

12/01/2012

LCG-France Tier-1 & AF

Réunion mensuelle de coordination

Pierre Girard

Pierre.girard@in2p3.fr



Les rats quittent le navire...



- Poste de responsable du projet LCG@CC.IN2P3.FR vacant au 1er septembre...
- Poste d'administrateur grille vacant au 1er septembre...
- Poste de développeur accounting grille vacant au 1er septembre...



- Nouvelles de LCG
 - CR du dernier GDB
- Technical Evolution Groups
- Disponibilité du site
- Dossiers en cours
- Événements



Nouvelles de LCG



■ GDB du 8 février 2012

- <http://indico.cern.ch/conferenceDisplay.py?confId=155065>

CR du dernier GDB



Grid Engine SIG

- Philippe Olivero and Rolf Rumler are investigating whether there is interest in forming a group to discuss issues around Grid Engine.
- Support for Grid Engine has moved from Sun to Oracle to U|niva. It has never been very well supported in the glite stack but is used and liked by an increasing number of sites.
- If you are interested contact Philippe
 - philippe.olivero@in2p3.fr
- He will say more next month

WLCG
Worldwide LHC Computing Grid



Science & Technology
Facilities Council

GDB

CR du dernier GDB



Today

- TEG
- Site Dashboards
- EMI
- Storage Accounting



WLCG
Worldwide LHC Computing Grid



Science & Technology
Facilities Council

GDB

CR du dernier GDB



EUGridPMA plan

- The EUGridPMA discussed the matter at their Jan meeting
 - <https://www.eugridpma.org/meetings/2012-01/summary.txt>
- Their current plan:
 - “The date by which SHA-2 production certs may be issued will be **NO LATER** than **January 2013** (and it is likely we will **RECOMMEND** CAs to move then, since it will take another 395 days to get rid of SHA-1 in a reasonable way)”
- This would give us 10 months to get ready for RFC and SHA-2
 - See next slide
- The TAGPMA will discuss the matter

Need MW to switch to RFC proxies
We have summer & autumn to deal with the bugs. Prepare for major production upgrades early 2013



Technical Evolution Groups



■ Mardi 7 février au CERN

- <http://indico.cern.ch/conferenceDisplay.py?confId=158775>

■ Constat

- Often more status quo / known defect discussions than a strategy one so far
- Often we know better in which direction to leave than where we want to arrive...
- Most of the strategy proposals originate from s/w providers

■ La suite prévue

- There is a need for an overall architecture diagram and concrete recommendations.
- Should document (the architecture of) both what we have now and what we want in the future."

■ Data placement responsibilities

- Site: Storage Management
 - Stable storage for placed data
 - Detection/Repairing of Mismatch between SE disk and SE catalogue (dark data)
- Experiments: Data Management
 - Geographical dataset distribution
 - Mismatch/Repairing between SE and experiment catalogue (grey data)
 - DM solution are used to overcome storage management problems

■ Catalogs

- Abandon programmé de LFC dans les 2 prochaines années
- Solution spécifique par VO

■ Data placement & Federations

- Push via FTS/xroot
- Now complemented with federated access via xroot to reduce impact of (temporarily or permanently) unavailable data. Relevance increasing towards smaller sites.
- Federation in WLCG:
 - Alice, US Atlas and US CMS
 - Xrootd used as federation infrastructure with mixed storage infrastructures: dCache, DPM, GPFS, HDFS, and native xrootd
- Protocoles considérés : xrootd, NFS v4 et HTTP

■ WAN Protocols

- Fonctionnalités attendues
 - third party copy, data integrity (checksum), efficiency (parallel streams), stable/dependable server/client implementations
 - Should be compatible with storage federation
- Candidats possibles
 - gridFTP, Xrootd, http(s)/WebDav, NFS-4.1/pNFS, S3
 - https considered as very promising

TEG Storage/Data Management (4)



■ SRM + Cloud discussions

- Problèmes de SRM dont son coût de maintenance
- Discussion “Cloud vs SRM”
 - Unclear if Cloud provides any benefit
 - Should consider USA's push toward virtualisation
- WebDav considéré aussi
 - It's a standard. Lots of clients, covering all platforms
 - Support in SEs coming with EMI-2.
- Need to separate SRM into **core functionality blocks**
 - Transfer management, Interacting with namespace, Aggregated space querying, Storage management
- Use this breakdown to compare alternatives
 - Analyse Cloud APIs and WebDAV
 - Identify their **limitations**

