



2011/11/30

# News of CCIN2P3-T1/AF

Rencontre LCG-France

*Lyon, November 30<sup>th</sup>-December 1<sup>st</sup> 2011*

Pierre Girard

[Pierre.girard@in2p3.fr](mailto:Pierre.girard@in2p3.fr)

dapnia

cea

saclay





- What has changed since June 2010
- WLCG metrics
- Site overview
- Major concerns
- Conclusions & Perspectives

# What has changed since june 2010

# ▶ Main changes at CC-IN2P3



- Moved from EGEE project to EGI project
- New organisation since F. Hernandez has left
  - CC-IN2P3 technical director
    - P.E. Macchi
  - LCG France technical director
    - F. Chollet-Le Flour
  - CCIN2P3-T1 representative
    - P. Girard
- New machine room
- New LRMS (Grid Engine)
- Manpower turnover

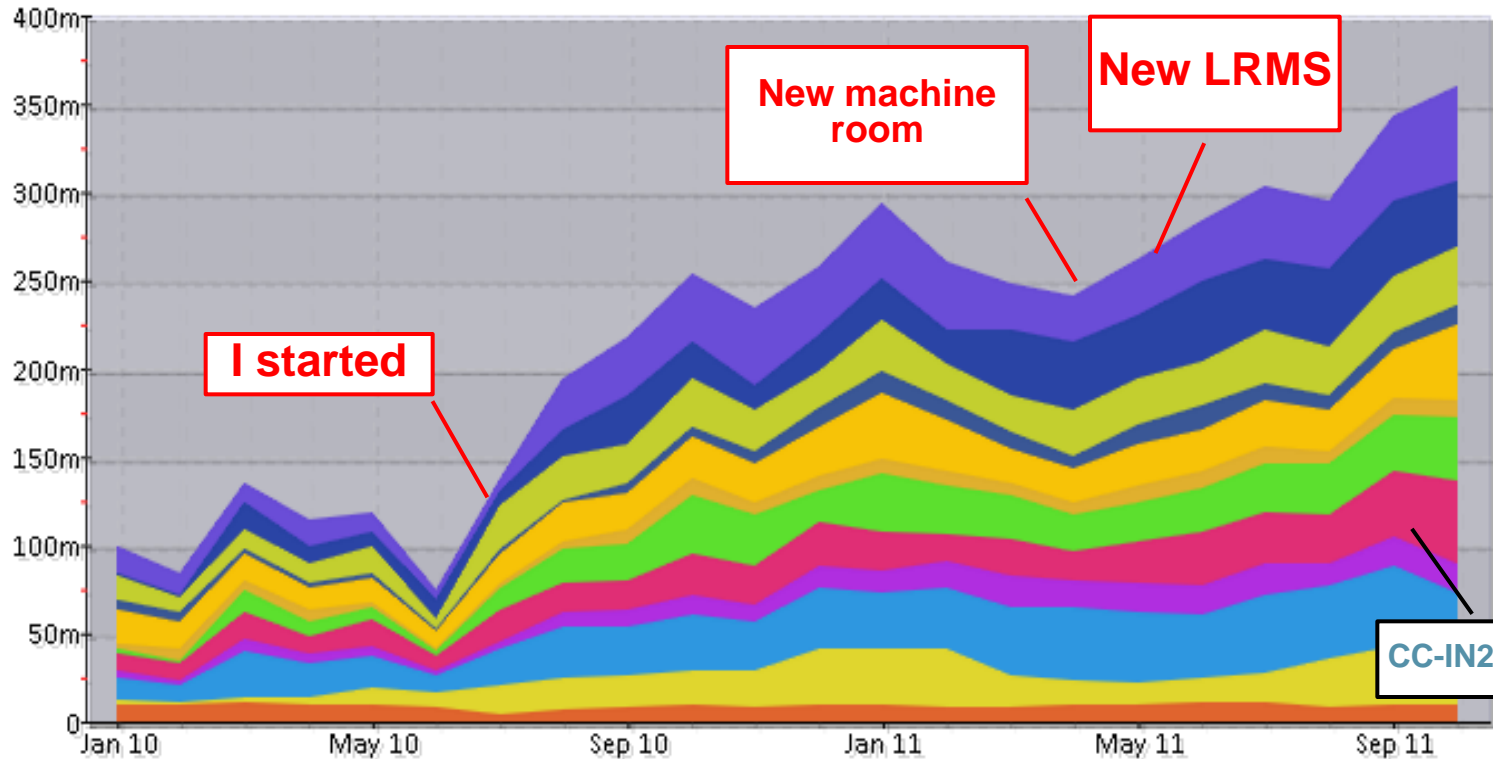


# Main changes at WLCG



Cumulative Normalised CPU time (HEPSPEC06) by TIER1 and DATE

- CA-TRIUMF
- CH-CERN
- DE-KIT
- ES-PIC
- FR-CCIN2P3
- IT-INFN-CNAF
- NDGF
- NL-T1
- TW-ASGC
- UK-T1-RAL
- US-FNAL-CMS
- US-T1-BNL



Source: [http://www3.egee.cesga.es/accounting/tier1\\_view.html](http://www3.egee.cesga.es/accounting/tier1_view.html)



**Important manpower turnover**

© CESGA 'EG1 View': TIER1 / normcpu-HEPSPEC06 / 2010:1-2011:10 / TIER1-DATE / Lhc (x) / AC 11-18 05:00

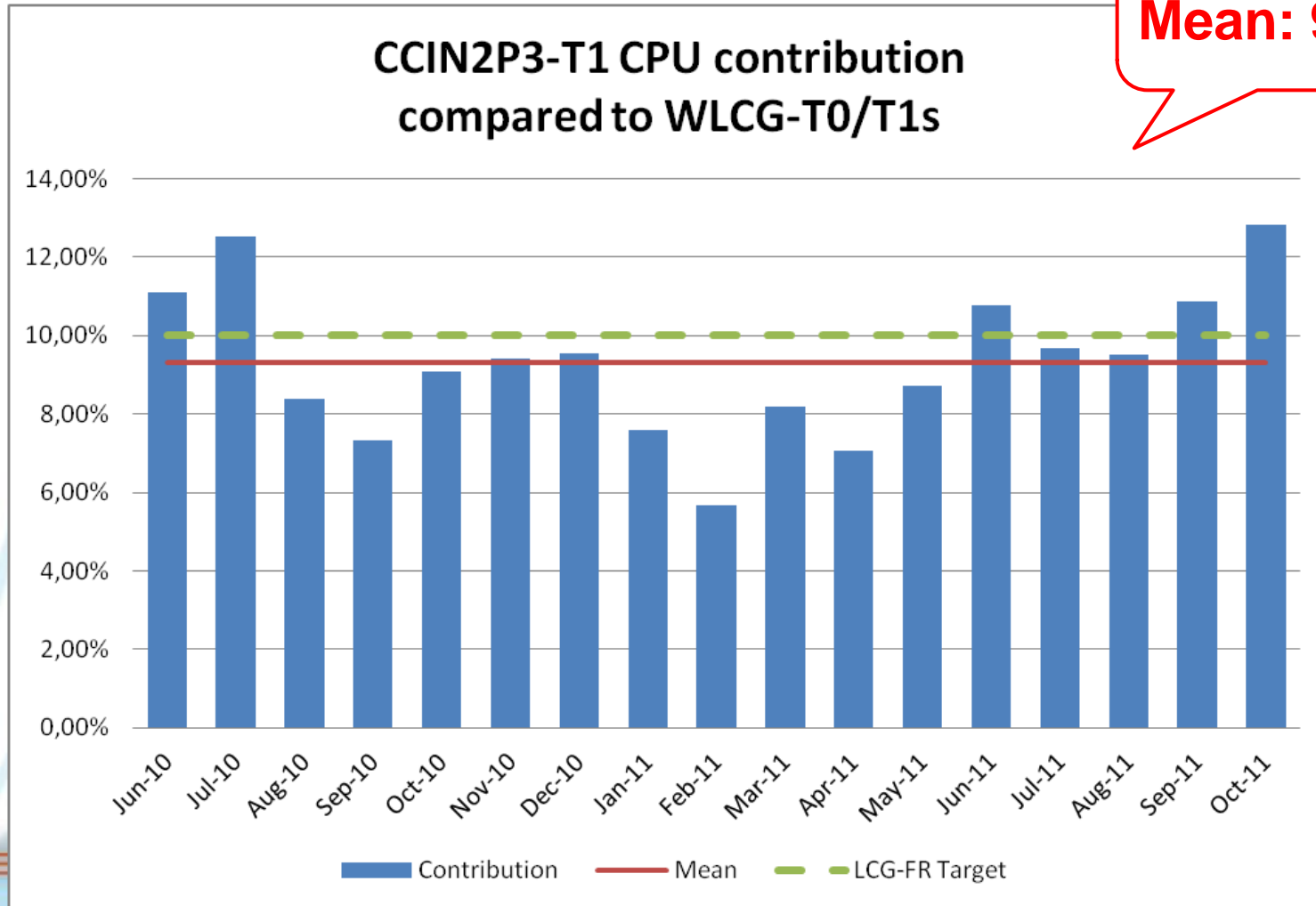
# WLCG Accounting metrics

# CCIN2P3-T1 CPU contribution



**Target: 10%**  
**Mean: 9.3%**

### CCIN2P3-T1 CPU contribution compared to WLCG-T0/T1s

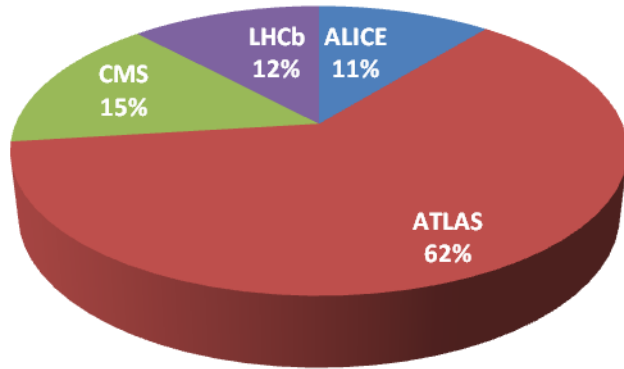


Source: [http://www3.egee.cesga.es/accounting/tier1\\_view.html](http://www3.egee.cesga.es/accounting/tier1_view.html)

