or 10 Gbps
 - GRIF: interconnection of sites at 10 Gbps (final phase)
 - LAPP: weak connectivity
 - ◆ Ongoing discussion with RENATER about the technical possibilities of circumventing the problem

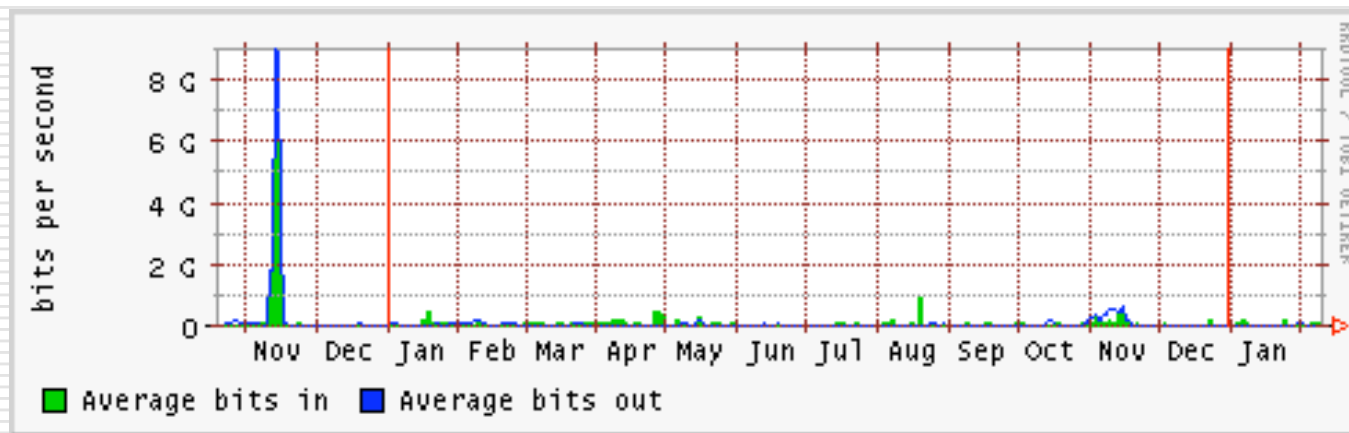


Wide Area Network (cont.)



CERN → CCIN2P3

Link: CCIN2P3 – CERN, 10 Gbps [Snapshot taken on 10/02/2009]

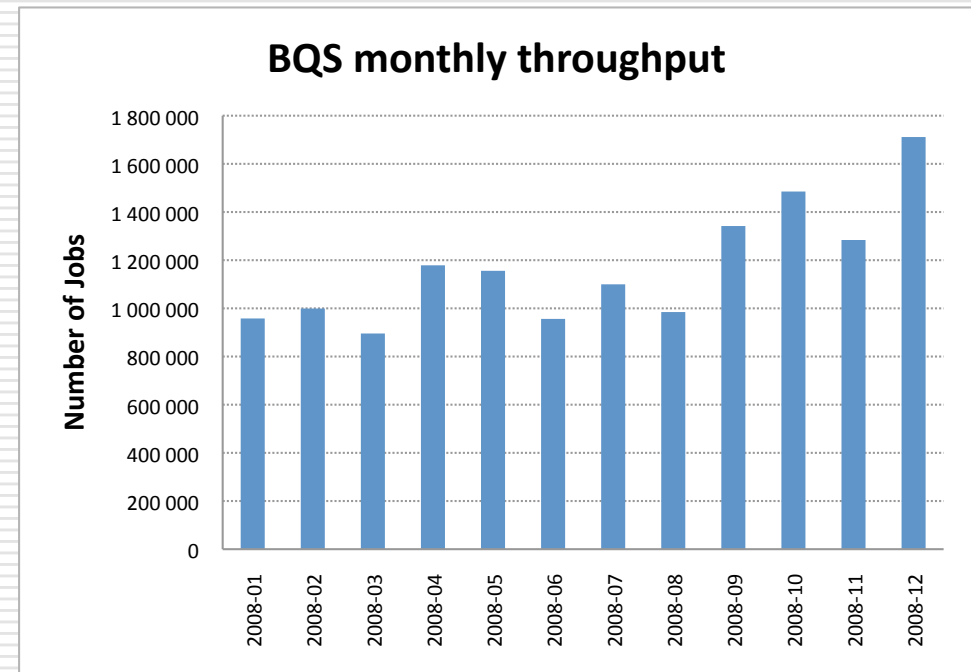


Link: CCIN2P3 – GRIDKA, 10 Gbps [Snapshot taken on 10/02/2009]

Batch Service

Batch System

- New features implemented to allow for better sharing of CPU allocation by jobs within the same VO
 - Privileged user of the VO can modify the shares
 - Useful for ATLAS and CMS
- BQS associates information relevant to grid jobs to the local jobs
 - Grid job id, VO name, grid user id, site type (T1, T2, AF)
 - Tools for querying this information and acting on it
- Modifications in the internal scheduler to speed it up
- Improvements of the regulation tools
 - Control of job submission by user, VO, submitter host, max running jobs per VO (and per share within the VO)



Computing Element

- Hardware and software upgrade of all grid nodes in 2008
 - More powerful hardware
 - SL4 64bits wherever possible
- DNS load balancing of information system nodes
 - Top BDII & Site BDII
 - ◆ 2 load-balanced machines
 - ◆ A single Site BDII publishes information for both tier-1 and tier-2

Computing Element (cont.)

- **Current setup**
 - Tier-1: 4 computing elements
 - Tier-2: 2 computing elements
- **Configuration for availability**
 - Each VO is supported by 2 computing elements (both for tier-1 & tier-2)
 - Updates performed without service interruptions
 - Configuration more complex
- **Stricter separation of roles**
 - Access restrictions applied to tier-1 and tier-2 according to the VOMS role of the submitter, for ATLAS and CMS
 - Intra-VO scheduling implemented

Computing Element (cont.)

- Work in progress
 - Improvement of the error handling of the BQS JobManager
 - ◆ Better tracking of local jobs in « unusual » states
 - ◆ Positive impact for ATLAS (PANDA jobs)
 - Modification of the BQS JobManager for exploiting the capabilities of BQS of scheduling jobs of several activities of the same experiment
 - Migration of all the grid nodes to new hardware, to alleviate an intermittent problem observed with some machines crashing when local disk (in RAID mode) used intensively
- Perspectives for 2009
 - Deployment of a CREAM-based CE for BQS
 - ◆ Slightly delayed due to a skying accident
 - Improve the monitoring of the grid nodes
 - Improve the tools for operating the grid nodes

Data Storage & Data Transfer Services

Mass Storage

- Consolidation of the storage team
 - Responsible for the HPSS, dCache, GPFS, xrootd services
 - HPSS: currently 2.5 FTE, 4 people
 - Better interaction between dCache and HPSS experts, and also with the Operations team
- Hardware improvements
 - 2 new cartridge libraries STK SL8500, 10.000 slots each
 - Migration of data to higher-density cartridges
 - +200 TB disk cache
 - ◆ Big impact on “small” files, which now reside mainly on disk (120 TB)

Mass Storage (cont.)