TEG Workload Management (1)



- Main areas tracked so far
 - Commonalities with pilots and frameworks
 - Streamed submission
 - Support of jobs requiring whole nodes or multiple cores
 - specifying whole node / multi-core requirements in the JDL/RSL
 - SGE support for these features in CREAM (provided in EMI by CESGA) is currently missing (?)
 - Support of CPU affinity and CPU sets
 - CPU Pinning, LRMS solution, SL6...
 - Tagging I/O vs. CPU intensive jobs
 - We propose to introduce a new parameter to specify whether a job is CPU- or I/Obound. The parameter is a scalar in the range [0,1] (0=CPU-bound; 1=I/O bound).
 - Requirements for a CE service
 - define a “CE alias” with a pool of individual CE’s behind the alias, where the individual CE’s share state data
 - Summary of the use of the IS by WLCG experiments

TEG Workload Management (2)



■ Use of virtualization technologies

- <https://twiki.cern.ch/twiki/bin/view/LCG/WLCGWIMgmtTEGVirtualization>
- Virtualization penalties must be carefully considered...
- Static virtualization of batch resources vs dynamic allocation of VMs
 - e.g. of automating deployment of intrusive updates, or of optimizing allocation and encapsulation of multi-core VMs used for multi-core jobs.
- Scalable mechanism, possibly linked to an LRMS, is required to properly handle dynamic allocations.
 - Early tests suggest that both e.g. Torque and LSF may have troubles scaling to several (tens of) thousands of VMs, at least if not properly configured.

TEG Workload Management (3)



■ Cloud computing

- <https://twiki.cern.ch/twiki/bin/view/LCG/WLCGWIMgmtTECloudComputing>
- VO's already expressed interest in accessing IaaS entry points to sites.
- Batch cloud factory
 - Adopt HEPiX recommendations (e.g. wrt VM capabilities) and consider the EGI policy for what regards running VMs.
 - Define common methods to access resources (e.g. via EC2 or OCCl) and common sizes to ensure consistency and avoid effort duplication; track possible integration into an IS.
 - Explore integration into an LRMS and with CREAM for the provisioning of cloud instances.
 - Track AuthN/AuthZ developments and integration of multiple cloud sites.
- Early tests on Cloud instantiations at WLCG sites include CERN's Ixcloud and INFN's WNoDeS.
 - participating to the EGI distributed cloud TF



WLCG TEG Workshop - 7 February 2012

Production and Data Mgt Databases (2)

- The ATLAS DDM (Data Management) team investigated many NoSQL options and found that Hadoop/Hbase is a very good solution for the accounting tools
 - A "private" Hadoop cluster of 12 nodes is now in production; data are imported daily from Oracle production DB
 - Accounting tools: 8 hours in Oracle, 25 mins in Hadoop/SLC5, 6 mins in Hadoop/SLC6
 - Comparing an Oracle application that was optimized over the course of one year together with IT-DB with a straightforward one-week Hadoop implementation
 - But Oracle application running on a busy cluster vs Hadoop on dedicated h/w
 - Hadoop MapReduce vs Oracle SQL
- Similarly, the ATLAS Panda (Distributed Production and Analysis) and TDAQ teams found Cassandra useful for storing operations logs and running monitoring applications
 - Both under test
- CMS has been using CouchDB and MongoDB in production tools on VOboxes managed jointly by CMS and CERN-IT.
- Deploying and running these services has not been a burden to the operations team, in part because the scope of these tools was defined appropriately, and their deployment is considered a "quite successful scale-out experience" by the operators

12



■ TEG Operations

- <http://indico.cern.ch/materialDisplay.py?contribId=5&materialId=slides&confId=158775>
- Draft version of report at
 - <https://twiki.cern.ch/twiki/bin/view/LCG/WLCGTEGOperations>

■ TEG Secutity

- <http://indico.cern.ch/materialDisplay.py?contribId=4&materialId=slides&confId=158775>
- Only serious option to take for the short-term is ID switching with gLExec
- **GlideinWMS** already handles the ID switch transparently via Condor