# Compared CPU usages at CCIN2P3-T1



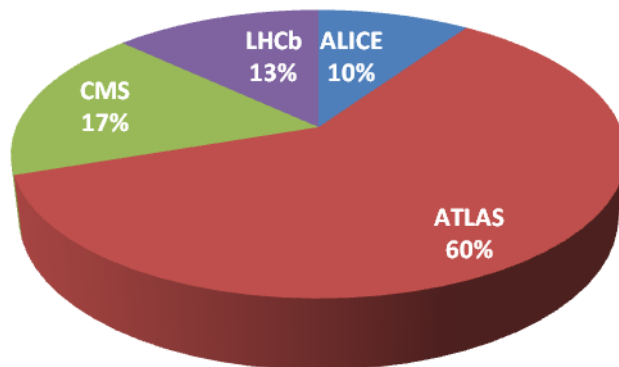
### Compared CPU usages for 2010 (06/2010-03/2011)



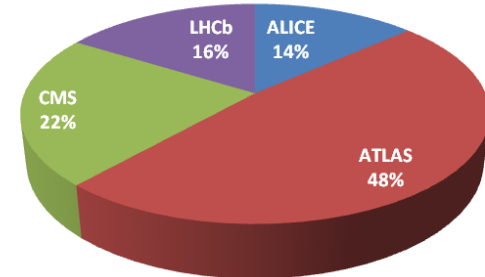
### To be compared with pledged CPU

- Pledges 2010
  - ALICE: 9%
  - ATLAS: 49%
  - CMS: 21%
  - LHCb:22%
- Pledges 2011
  - ALICE: 11%
  - ATLAS:46%
  - CMS: 15%
  - LHCb:28%

### Compared CPU usages for 2011 (04/11-10/11)



### Compared CPU usages for October 2011



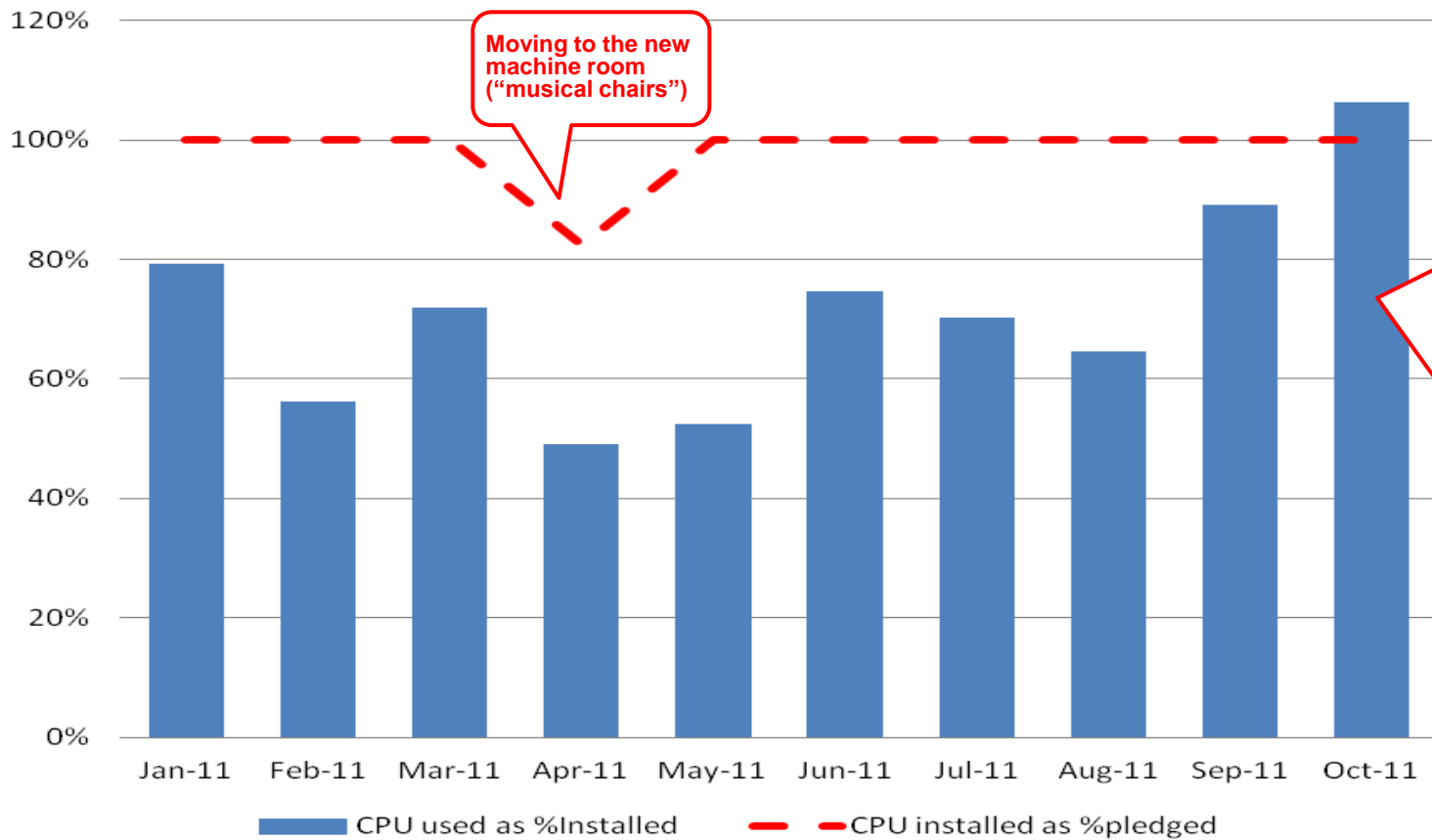
Source: <http://lcg.web.cern.ch/lcg/accounts.htm>



# CPU usage vs CPU pledged



## Used, Installed and Pledged CPU



Moving to the new machine room ("musical chairs")

Including the CPU efficiency factor of 85%

CPU efficiency has significantly increased during the last 2 months

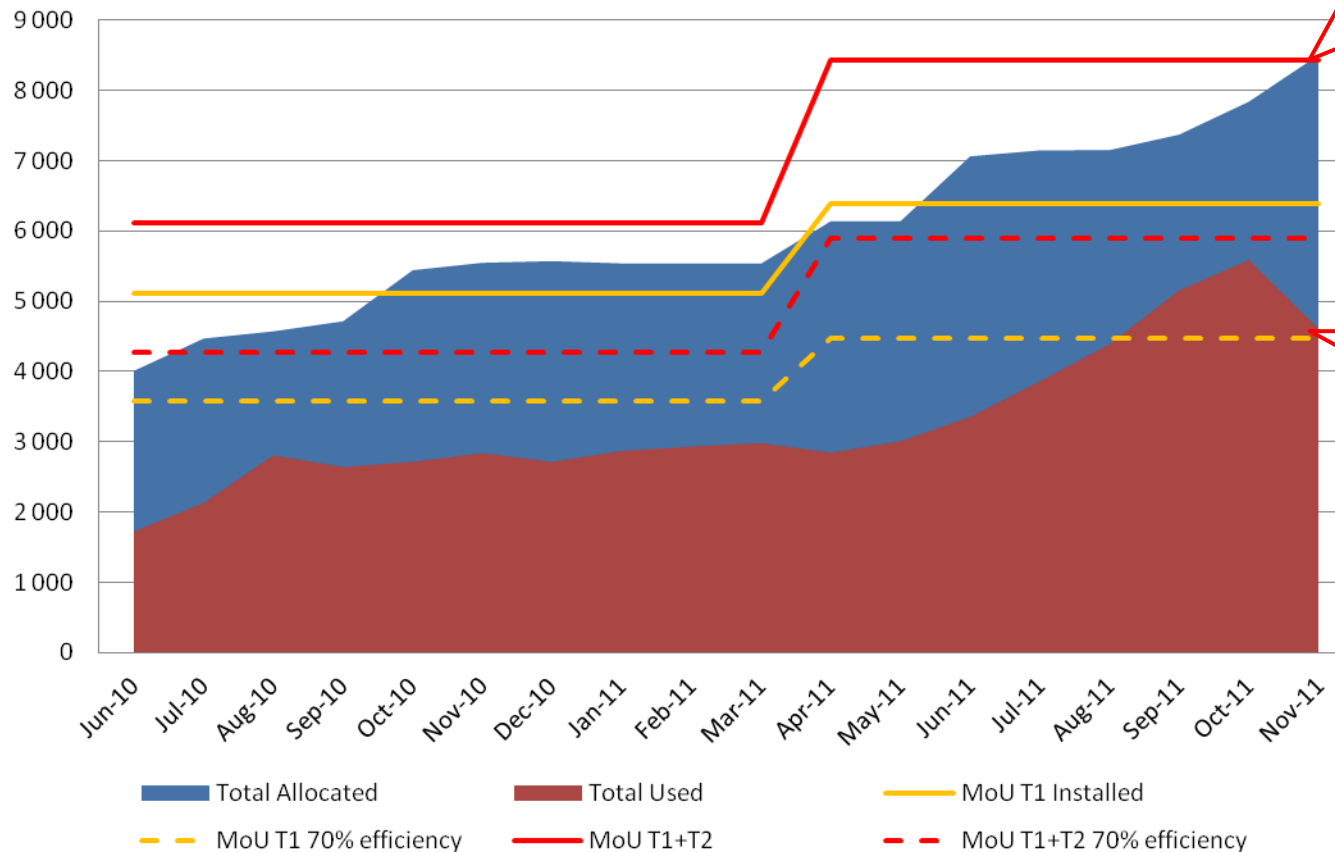
See major concerns

Source: <http://lcg.web.cern.ch/lcg/accounts.htm>

# Disk Usage vs Disk pledge



## Installed/Pledged Disk (TB)



Most of the experiments share storage between T1 and T2.