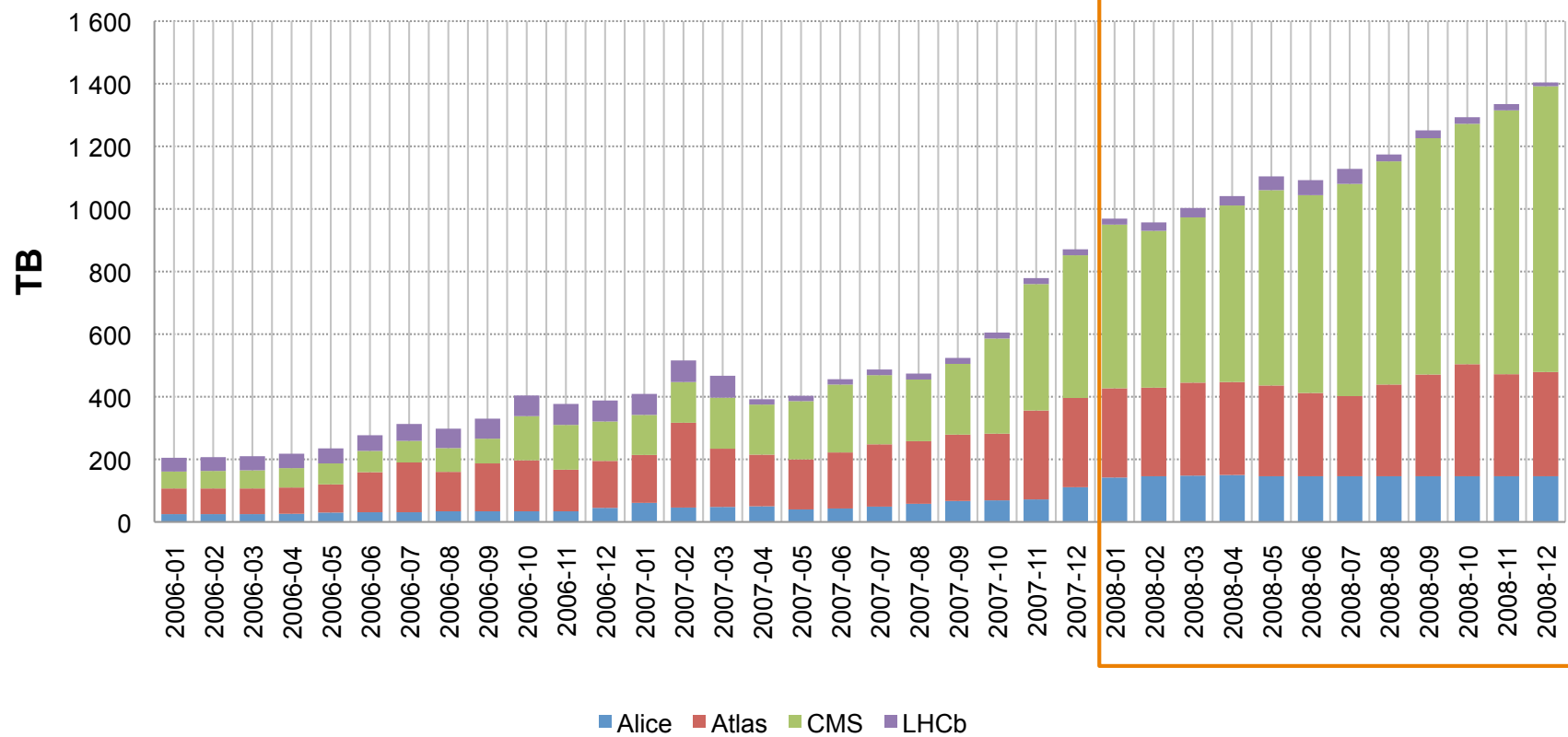
- **Hardware improvements (cont.)**
 - Uniform hardware per component: tape movers and disk movers
- **Issues**
 - Sub-optimal interaction between dCache and HPSS for staging in data
 - ◆ Observed rates of about 400 GB/hour not enough for ATLAS and CMS
 - ◆ Internal tests show that too much time spent in repositioning the medium for reading files on the same cartridge

Mass Storage (cont.)

- **Issues (cont.)**
 - Work started for adapting a scheduler which will mediate the interaction between dCache and HPSS, so that staging requests issued by dCache are presented to HPSS in good order to optimise cartridge mounts and tape movement
 - Based on a work done at Oakridge, currently in use at BNL
- **Perspectives**
 - Deployment of HPSS v6.2 (during next summer)
 - ◆ 30 additional tape drives will then be put in production
 - ◆ New core server ordered
 - ◆ This deployment is a major operation
 - Improve the tools for monitoring the system

Mass Storage (cont.)

Evolution of LHC data managed by HPSS



Source: [WLCG CERN & Tier-1 Sites Accounting Report](#)

Mass Storage (cont.)

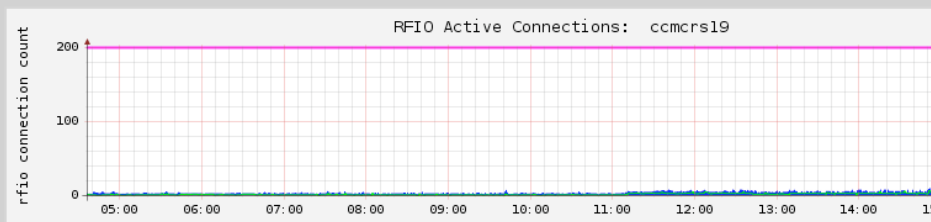
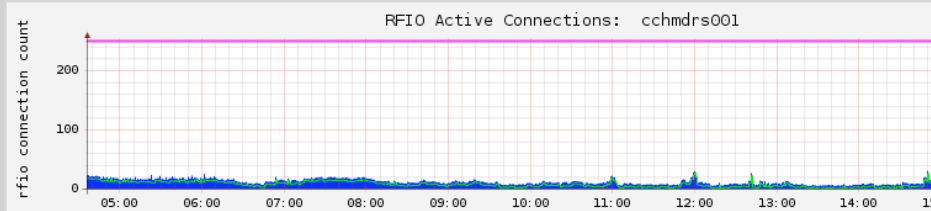
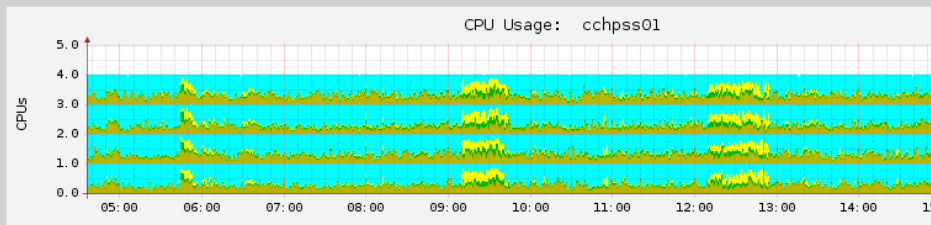
HPSS DASHBOARD

Monday 9 Feb 2009

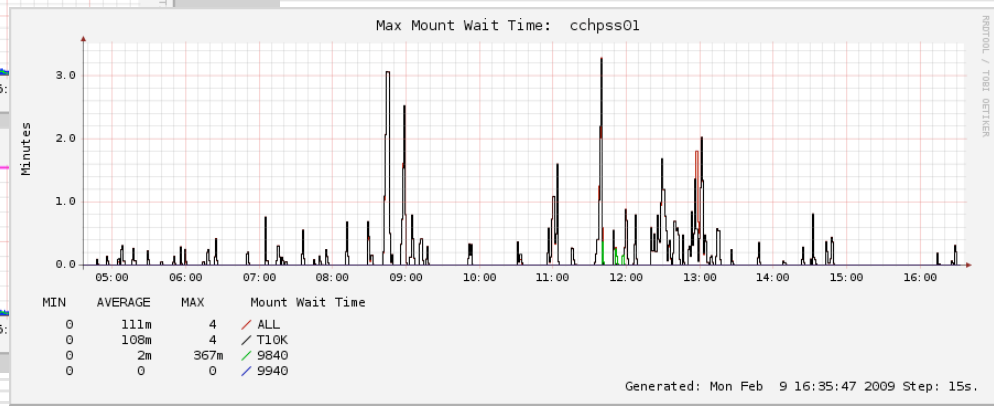
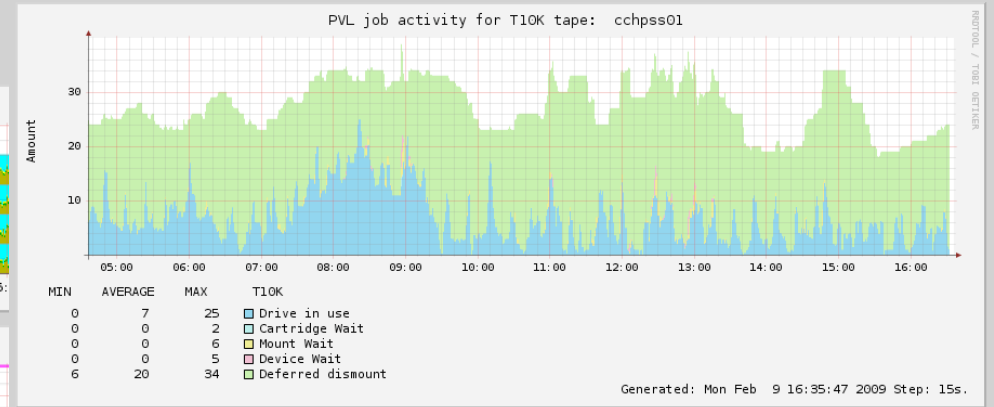
Usefull links