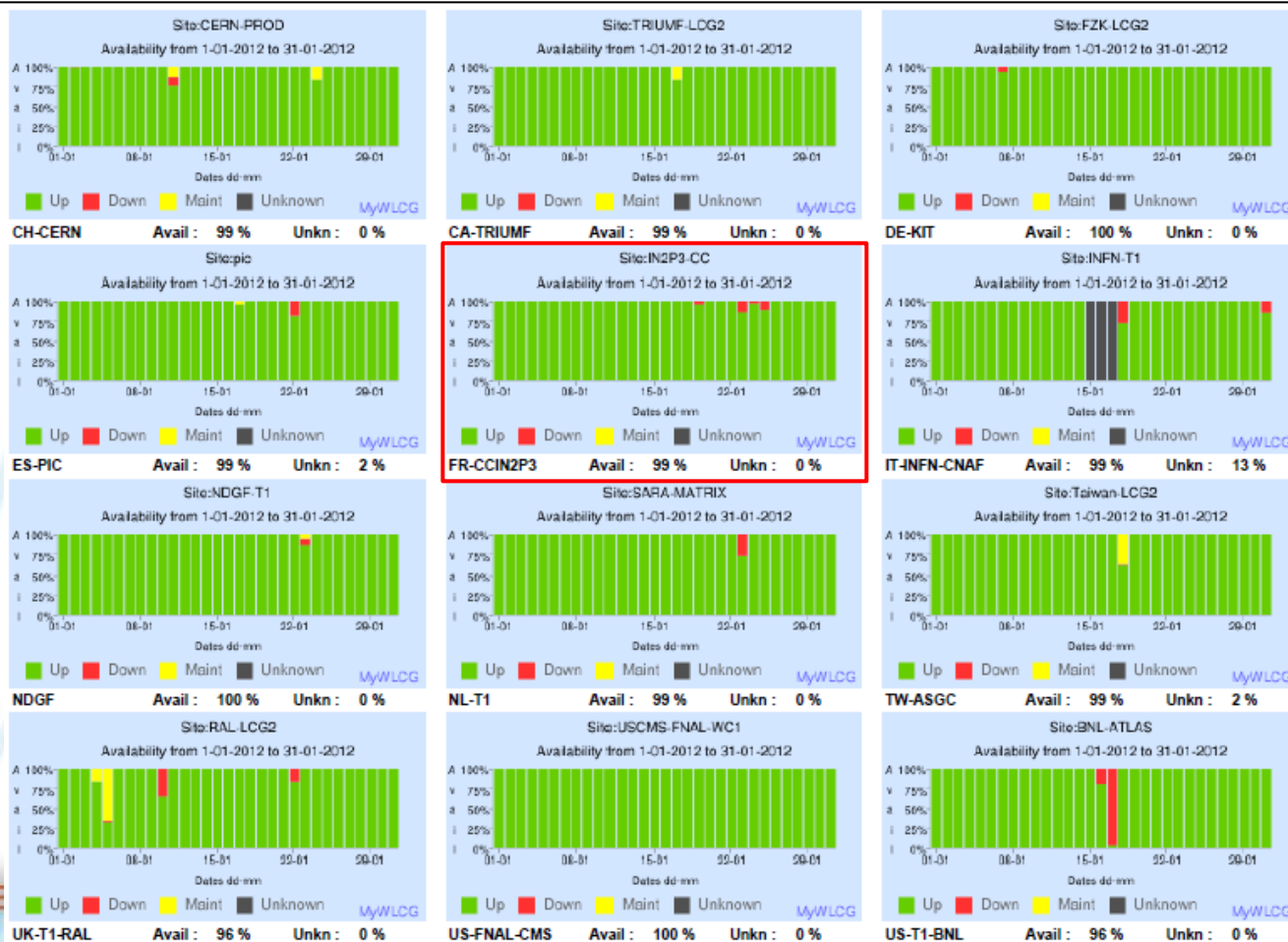


Résultats du site

Résultats Janvier 2012



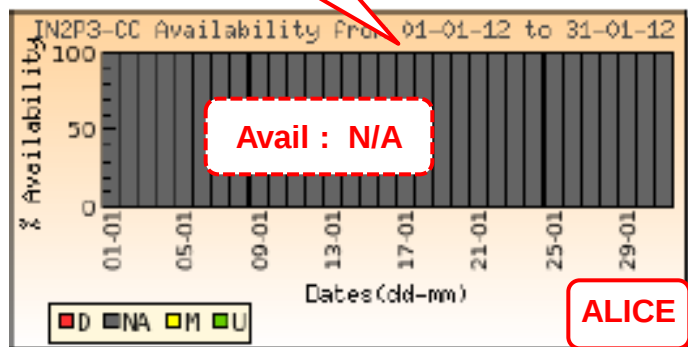
■ <https://espace.cern.ch/WLCG-document-repository/ReliabilityAvailability/Forms/AllItems.aspx>



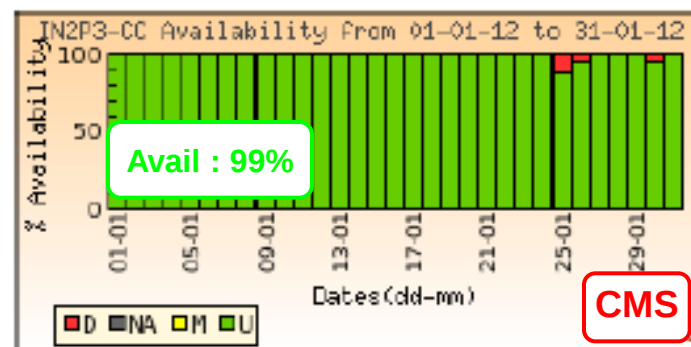
Disponibilités des VOs pour janvier 2012



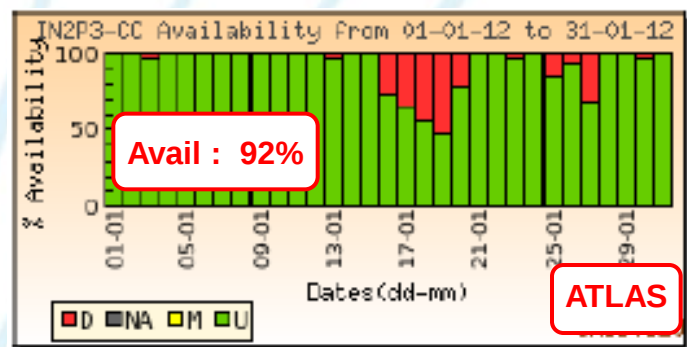
Fin des LCG-
CEs et pas
de SE



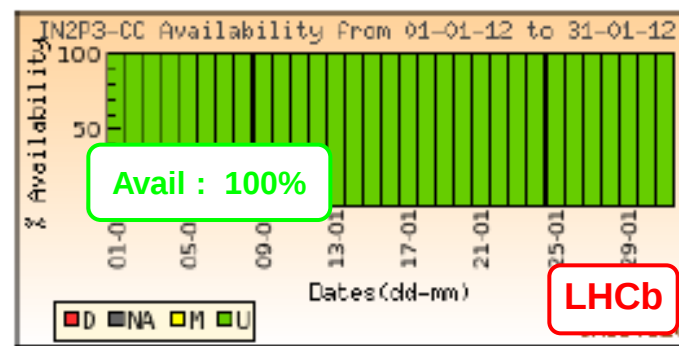
FR-CCIN2P3 Avail: N/A Unkn: 100 %



FR-CCIN2P3 Avail: 99 % Unkn: 0 %



FR-CCIN2P3 Avail: 92 % Unkn: 0 %



FR-CCIN2P3 Avail: 100 % Unkn: 0 %



Dossiers en cours

Pb d'accounting GE



**Décisionnel,
portail CESGA et
accounting LCG
faux depuis juin
2011.**

Mozilla Firefox

Eichier Édition Affichage Historique Marque-pages Outils 2

VidyoConferencing x VidyoConferencing x https://grid.in2p3...6.days.2011.html x +

in2p3.fr https://grid.in2p3.fr/LCGFrAccounting/CPU/2011/LCG-T1-total.hs06.days.2011.html

Les plus visités Mappy - plans, itinérail... Prévisions météo de M... MySQL :: MySQL 5.0 R... Indico [Grid Deploye... French Translation of ... FR-cloud Regional Cen...