ALICE required to migrate its storage from SOLARIS to Linux

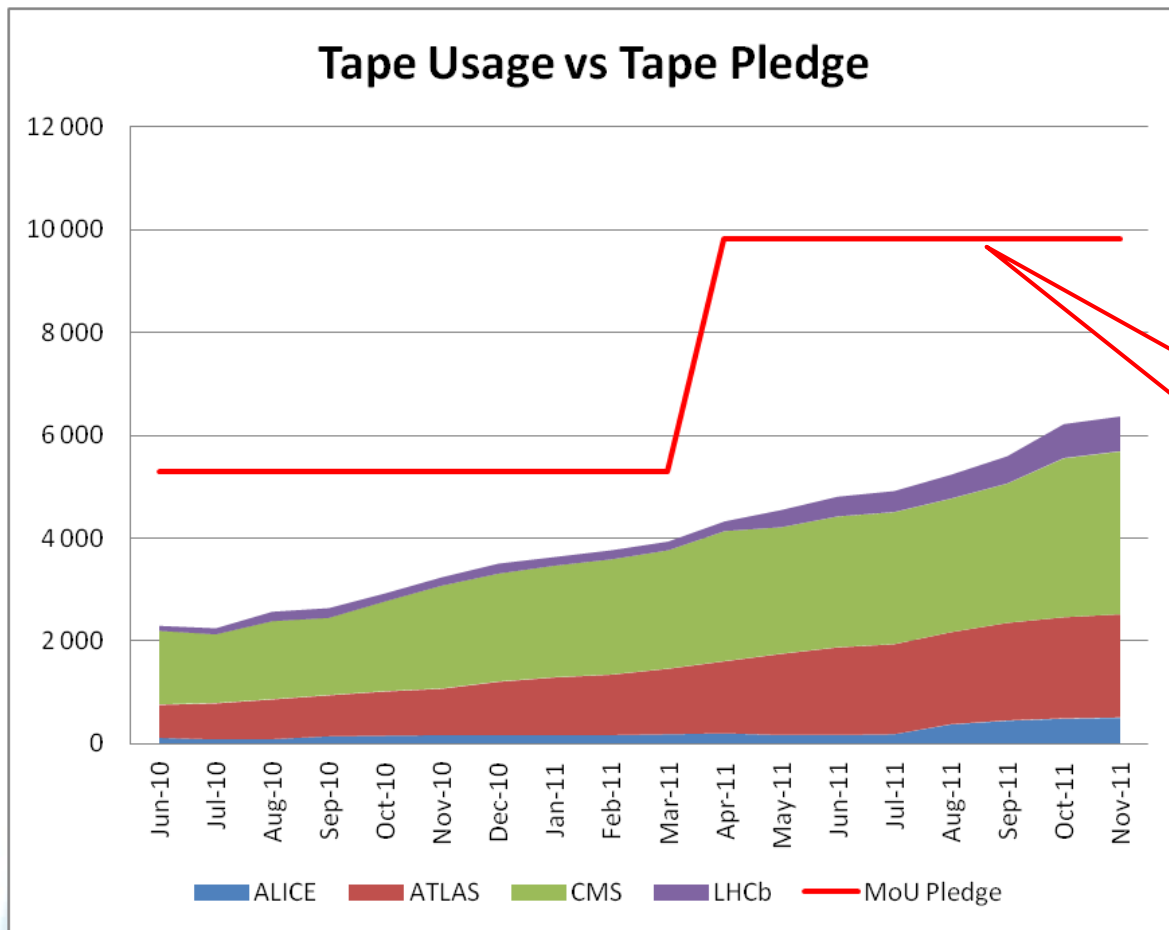
CIN2P3 warned ATLAS about the number of stored files. Since October, ATLAS has removed 1 PB

Source: <http://lcg.web.cern.ch/lcg/accounts.htm>

# Tape Usage vs Tape Pledge



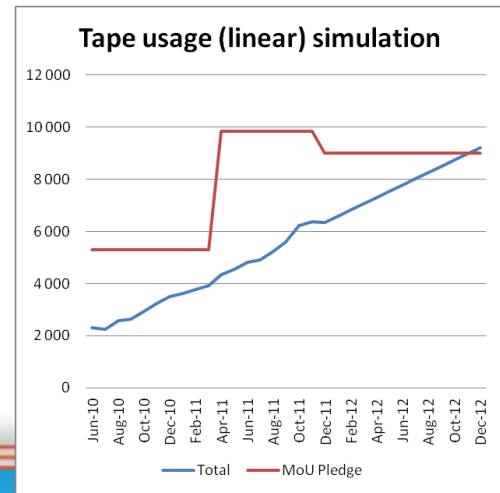
Source: <http://lcg.web.cern.ch/lcg/accounts.htm>



To be compared with pledges for 2011

- ALICE: 47%
- ATLAS: 48%
- CMS: **88%**
- LHCb: 67%

The Tape capacity will be reduced from 850 TB in 2012 on demand of ALICE and ATLAS





# WLCG Service metrics

# Availability/Reliability of CCIN2P3-T1

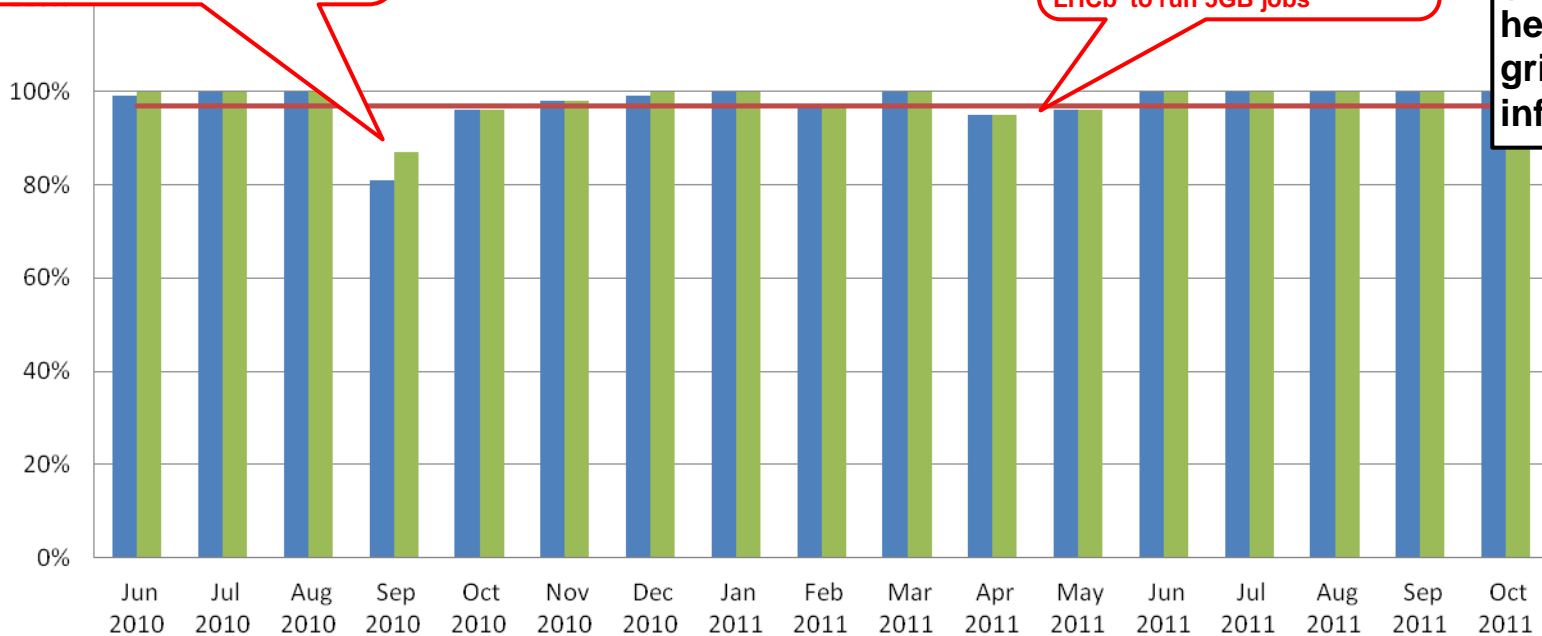


Global security alert which led to an unscheduled downtime of several days

Site Availability/Reliability

Mostly due to a monitoring problem after tricky configuration changes to help LHCb to run 5GB jobs

Using OPS VO which gives us a good idea of the health of both grid services and infrastructure



	Jun 2010	Jul 2010	Aug 2010	Sep 2010	Oct 2010	Nov 2010	Dec 2010	Jan 2011	Feb 2011	Mar 2011	Apr 2011	May 2011	Jun 2011	Jul 2011	Aug 2011	Sep 2011	Oct 2011
OPS-Av	99%	100%	100%	81%	96%	98%	99%	100%	97%	100%	95%	96%	100%	100%	100%	100%	100%
OPS-Rel	100%	100%	100%	87%	96%	98%	100%	100%	97%	100%	95%	96%	100%	100%	100%	100%	100%
Target	97%	97%	97%	97%	97%	97%	97%	97%	97%	97%	97%	97%	97%	97%	97%	97%	97%