[PVLJobs](#) [List of Drives](#) [Débits HPSS depuis dCache](#) [Débits HPSS depuis dCache par COS](#)

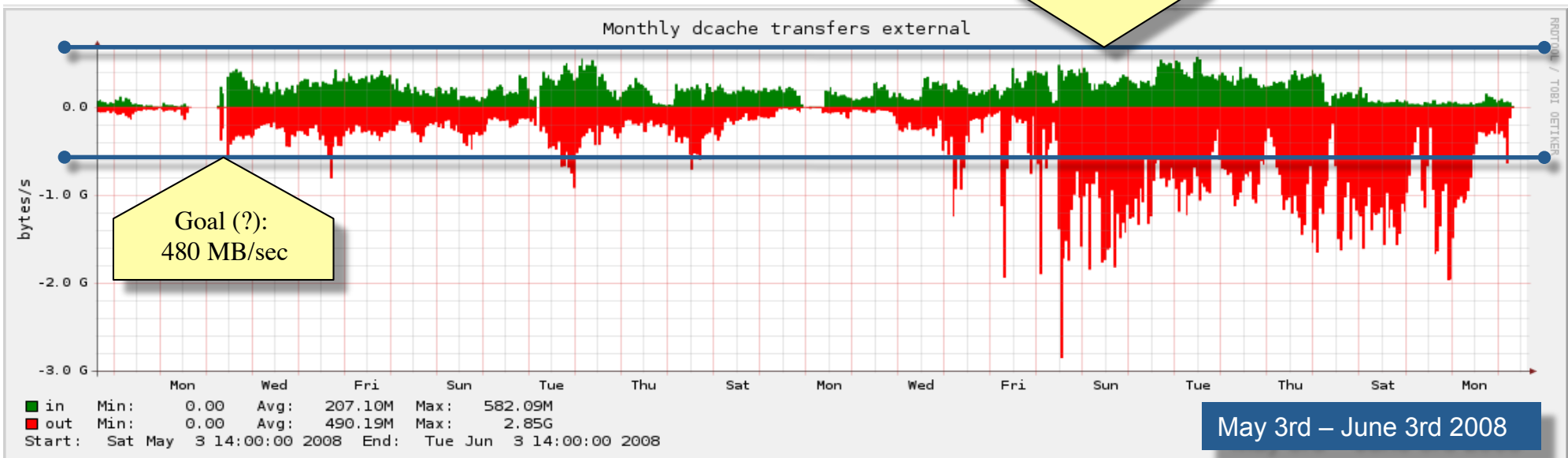
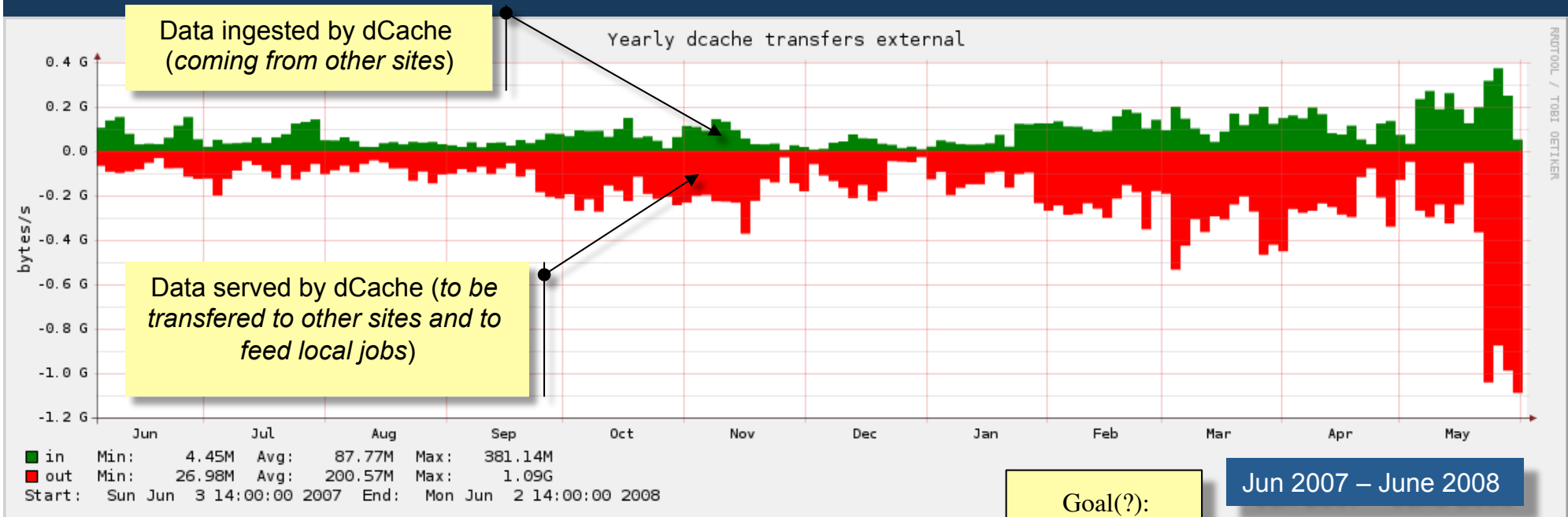
Main Metrics



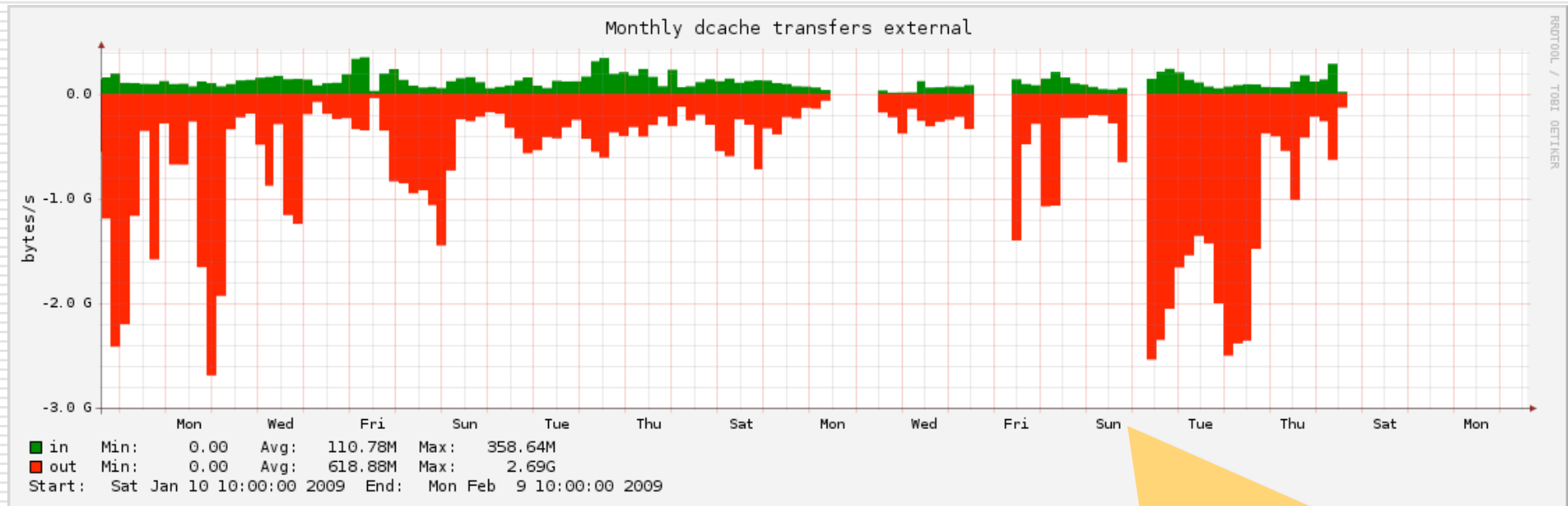
PVL Job Activity



dCache/SRM



dCache/SRM (cont.)



Throughput delivered by dCache increased with respect to what we demonstrated during CCRC08.

Still to demonstrate we can sustain it over longer periods.

dCache/SRM (cont.)