CPU Accounting of CCIN2P3-T1 for 2011

			2011/01	2011/02	2011/03	2011/04	2011/05	2011/06	2011/07	2011/08	2011/09	2011/10	2011/11	2011/12
BQS	alice	cpu	65 022,07	41 831,14	52 476,98	64 281,44	14 267,13	34 048,52	77 987,42	49 450,17	46 631,39	89 583,75	49 559,68	0,00
		wall	91 486,71	73 480,99	86 167,35	110 039,09	37 315,14	116 091,70	136 161,47	99 978,13	57 277,88	98 892,08	53 515,68	0,00
	atlas	cpu	525 073,13	296 595,13	465 412,44	433 984,68	553 718,66	1 030 928,52	663 798,66	529 415,65	322 432,49	420 506,02	243 727,67	0,00
		wall	586 931,25	391 341,16	595 985,94	591 878,21	813 381,14	1 169 816,13	824 136,64	760 730,58	390 175,84	457 411,15	264 079,71	0,00
	cmsf	cpu	207 502,19	159 341,37	223 040,78	167 675,69	195 049,52	117 146,27	85 729,69	122 481,45	163 515,75	202 707,49	78 828,99	0,00
		wall	279 742,28	202 359,23	272 925,11	226 259,08	227 445,67	140 373,17	96 577,14	137 821,42	190 685,63	272 218,52	84 500,42	0,00
	lhcb	cpu	124 287,13	113 709,80	96 790,56	36 877,90	179 805,91	110 743,48	129 693,10	153 701,21	112 425,26	102 192,60	56 902,71	0,00
		wall	136 794,75	124 116,78	110 064,00	44 221,72	208 989,03	137 950,07	156 434,71	181 472,15	126 087,97	117 225,94	64 275,71	0,00
	TOTAL	cpu	921 884,52	611 477,44	837 720,76	702 819,71	942 841,22	1 292 866,79	957 208,87	855 048,48	645 004,89	814 989,86	429 019,02	0,00
		wall	1 094 954,99	791 298,16	1 065 142,40	972 398,10	1 287 130,98	1 564 231,07	1 213 309,96	1 180 002,28	764 227,32	945 747,69	466 371,52	0,00
GE	alice	cpu	0,00	0,00	0,00	0,00	0,00	0,00	0,00	62 692,84	117 129,90	148 705,36	151 704,48	266 532,59
		wall	0,00	0,00	0,00	0,00	0,00	0,00	103 891,28	229 701,53	176 782,40	169 752,62	301 721,75	238 578,04
	atlas	cpu	0,00	0,00	0,00	0,00	0,00	2 563,17	192 466,97	243 808,78	489 407,40	481 329,52	720 552,50	868 999,20
		wall	0,00	0,00	0,00	0,00	0,00	2 743,92	201 104,10	268 231,11	567 585,25	541 734,09	810 263,36	956 383,21
	cmsf	cpu	0,00	0,00	0,00	0,00	0,00	0,00	30 835,18	103 457,03	138 579,50	207 678,37	142 397,46	234 673,41
		wall	0,00	0,00	0,00	0,00	0,00	0,00	33 613,51	115 673,68	156 849,12	285 912,91	157 836,48	277 639,53
	lhcb	cpu	0,00	0,00	0,00	0,00	0,00	0,00	19 836,54	22 355,74	64 492,90	198 003,07	190 505,42	232 496,70
		wall	0,00	0,00	0,00	0,00	0,00	0,00	20 968,14	25 649,15	69 774,47	244 828,51	217 693,51	250 440,36
	TOTAL	cpu	0,00	0,00	0,00	0,00	0,00	2 563,17	305 831,53	486 751,45	841 185,16	1 038 715,44	1 319 987,97	1 522 448,32
		wall	0,00	0,00	0,00	0,00	0,00	2 743,92	359 577,03	639 255,47	970 991,24	1 242 228,13	1 487 515,10	1 723 041,14
ALL	alice	cpu	65 022,07	41 831,14	52 476,98	64 281,44	14 267,13	34 048,52	140 680,26	166 580,07	195 336,75	241 288,23	316 092,24	186 279,01
		wall	91 486,71	73 480,99	86 167,35	110 039,09	37 315,14	116 091,70	240 052,75	329 679,66	234 060,28	268 644,70	355 237,43	238 578,04
	atlas	cpu	525 073,13	296 595,13	465 412,44	433 984,68	553 718,66	1 033 491,69	856 265,63	773 224,43	811 839,89	901 835,54	964 280,17	868 999,20
		wall	586 931,25	391 341,16	595 985,94	591 878,21	813 381,14	1 172 560,05	1 025 240,74	1 028 961,69	957 761,09	999 145,24	1 074 343,07	956 383,21
	cmsf	cpu	207 502,19	159 341,37	223 040,78	167 675,69	195 049,52	117 146,27	116 564,87	225 938,48	302 095,25	410 385,86	221 226,45	234 673,41
		wall	279 742,28	202 359,23	272 925,11	226 259,08	227 445,67	140 373,17	130 190,65	253 495,10	347 534,75	558 131,43	242 336,90	277 639,53
	lhcb	cpu	124 287,13	113 709,80	96 790,56	36 877,90	179 805,91	110 743,48	149 529,64	176 056,95	176 918,16	300 195,67	247 408,13	232 496,70
		wall	136 794,75	124 116,78	110 064,00	44 221,72	208 989,03	137 950,07	177 402,85	207 121,30	195 862,44	362 054,45	281 969,22	250 440,36
	CPU Total	cpu	921 884,52	611 477,44	837 720,76	702 819,71	942 841,22	1 295 429,96	1 263 040,40	1 341 799,93	1 486 190,05	1 853 705,30	1 749 006,99	1 522 448,32
		wall	1 094 954,99	791 298,16	1 065 142,40	972 398,10	1 287 130,98	1 566 974,99	1 572 886,99	1 819 257,75	1 735 218,56	2 187 975,82	1 953 886,62	1 723 041,14