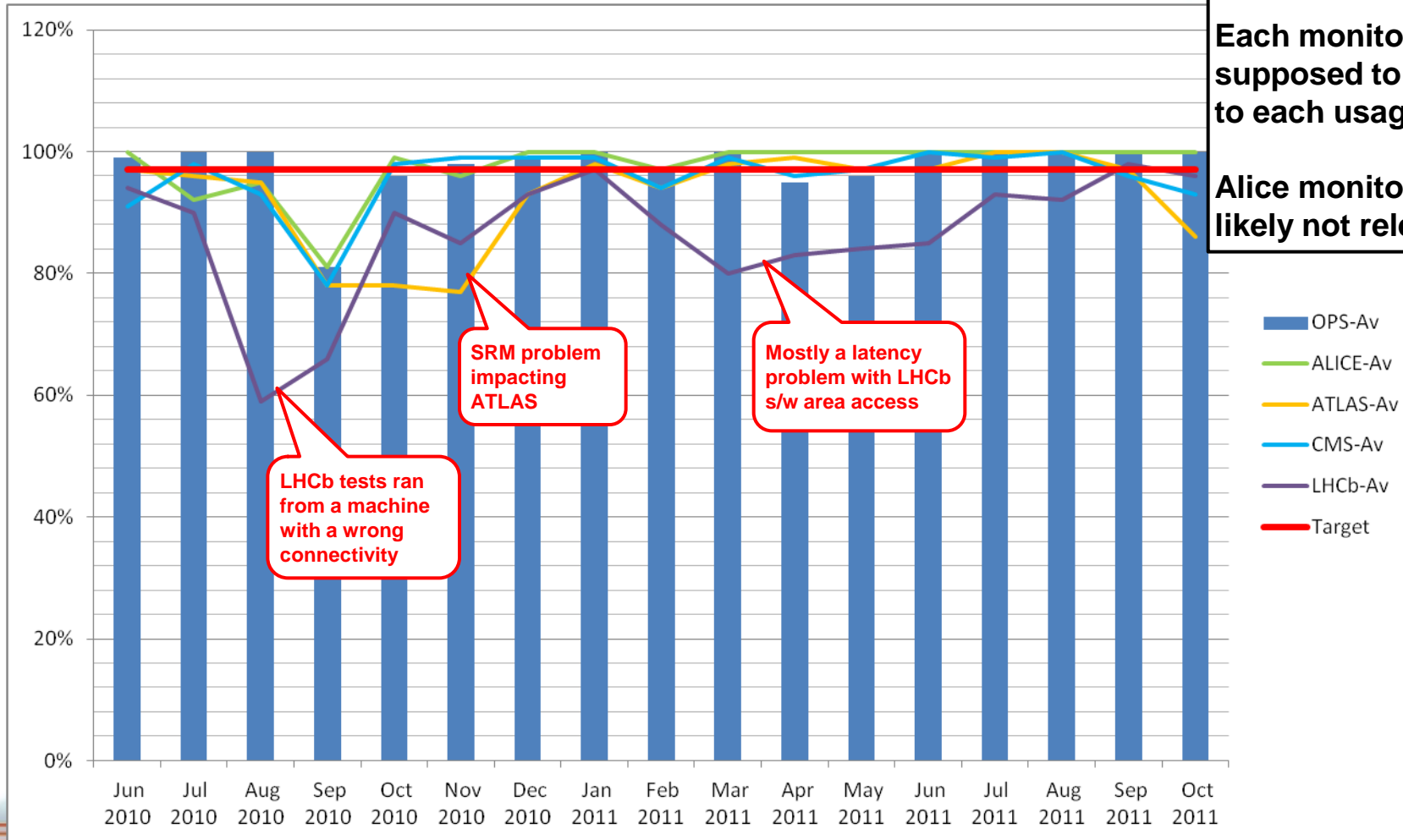


# VOs Availability of CCIN2P3-T1

Same infrastructure,  
different usages.

Each monitoring is  
supposed to be adapted  
to each usage

Alice monitoring is  
likely not relevant.



# WLCG Service Incident Reports



Year	Quarter	Area	Type	Duration
2010	Q4	Storage	Degradation	2 months
2011	Q1	Shared s/w area	Degradation	6 months
		Network	Outage	3 hours
		Power Cut	Outage	13 hours
		Network	Outage	40 mn
	Q2	Storage	Degradation	3.5 hours
		Power cut	Outage	5 hours
		Q3	Database / Hardware	Data loss
Power cut	Outage		20 hours	
		Power cut / Cooling system	Outage	7.5 hours

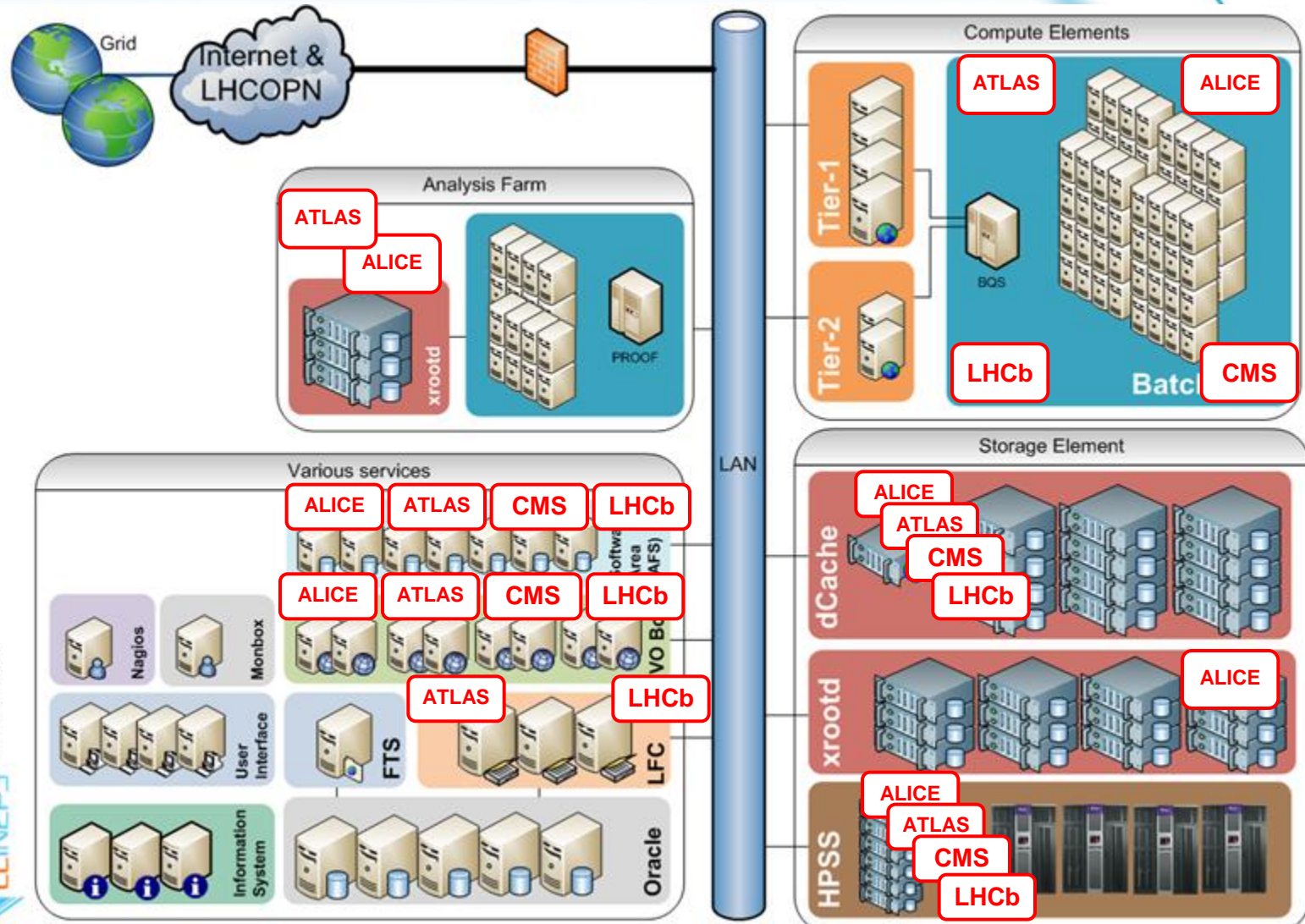
- **Good responsiveness for infrastructure incidents**
- **Degradation incidents**
  - **Related to grid activities**
  - **Hard to understand and then to solve**
  - **Require a thin collaboration with external people (Middleware/VO)**