- Improvements in 2008
 - Add a 4th core server
 - Hardware upgrade of all core servers
 - Software upgrade: SL4 64bits, PostgreSQL, Java VM
 - 5 upgrades dCache
 - Approaching 100 file servers in production, managed by dCache
 - Isolation of the infrastructure for data import from tier-0
 - ◆ To make sure we can continue importing data, even in case of major shutdown

dCache/SRM (cont.)

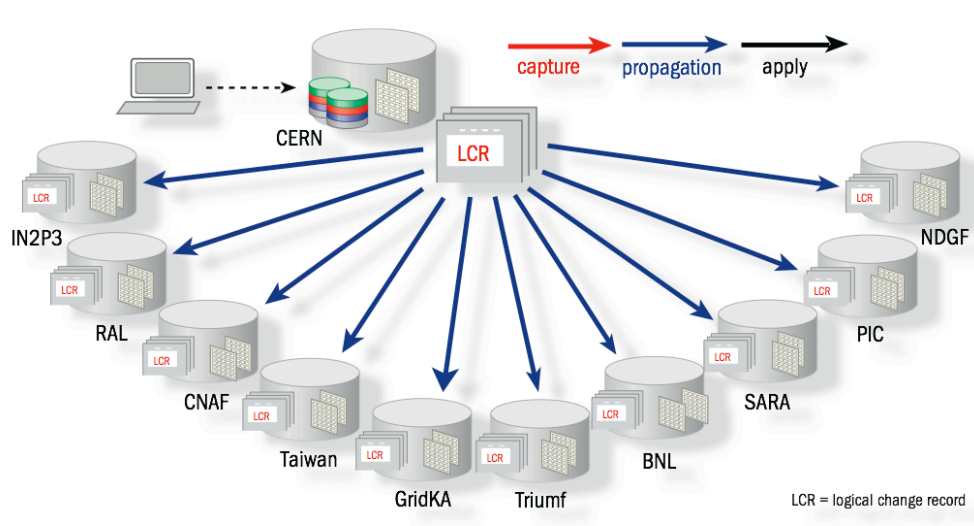
- **Issues**

- Chronic instability of SRM, leading to frequent upgrades of the dCache software
 - ◆ With the unavoidable interruptions of service
- Inconsistency of the dCache internal catalog: ATLAS files missing on disk
 - ◆ Cause of the problem traced back to an modification in the dCache configuration several months ago
 - ◆ Full check of the catalogue performed and complete list of missing files identified: catalogue updated and ATLAS informed of the result
- Core services reached their limits, as a consequence of the increase in the load

dCache/SRM (cont.)

- Perspectives for 2009
 - Migration to Chimera, as a replacement for PFNS (internal catalogue)
 - ◆ Coordinated operation with dCache developers and other tier-1s
 - Separation of meta-data: one instance per experiment
 - Explore the possibility of setting up a configuration consisting of a dedicated SRM server per experiment
 - Improve our understanding of the internals of dCache
 - ◆ Common effort with other tier-1s and dCache developers

Databases



- Setup of 2 dedicated Oracle clusters: one for ATLAS and another for LHCb databases
 - In addition to the one used for generic services (cloud LFC, FTS, ...)

- Redundant hardware setup
 - Both for availability and performance reasons
- Replication from master database at CERN
 - ATLAS: tags, COOL
 - LHCb: read-only LFC

Databases (cont.)

- ATLAS Metadata Interface (AMI)
 - <http://ami.in2p3.fr>
 - Primary physicist interface to the ATLAS metadata and dataset selection catalog
 - Master catalog is hosted and operated at CCIN2P3
 - Database replicated to CERN, by using the Oracle STREAMS technology
 - ◆ The same used by LCG 3D
 - ◆ Replication mechanism in test since more than 1 month
 - Operations of this service is a shared responsibility between the AMI developers at LPSC and CCIN2P3

Data Transfer

- **Upgraded FTS service**
 - v2.0 installed early 2008
 - November 2008: introduction of gridFTP2
 - Introduction of a load-balanced third machine for handling the load: as a result of the 10 million files transfer exercise by ATLAS early 2009
- **Perspectives for 2009**
 - Imminent deployment of FTS v2.1 on SL4 64bits
 - Improve monitoring of the file transfer activity

File Catalogue

- Early 2008: deployment of a new version on SL4 64bits
- ATLAS file catalogue (based on LFC) stores information on the files in the cloud
- Instabilities observed during CCRC08
 - Local turnaround implemented (additional load-balanced machine and watchdog)
 - Software problem identified and fixed by the LFC developers
 - ◆ patch deployed early July 2008
- Perspectives for 2009
 - Intensive testing of failure recovery features already present in the LFC client and not used yet

Software Area

- Software area for LHC experiments on top of AFS
 - Advantages: area shared by all the worker nodes of the site (tier-1 and tier-2), single installation
- On 2007, problems observed due to high load created by thousands of jobs accessing simultaneously the same set of files
 - Also, some misbehaving AFS clients could freeze the servers
- Big effort on 2008 for improving the installation
 - Software update of all the AFS servers and clients
 - Increase of the cache size on clients (now 20 GB)
 - Dedicated AFS servers for some experiments with replication of hot areas (read-only)
 - ♦ ATLAS: 2 dedicated servers + 3 read-only replicas
 - ♦ CMS, LHCb: 1 dedicated server
 - ♦ Alice: 1 shared server
- Perspectives for 2009
 - Renew the hardware for AFS file servers and improve their connectivity

Perspectives for 2009

- Deploy a prototype analysis infrastructure
- Improve the stability of the storage services, in particular dCache/SRM
- Perform the risk analysis, according to the framework proposed by IN2P3
- Monitoring, Monitoring, Monitoring !!!

Conclusions

- Increased project visibility within the WLCG collaboration
- Deployment of resources has been performed as planned
- Several improvements in the operations of all the relevant services have been implemented
- Still, more effort is needed to identify and fix problems related to the stability of the SRM service
- Analysis infrastructure is a priority in 2009

Questions & Comments



Acknowledgements

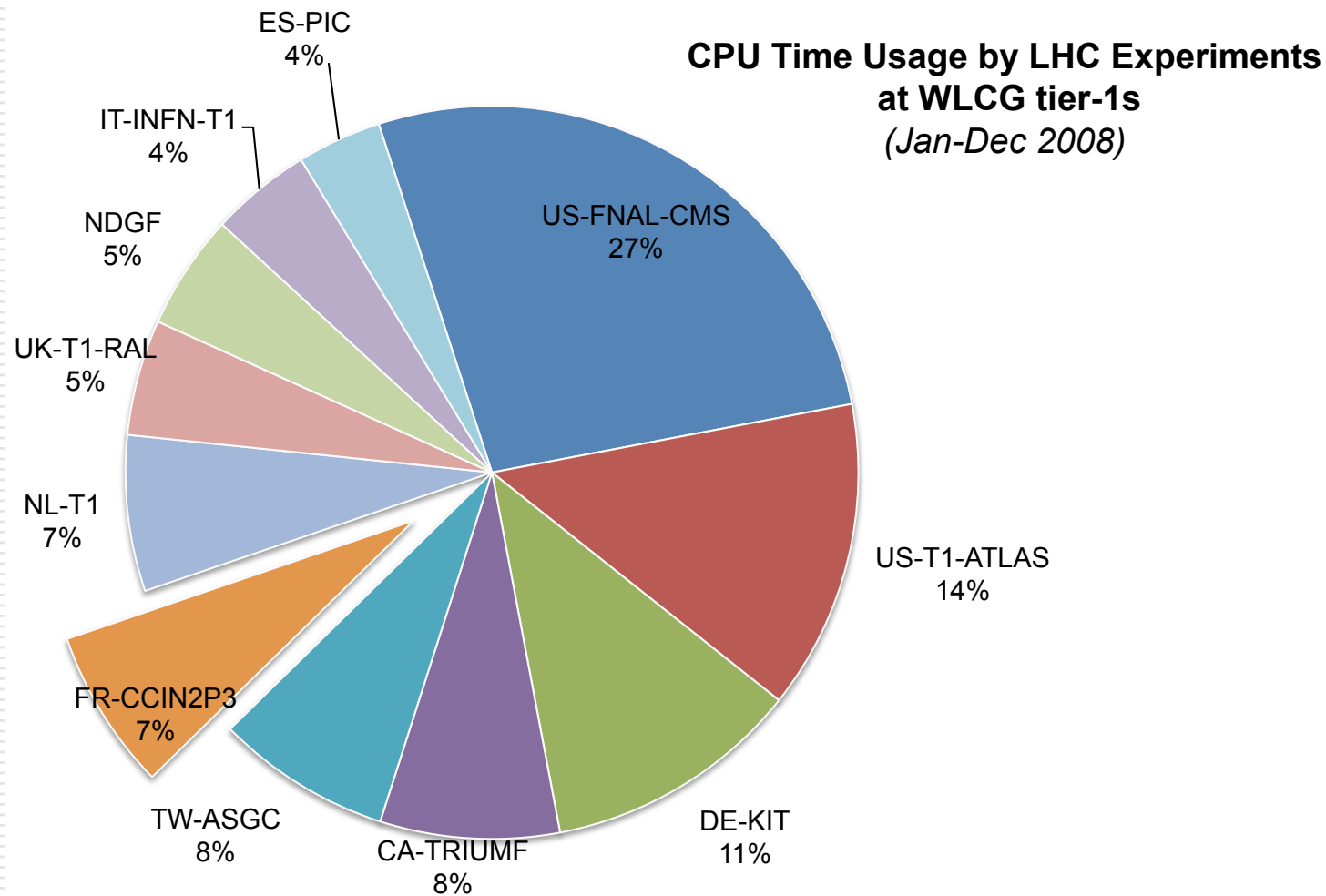
Thanks to all the people who provided input for this talk, including:

Pierre Girard, Lionel Schwarz, Pierre-Emmanuel Brinette, Pierre-Etienne Macchi, Hèlène Cordier, Cyril L'Orphelin, David Bouvet, Xavier Canehan, Christine Durand, Bernard Chambon, Suzanne Poulat, Laurent Caillat-Vallet

Annex

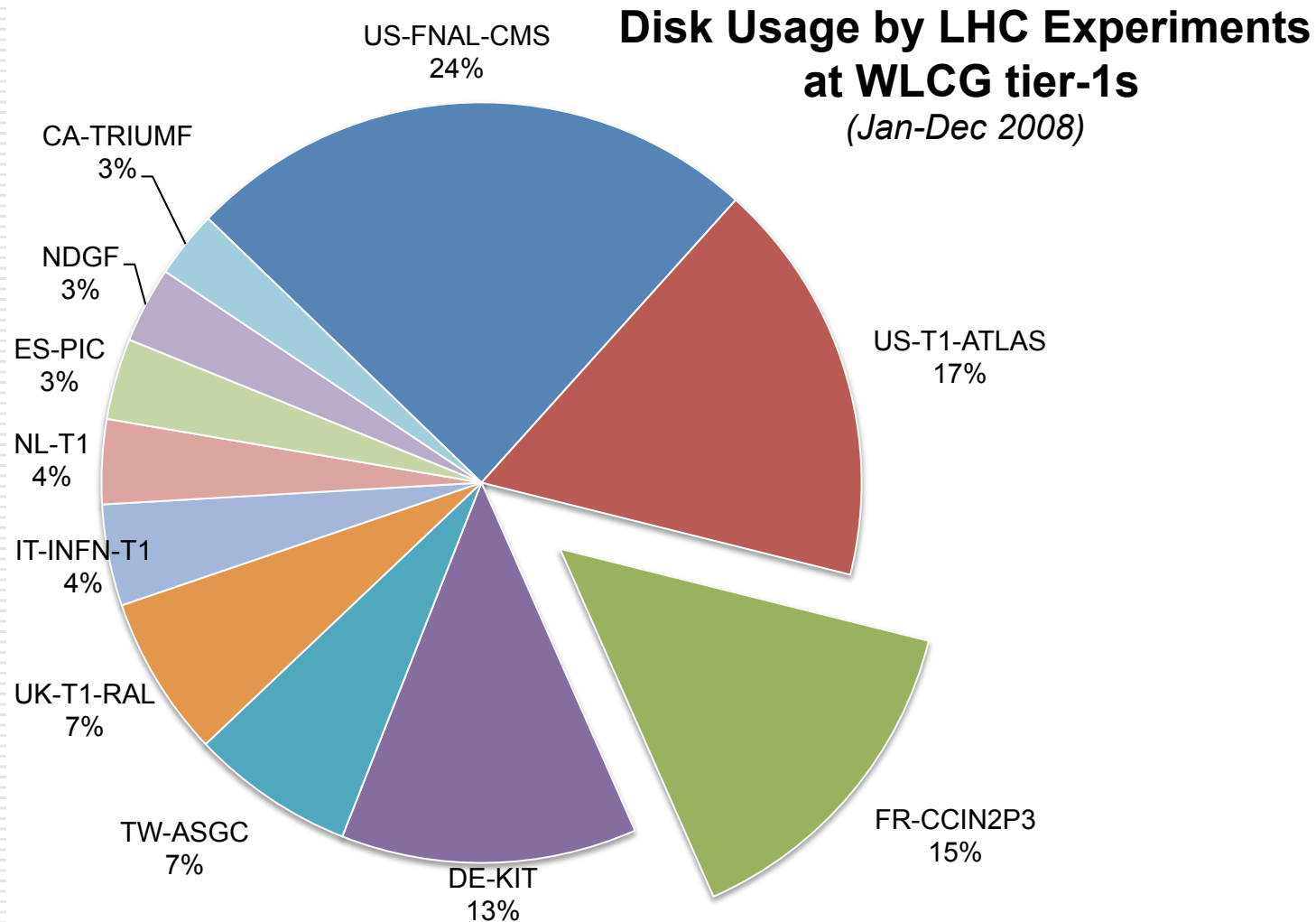
Contribution LCG-France

Contribution by LCG-France (cont.)



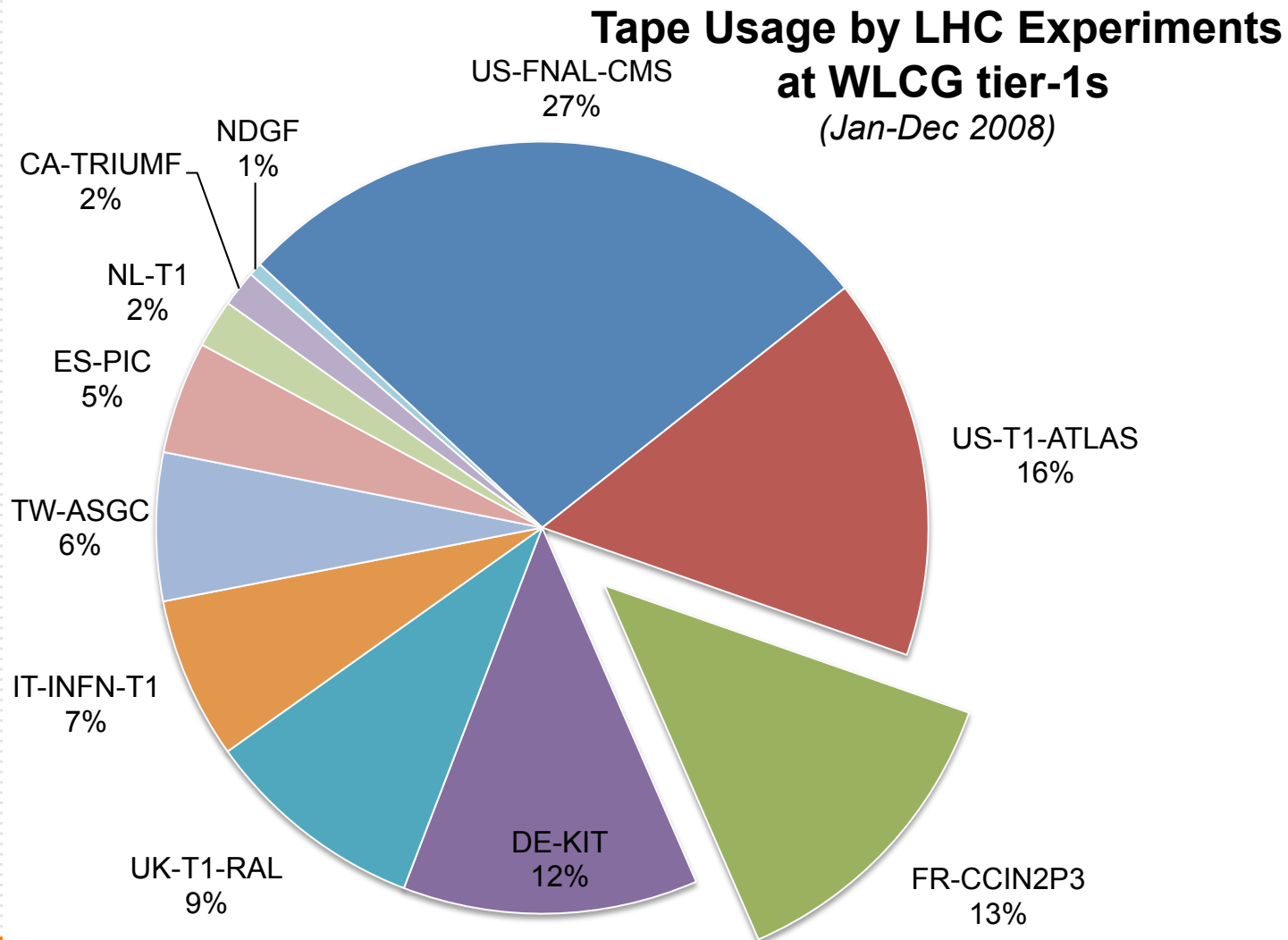
Source: [WLCG CERN & Tier-1 Sites Accounting Report](#)

Contribution by LCG-France (cont.)