<https://grid.in2p3.fr/LCGFrAccounting/CPU/2011/>

Pb d'accounting GE



Décisionnel,
portail CESGA et
accounting LCG
faux depuis juin
2011.

Mozilla Firefox

https://grid.in2p3.fr/LCGFrAccounting/CPU/2011/LCG-T1-total.hs06.days.2011.html

CCIN2P3-T1 for 2011

CPU used - HEPSPEC06-days	2011/01	2011/02	2011/03	2011/04	2011/05	2011/06	2011/07	2011/08	2011/09	2011/10	2011/11	2011/12
alice	cpu	65 022,07	41 831,14	52 476,98	64 281,44	14 267,13	34 048,52	77 987,42	49 450,17	46 631,39	89 583,75	49 559,68
	wall	91 486,71	73 480,99	86 167,35	110 039,09	37 315,14	116 091,70	136 161,47	99 978,13	57 277,88	98 892,08	53 515,68
BQS												
GE												
TOTAL	cpu	0,00	0,00	0,00	0,00	0,00	2 563,17	305 831,53	486 751,45	841 185,16	1 038 715,44	1 319 987,97
	wall	0,00	0,00	0,00	0,00	0,00	2 743,92	359 577,03	639 255,47	970 991,24	1 242 228,13	1 487 515,10
ALL	alice	cpu	65 022,07	41 831,14	52 476,98	64 281,44	14 267,13	34 048,52	140 680,26	166 580,07	195 336,75	241 288,23
	wall	91 486,71	73 480,99	86 167,35	110 039,09	37 315,14	116 091,70	240 052,75	329 679,66	234 060,28	268 644,70	355 237,43
	atlas	cpu	525 073,13	296 595,13	465 412,44	433 984,68	553 718,66	1 033 491,69	856 265,63	773 224,43	811 839,89	901 835,54
	wall	586 931,25	391 341,16	595 985,94	591 878,21	813 381,14	1 172 560,05	1 025 240,74	1 028 961,69	957 761,09	999 145,24	1 074 343,07
	cmsf	cpu	207 502,19	159 341,37	223 040,78	167 675,69	195 049,52	117 146,27	116 564,87	225 938,48	302 095,25	410 385,86
	wall	279 742,28	202 359,23	272 925,11	226 259,08	227 445,67	140 373,17	130 190,65	253 495,10	347 534,75	558 131,43	242 336,90
	lhcb	cpu	124 287,13	113 709,80	96 790,56	36 877,90	179 805,91	110 743,48	149 529,64	176 056,95	176 918,16	300 195,67
	wall	136 794,75	124 116,78	110 064,00	44 221,72	208 989,03	137 950,07	177 402,85	207 121,30	195 862,44	362 054,45	281 969,22
	CPU Total	cpu	921 884,52	611 477,44	837 720,76	702 819,71	942 841,22	1 295 429,96	1 263 040,40	1 341 799,93	1 486 190,05	1 853 705,30
	wall	1 094 954,99	791 298,16	1 065 142,40	972 398,10	1 287 130,98	1 566 974,99	1 572 886,99	1 819 257,75	1 735 218,56	2 187 975,82	1 953 886,62