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



# Site overview

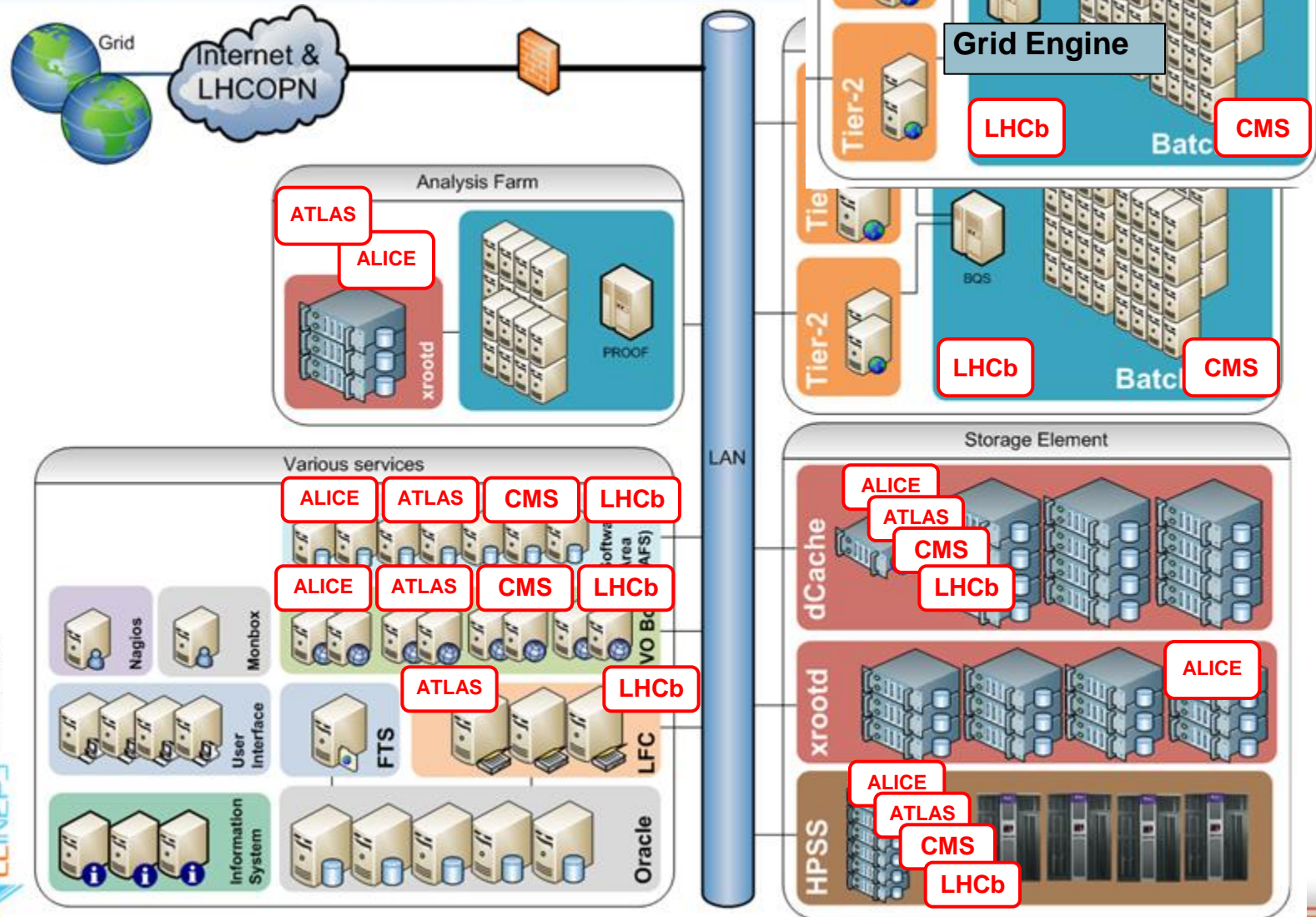
# As it was at the last COS/ESC



[F. Hernandez]