Source: [WLCG CERN & Tier-1 Sites Accounting Report](#)

Contribution by LCG-France (cont.)

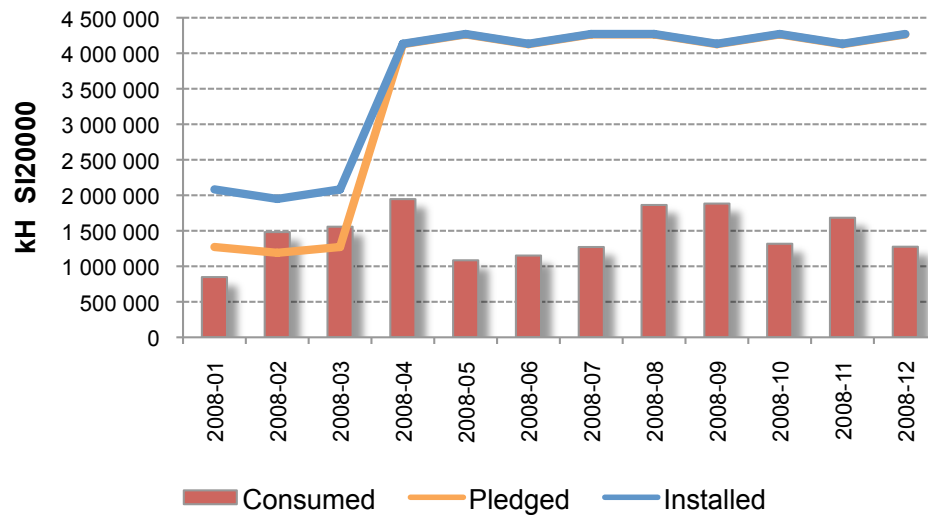


Source: [WLCG CERN & Tier-1 Sites Accounting Report](#)

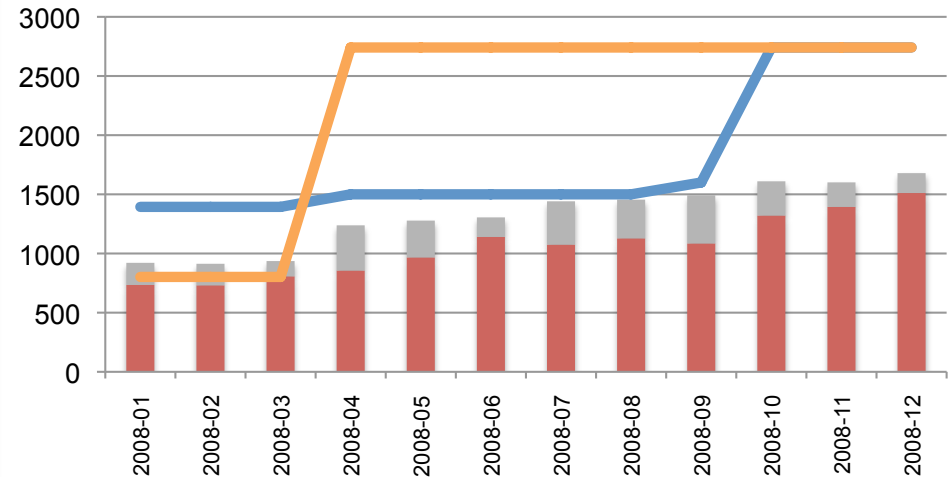
Resource Deployment & Utilisation

Resource Deployment & Utilisation

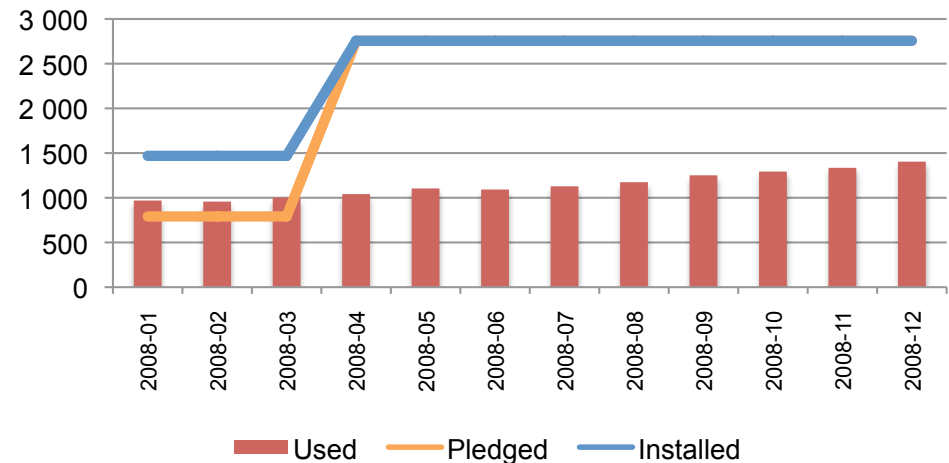
LHC - Consumed vs. Pledged vs. Installed CPU Capacity
(including efficiency factor)



Disk - All LHC Experiments



Tape -- All LHC experiments



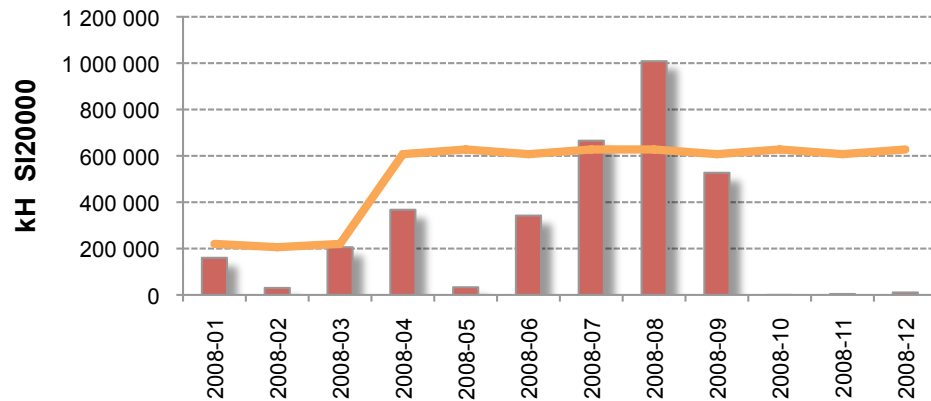
CPU and Tape capacity delivered according to the pledges, on time.

Disk capacity reached the pledges by October.