Préoccupation pour 2012:
migration (et reprise) de l'accounting
grille vers la nouvelle infrastructure
d'accounting de WLCG/EGI ?

<https://grid.in2p3.fr/LCGFrAccounting/CPU/2011/>

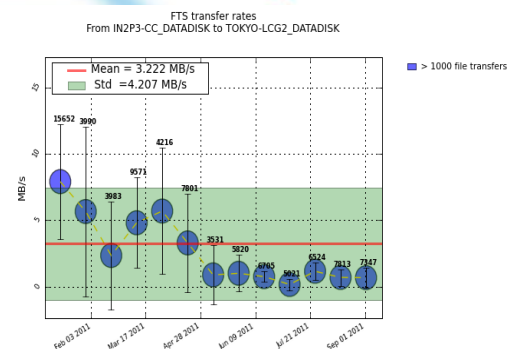


- CREAM GEs
 - Cf. talk des cemasters

Chantiers en cours



- Problèmes réseaux (update)
 - Avec les sites étrangers qui ne sont pas sur LHCOPN
 - Impact sur ATLAS et CMS
 - https://ggus.eu/ws/ticket_info.php?ticket=75983
 - https://ggus.eu/ws/ticket_info.php?ticket=74268
- Reconfiguration du LAN en janvier
 - Amélioration des transferts T2s->CCIN2P3
- Dernier update des tickets le 26/01 par Jérôme
 - We put in production a second 10Gb/s interface provided by RENATER in order to separate LHCONe and generic IP traffics.
Let's see the evolution of our transfers on the generic IP network.
 - Des progrès notés ?





Chantiers en cours



■ Important UPGRADE/UPDATE de dCache

– Documentés

- <https://cctools.in2p3.fr/wiki/exploitation:arrets:arret20120207>
- <http://cctools.in2p3.fr/elog/dCache/910>

– Principaux changements effectués

- Postgres 9.1.
- 2 core servers
 - chimera sur ccdcamcli04 + SRM sur ccdcamcli05
 - 96GB de RAM, disques SSD 200GB en RAID-10
- Passage a la deuxième golden release (aka 1.9.12)

– Qques problèmes rencontrés

- Prolongation du downtime jusqu'au 08/02/2011 10:00
- Pb de checksum sur le Pool2Pool
- Pb d'authentification pour certains DNs de LHCb
- Etc.



- Migration du LFC Atlas prévue 20/22 mars
 - Suivi par Emmanouil et David
- Migration du MW vers EMI-1/2
 - “Very urgent”
- Recensement des service grille à entamer
 - Etat, Hardware, Configuration, Monitoring, Contacts...



Prochaine T1/AF



- Date et lieu
 - 22 mars en salle 202
- Proposition de sujets
 - “Whole node” job
 - Sébastien ?
 - Frontier/Squid
 - Emmanouil ?
 - Etat d’avancement du déploiement d’OTRS ?
 - David ?
 - Autres ?



- A venir (à mettre à jour)
 - OGF 34, Oxford, 11-14 March 2012
 - HEPiX Spring Meeting, Prague, 23-27 April 2012
 - ISGC, Taipei, 26 February - 2 March 2012
 - EGI User Forum, Munich, 19-23 March
 - WLCG Workshop, NYC, 19-20th May 2012
 - CHEP2012, NYC, 21-25th May 2012
 - EGI Technical Forum, Prague, September 2012
 - HEPiX Fall Meeting, IHEP, October 2012