# As it is today (1)

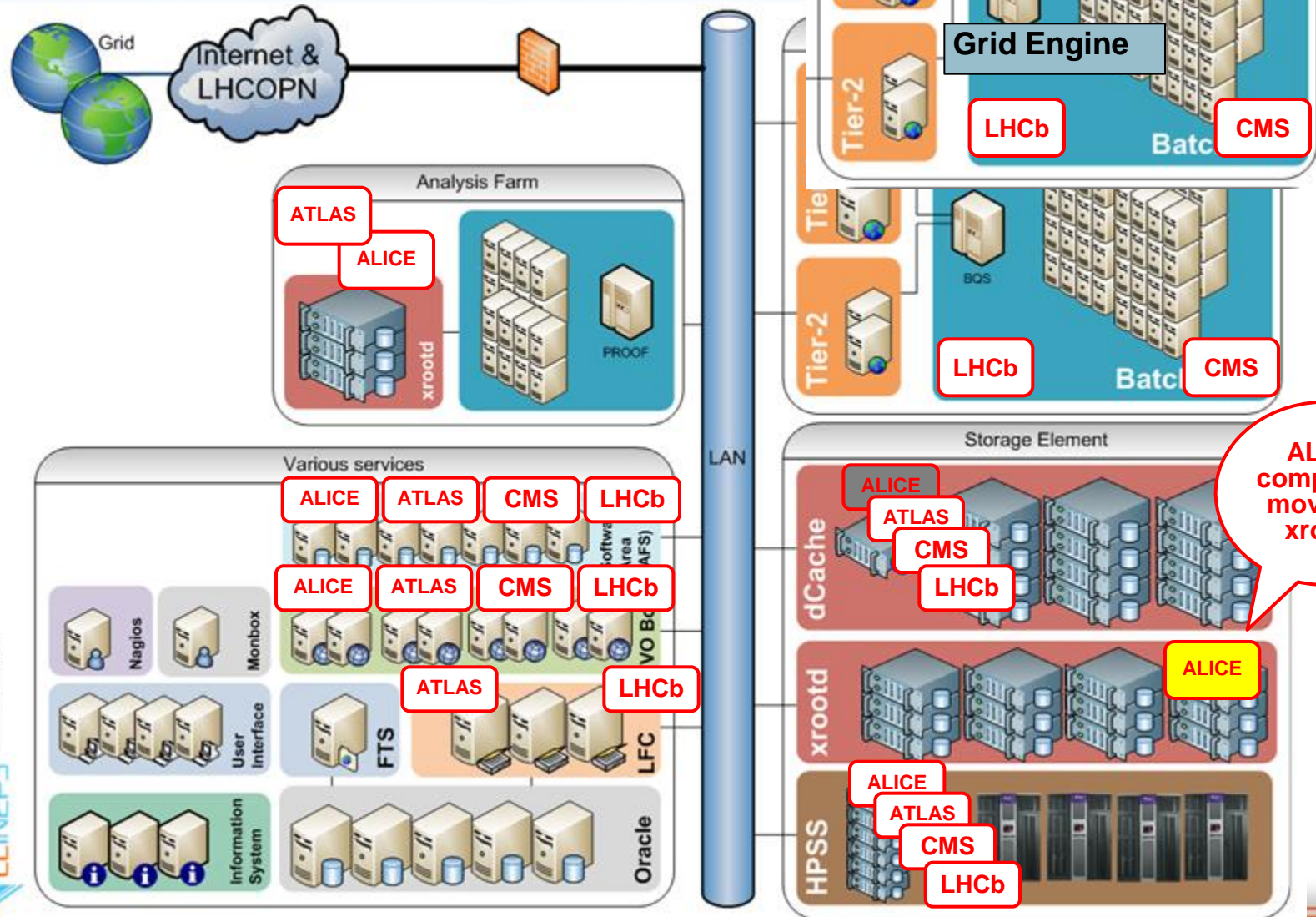


Author: Fabio Hernandez  
Last Updated: 2009-05-29



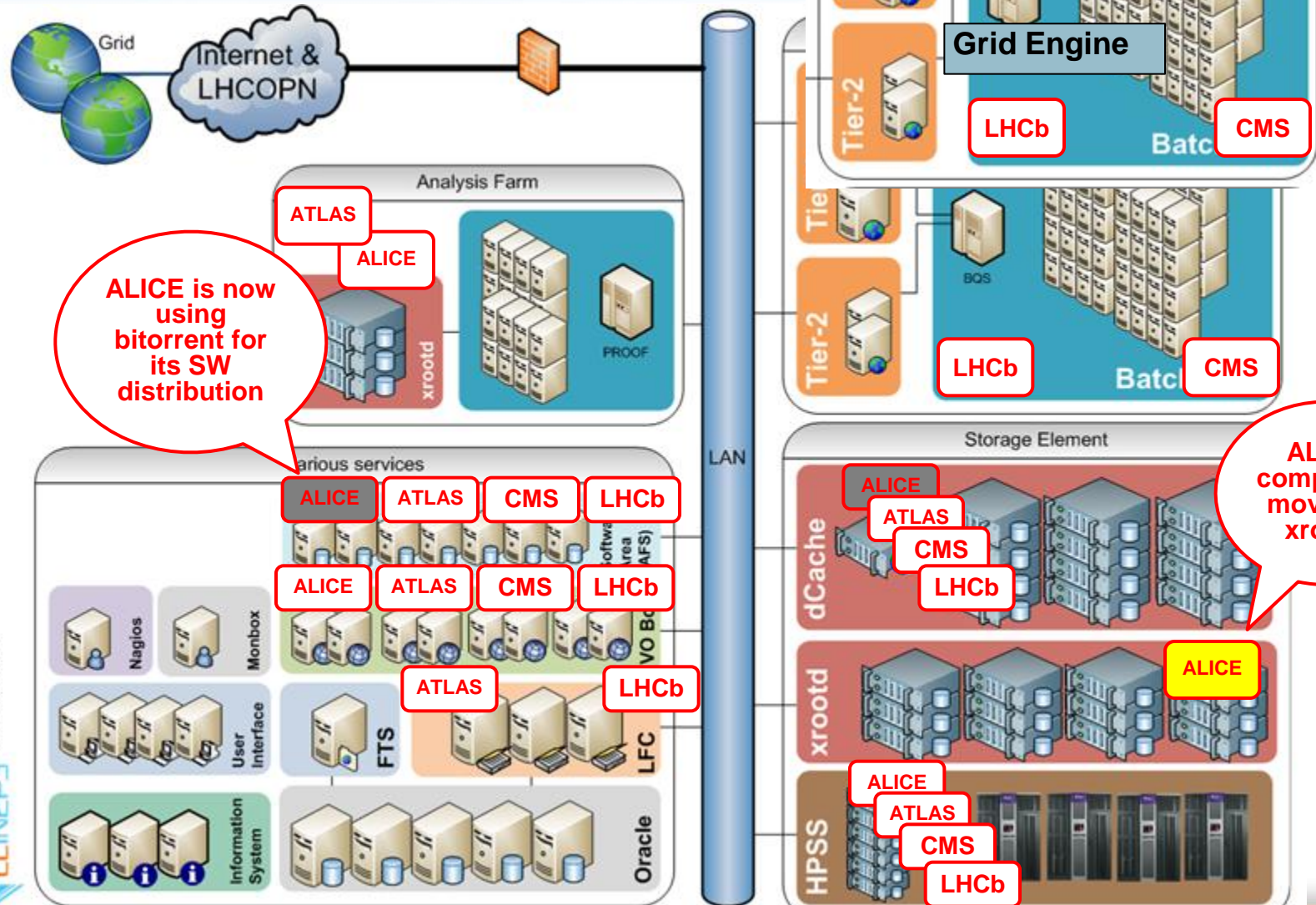


# As it is today (2)



Author: Fabio Hernandez  
Last Updated: 2009-05-29  
CCIN2P3

# As it is today (3)



Author: Fabio Hernandez  
Last Updated: 2009-05-29  
CCIN2P3

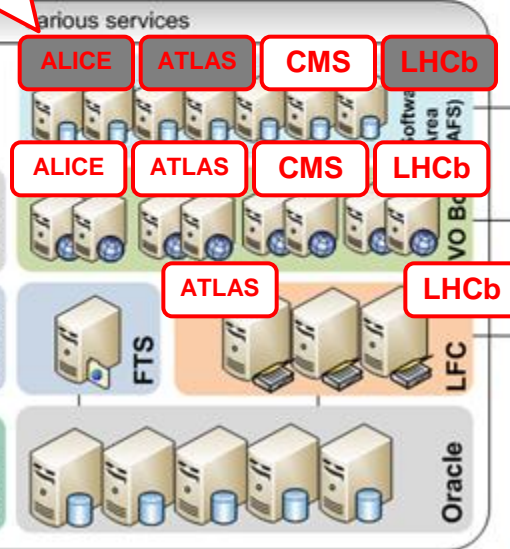
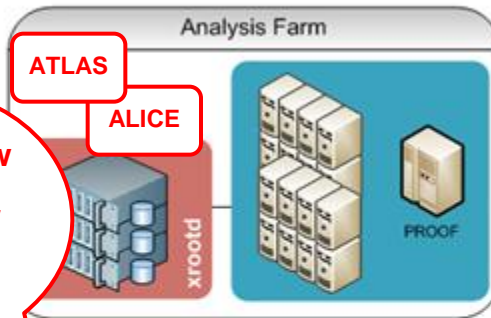
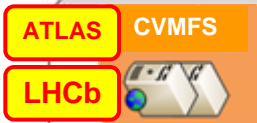


# As it is today (4)

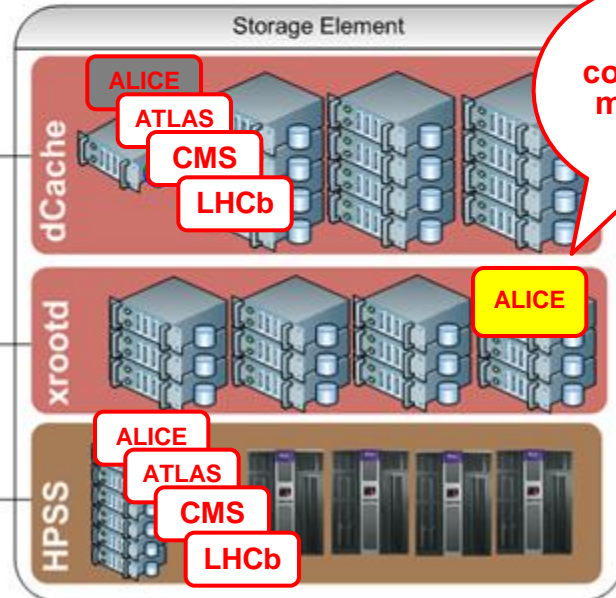
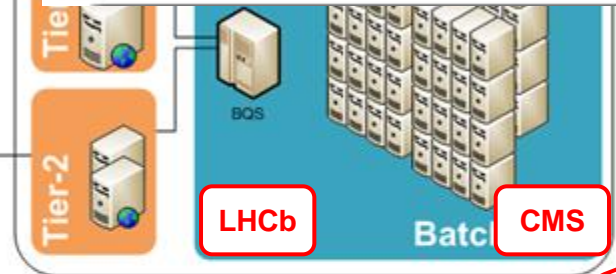
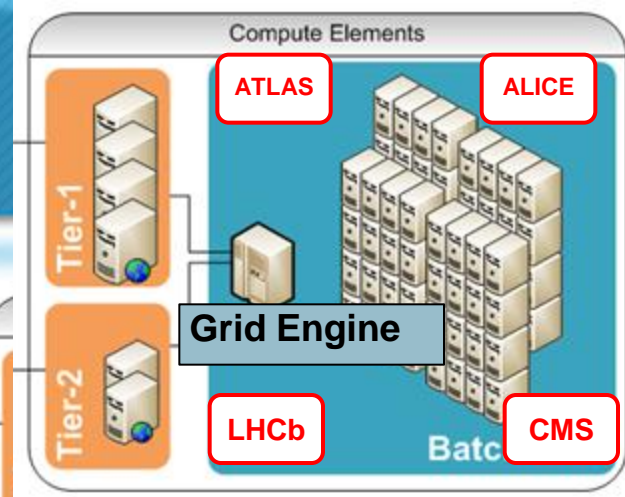