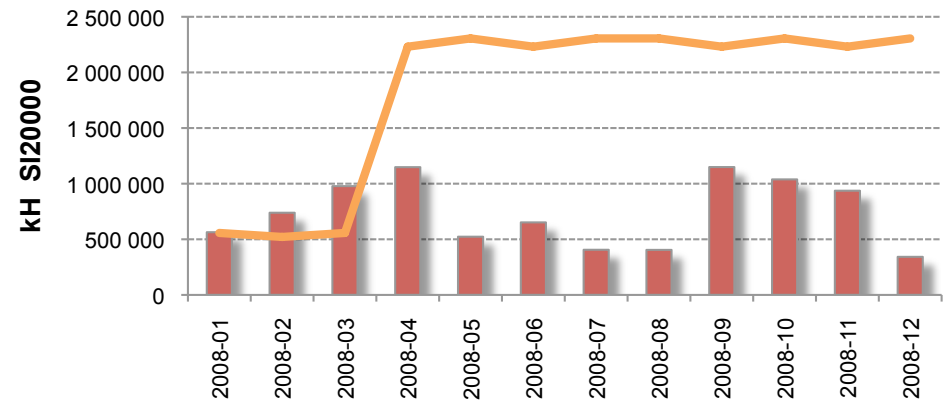


Resource Deployment & Utilisation: CPU

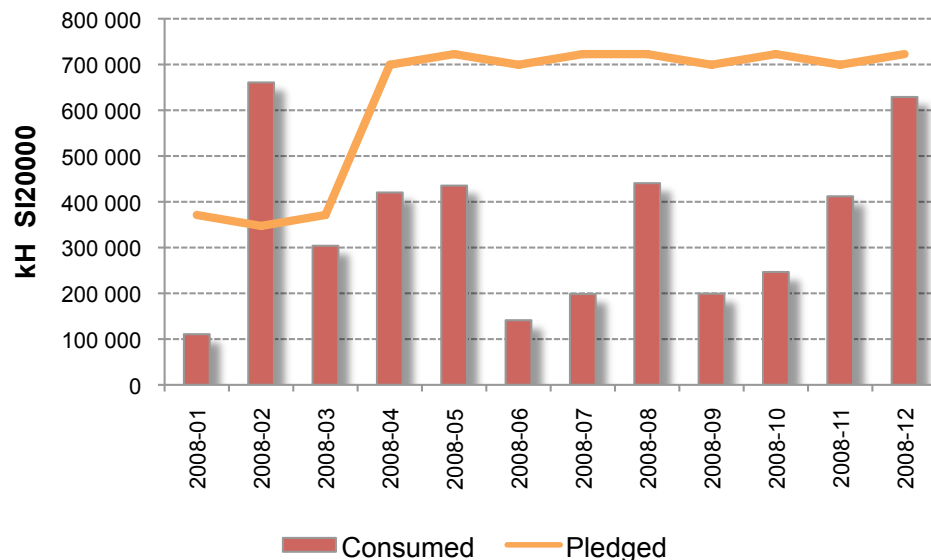
Alice - Consumed vs. Pledged CPU Capacity
(including efficiency factor)



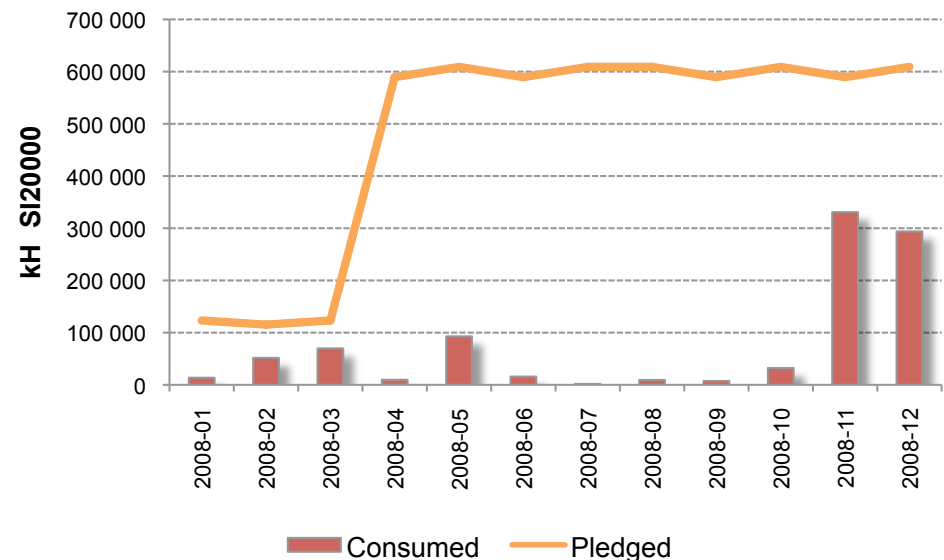
ATLAS - Consumed vs. Pledged CPU Capacity
(including efficiency factor)



CMS -- Consumed vs. Pledged CPU Capacity
(including efficiency factor)

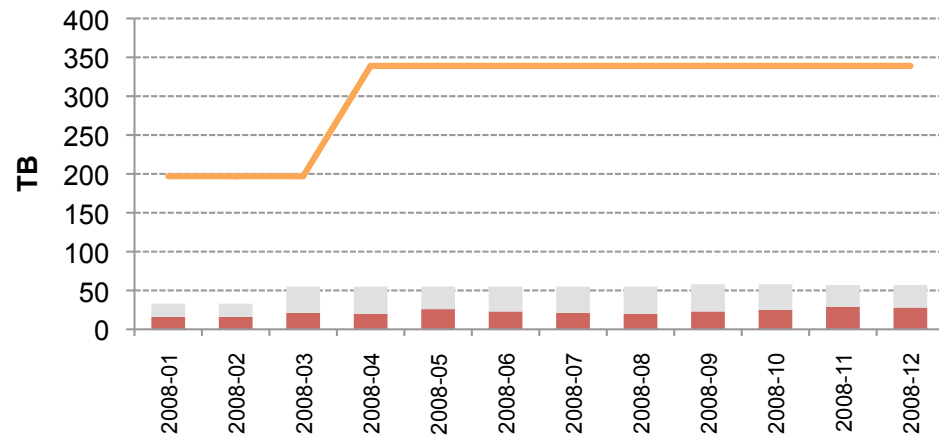


LHCb -- Consumed vs. Pledged CPU Capacity
(including efficiency factor)