ATLAS/LHCb moved their SW Area to CVMFS

ALICE is now using bittorrent for its SW distribution



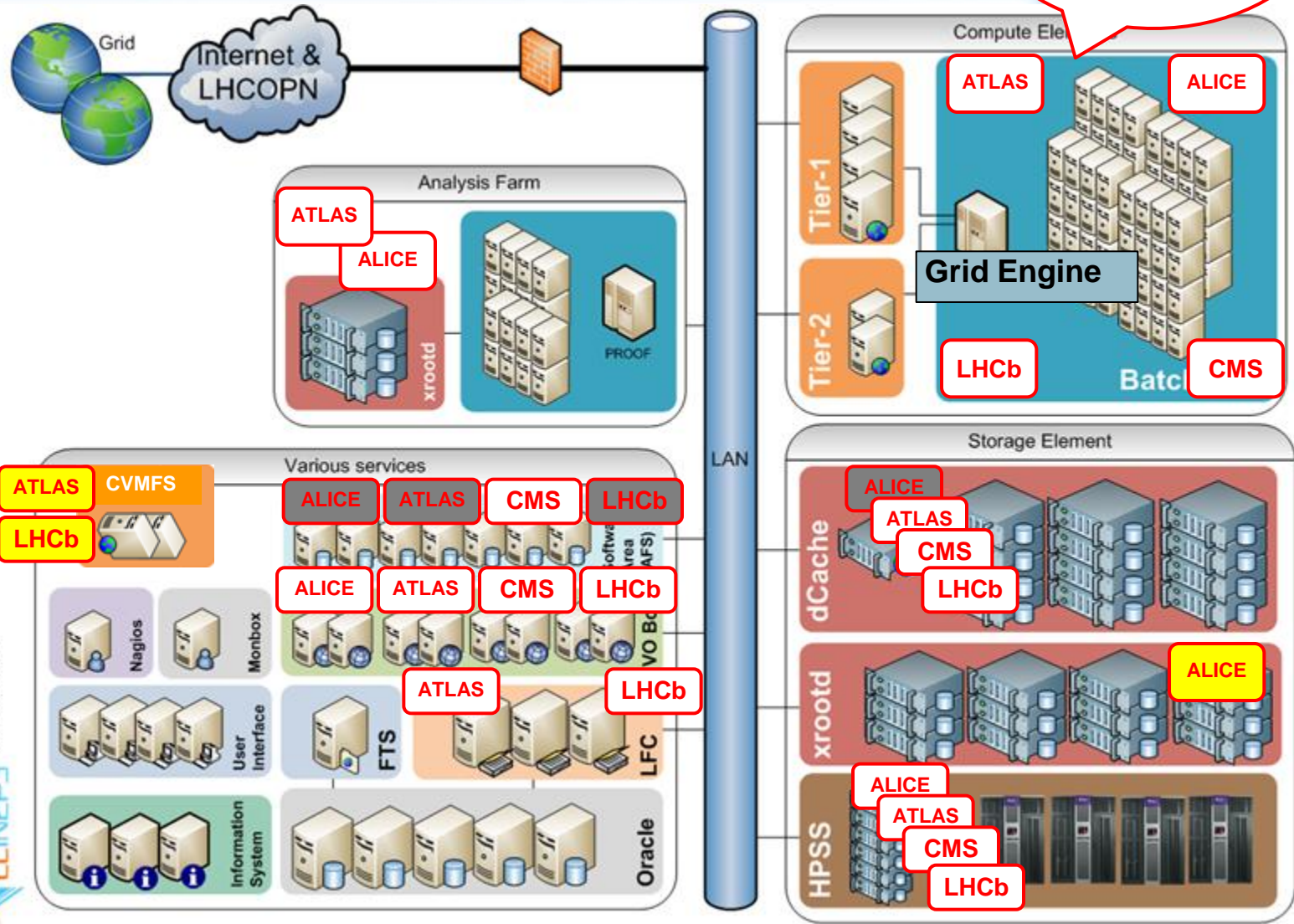
LAN



ALICE completely moved to xrootd

# As it is today (5)

BQS will be stopped on Dec. 6th. But no more BQS CE for LHC VOs yet



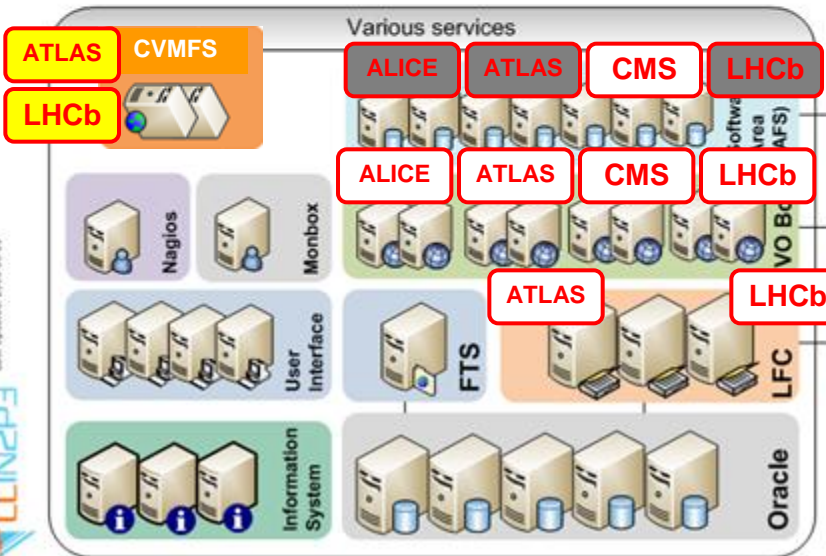
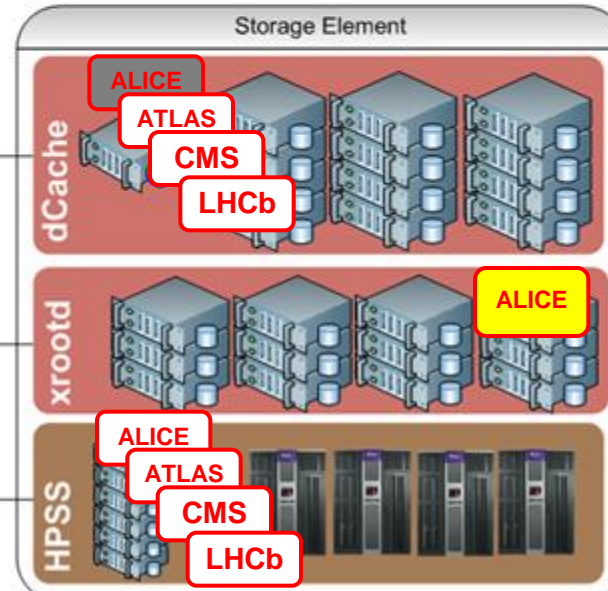
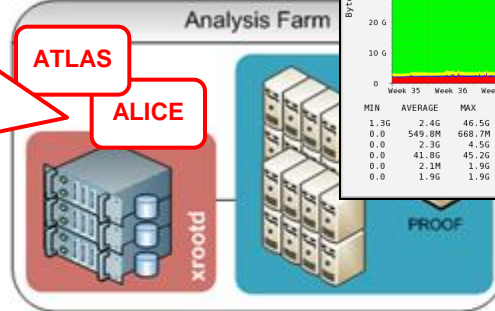
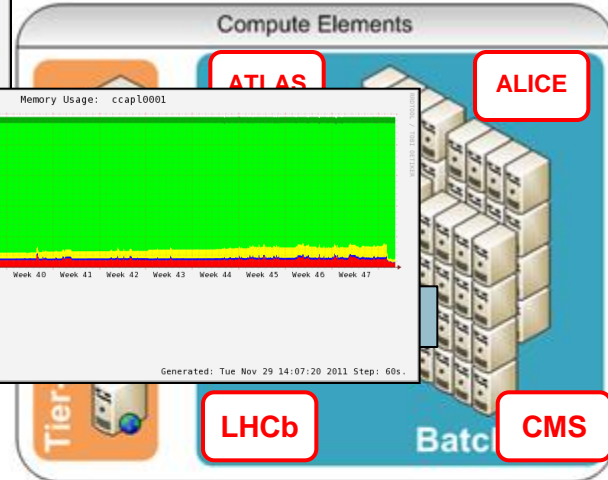
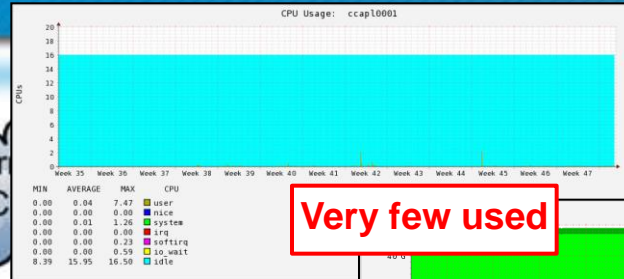
Author: Fabio Hernandez  
Last Updated: 2009-05-29  
CCIN2P3



# As it will be tomorrow (1)

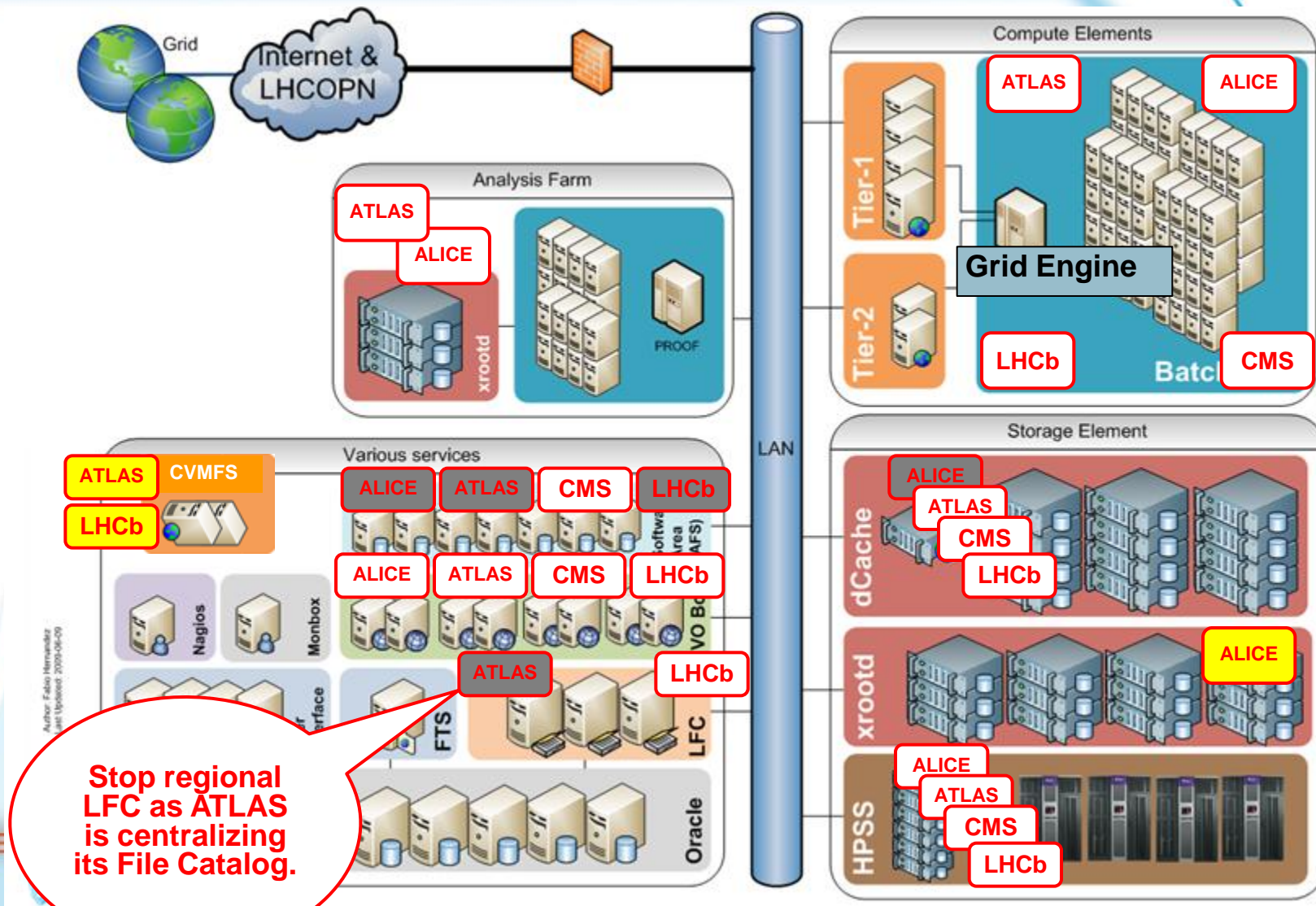
ALICE would like to use it more but extra (ALICE) software is required for.

Ongoing discussion with ATLAS people to understand if they still plan to use the LAF.





# As it will be tomorrow (2)





# Major concerns

# Top 5 issues

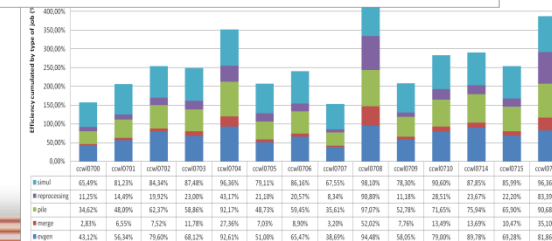
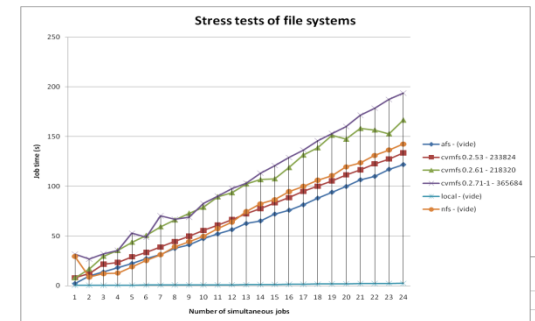


- Low transfers with dCache server during ATLAS reprocessing campaign (**Solved**)
  - Regular meetings with dCache.org developers
  - Regular meetings/reports with ATLAS operations team
  - Problem identified by our administrators and solved with the help of dCache.org people
  - Thin collaboration with French ATLAS people
    - To reproduce the problem
    - To make sure we get it under control

Plage Horaire (UTC)	MCDISK -> DATADISK				* -> DATADISK			
	efficacité	throughput	nombre de transferts en succès	en echec	efficacité	throughput	nombre de transferts en succès	en echec
08:00-9:00	100 %	50 MB/s	132	0	99 %	1025 MB/s	4729	24
09:00-10:00	100 %	88 MB/s	471	7	98 %	485 MB/s	3684	73
10:00-11:00	100 %	146 MB/s	822	4	93 %	316 MB/s	4100	315
11:00-12:00	100%	50 MB/s	928	1	100 %	282 MB/s	3938	13
12:00-13:00	98%	31 MB/s	198	4	100 %	264 MB/s	1685	8