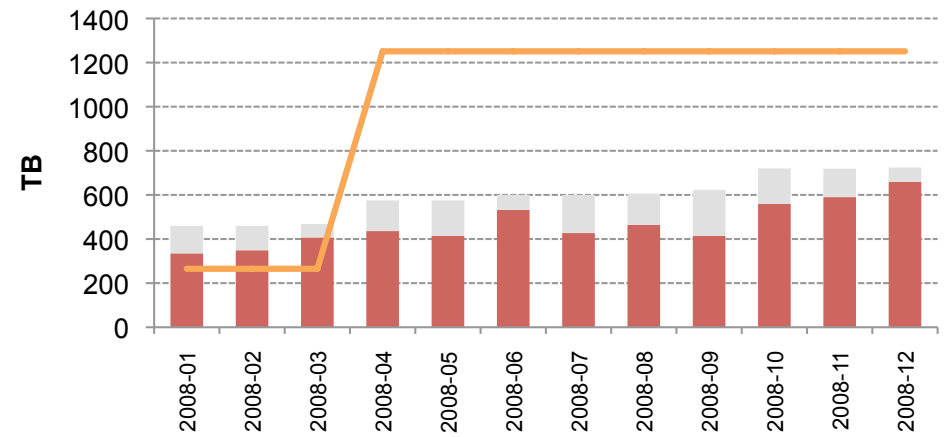


Resource Utilisation: Disk

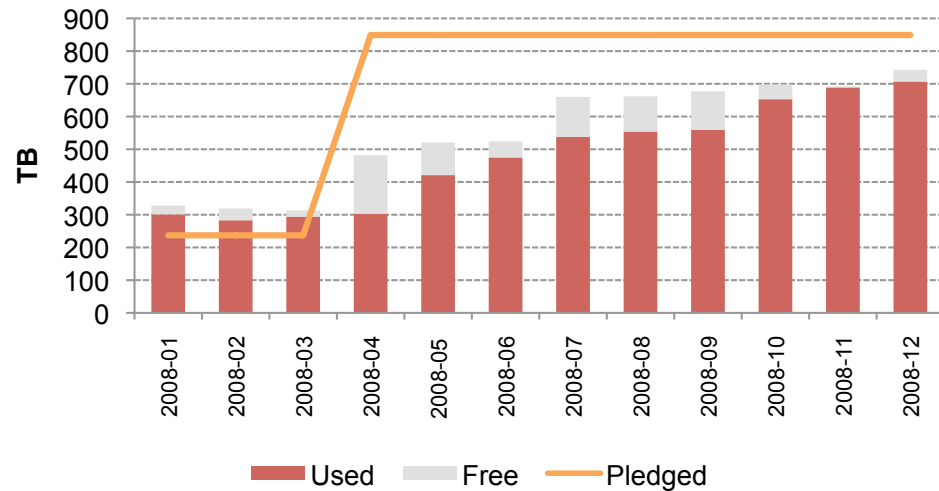
ALICE



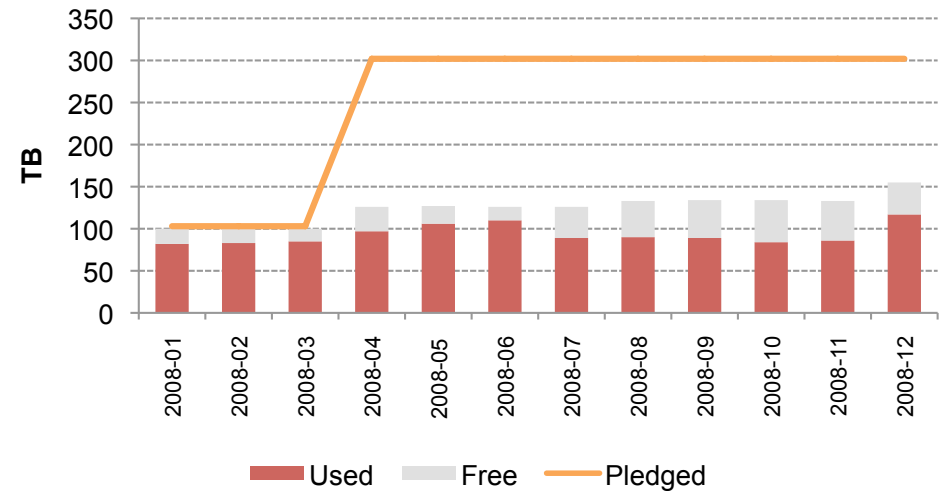
ATLAS



CMS

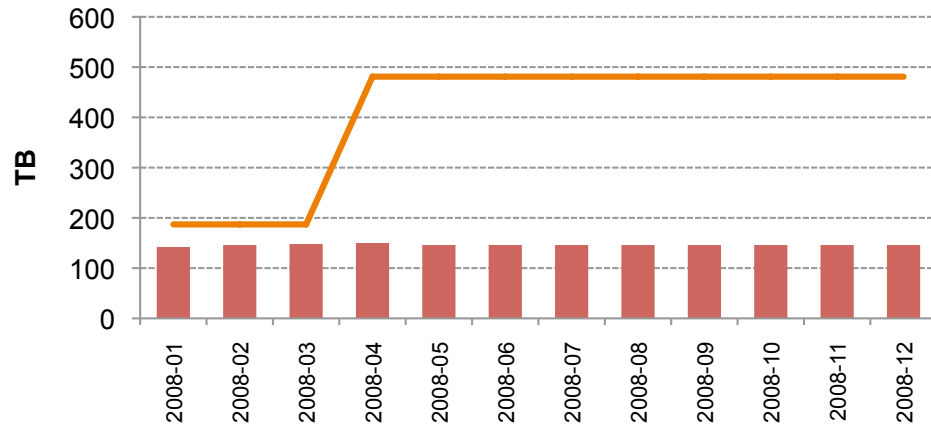


LHCb

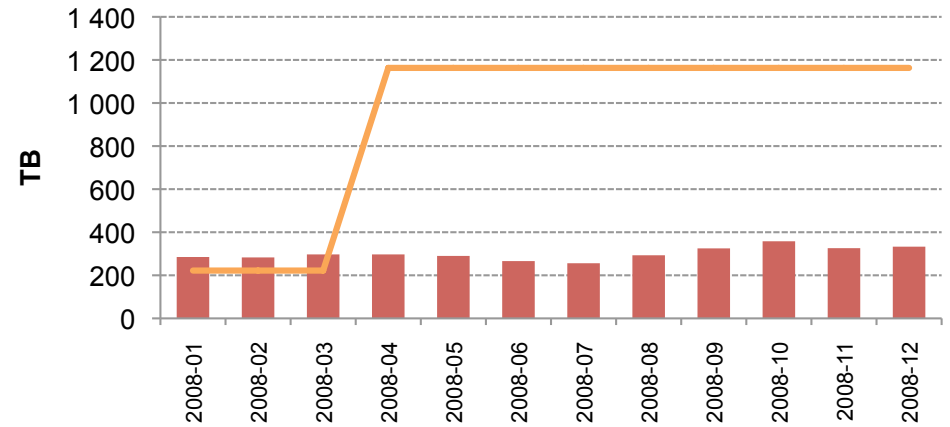


Resource Utilisation: Tape

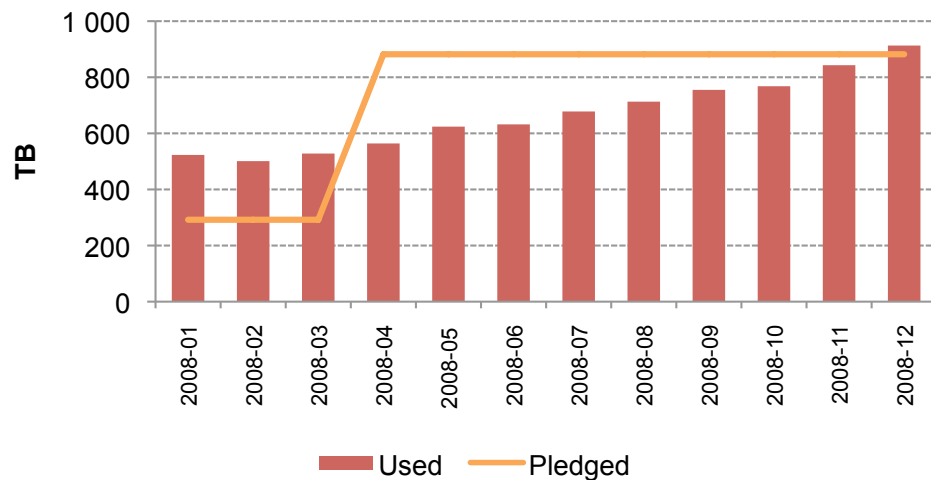
ALICE



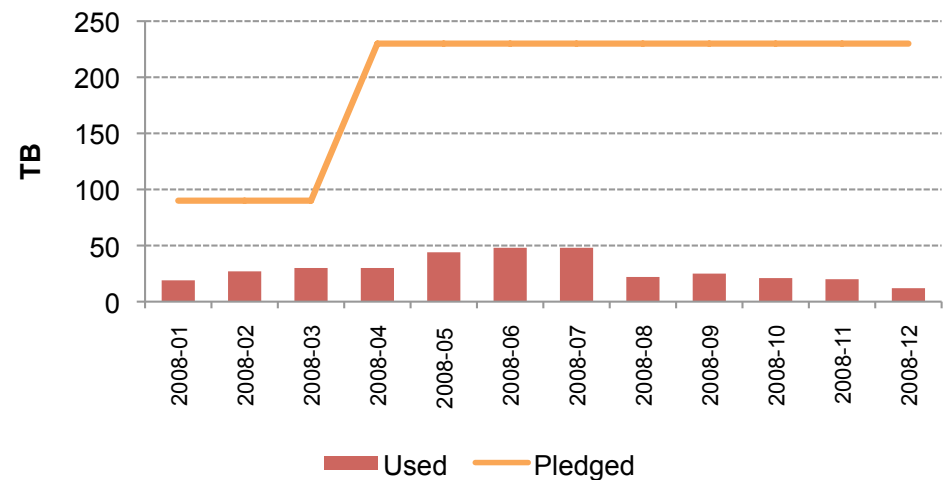
ATLAS



CMS



LHCb



VO Boxes Service Level Agreement

- Formal agreement between CCIN2P3 and the 4 LHC experiments regarding the level of service provided for the VO Boxes operated by the site
 - WLCG High-level milestone
 - Approved by ATLAS, ALICE and LHCb



SERVICE LEVEL AGREEMENT FOR VO BOXES OPERATED BY FR- CCIN2P3

Last updated: 17/12/2008

Document status: DRAFT

Document identifier: <https://edms.in2p3.fr/document/I-014543>

Document version number: 1.3

Abstract:

This document describes the general Service Level Agreement (SLA) between the virtual organizations (VO) and the FR-CCIN2P3 site regarding the operation of VO boxes for the LHC experiments, within the framework of the WLCG project.

DRAFT

1 / 15

Source: <https://edms.in2p3.fr/document/I-014543>