- Poor ATLAS/LHCb jobs efficiency due to AFS latency (**Solved**)
  - Originally an experiment S/W problem (“insane” job setup script)
  - Becoming particularly critical with the new hardware (more cores = more jobs)
  - Thin collaboration with French ATLAS people
  - Solved by introducing a new solution (CVMFS) at sites level on demand of VOs
    - After a lot of investigations, stress tests, etc.
    - Final results presented at the ATLAS SW week in July 2011



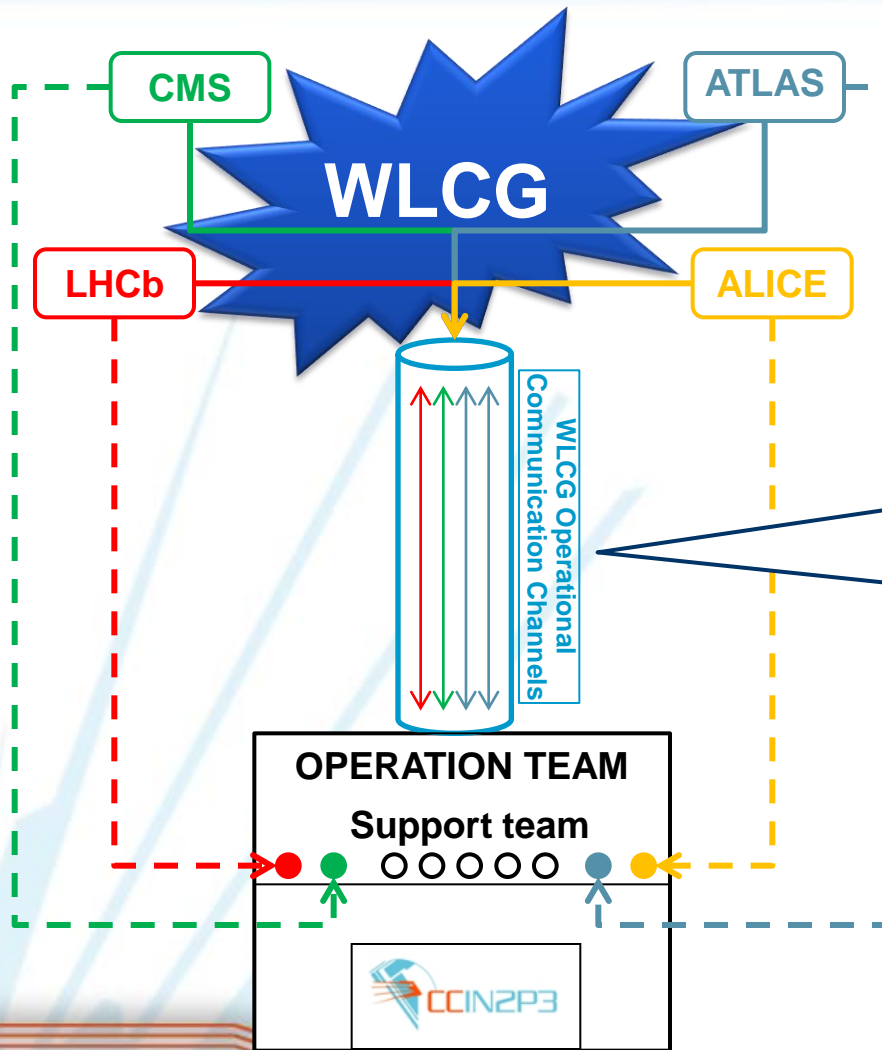


# Top 5 issues



- Oracle DB crash with loss of data (Ongoing)
  - Storage hardware problem
  - Actually **solved** by moving to other storage system
  - But still in discussion with Oracle support
  - Improvement of incident procedures
    - To speed up the restoring of DB backup
    - Has proved to be very useful in October
  
- CREAM GE (Grid Computing Element) instabilities (**Solved ?**)
  - Led to temporary jobs shortage for ATLAS during a reprocessing campaign
  - Thin collaboration with CREAM GE developers
    - 9 tested patches from May to September
  
- Low transfers between CCIN2P3 and foreign T2s since May (Ongoing)
  - Within WLCG Project, Tier1 is responsible for the incident follow-up
  - Collaboration with RENATER (French NREN)
  - First progress obtained last week
    - By disabling the QoS tagging of our network packets

# Operational issues



## Context

We figured out at the end of September that LHCb underused CCIN2P3 resources during months

## Main reason

- LHCb based its operations on the sites contacts
- LHCb-dedicated person left CCIN2P3 early this year, months were needed to find a new person
- LHCb problems did not come up through usual WLCG channels

**This is not optional**

- CCIN2P3 operation people apply the WLCG procedures.
- GGUS system is interfaced with our ticketing system
- CCIN2P3 people attend the WLCG meetings

**Dedicated support is for facilitating the life of both the VO and the site.**

**CCIN2P3 cannot sustain the effort of dealing with the operations of 4 VOs**

**See Ghita's talk about the manpower turnover problem**

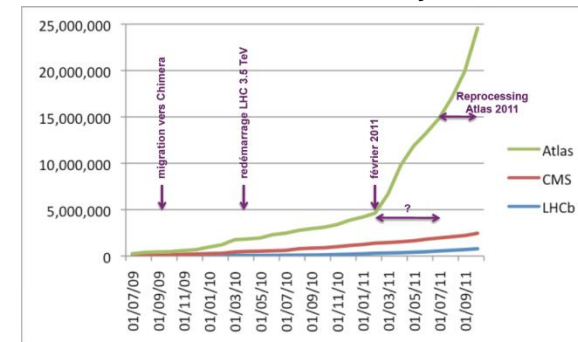




# Conclusions & Perspectives

- Short/medium terms
  - Address near future storage scalability issues
  - Clarify the status of the national Analysis Facility
  - Improve our service monitoring
  
- Longer terms
  - Follow up and participate as much as possible at the WLCG Technical Evolution Groups
    - Those groups are addressing most of our current concerns
  - Investigate new solutions and technologies
    - Whole node jobs in collaboration with both ATLAS and CMS
    - Virtualization and Cloud computing/storage

dCache entries history



# Conclusions



- Lot of work done since june 2010
- The site was able to face the ramp up of the LHC
  - Despite many big changes at CC-IN2P3
  - Despite many changes from the VOs
- The site contributed to LHC Computing in accordance with LCG France plans
- We are far away from the initial grid model based on shared services
  - Computing model of each VO is evolving
  - CCIN2P3 is consequently used in separate ways
- We must take care at the increasing entropy
  - Manpower consuming
  - Site is sized for sharing/mutualization



- When I made this presentation, I realized what we had to cope with and all the work which has been achieved for the last year, so many thanks to
  - My CCIN2P3 colleagues
  - Our LCG-France colleagues
  - Our WLCG collaborators
  - Our old good friend BQS (RIP)

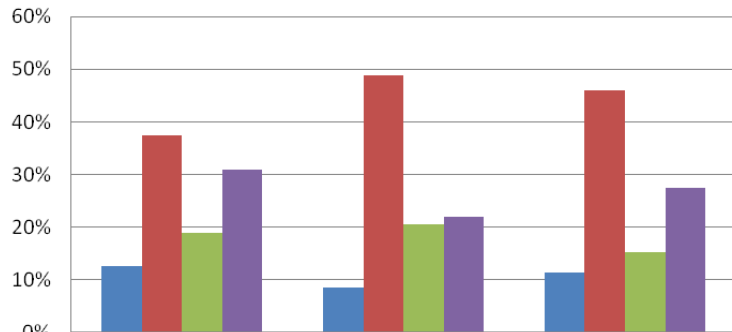


# Backup slides

# CCIN2P3-T1 Pledged Resources

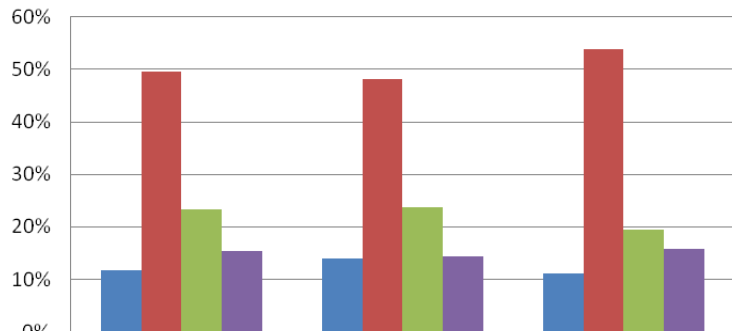


## CCIN2P3-T1 CPU Share



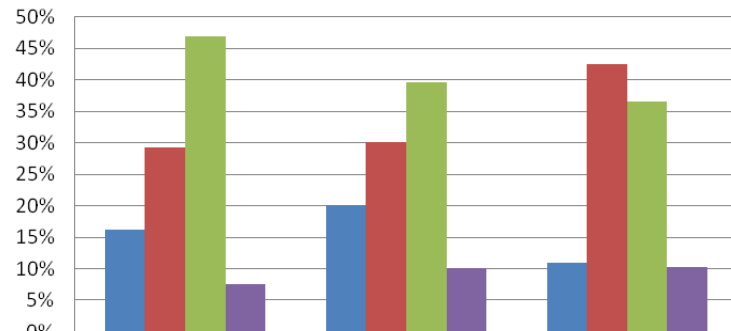
	2009	2010	2011
ALICE	13%	9%	11%
ATLAS	37%	49%	46%
CMS	19%	21%	15%
LHCb	31%	22%	28%

## CCIN2P3-T1 Disk Share



	2009	2010	2011
ALICE	12%	14%	11%
ATLAS	50%	48%	54%
CMS	23%	24%	19%
LHCb	15%	14%	16%

## CCIN2P3-T1 MSS Share



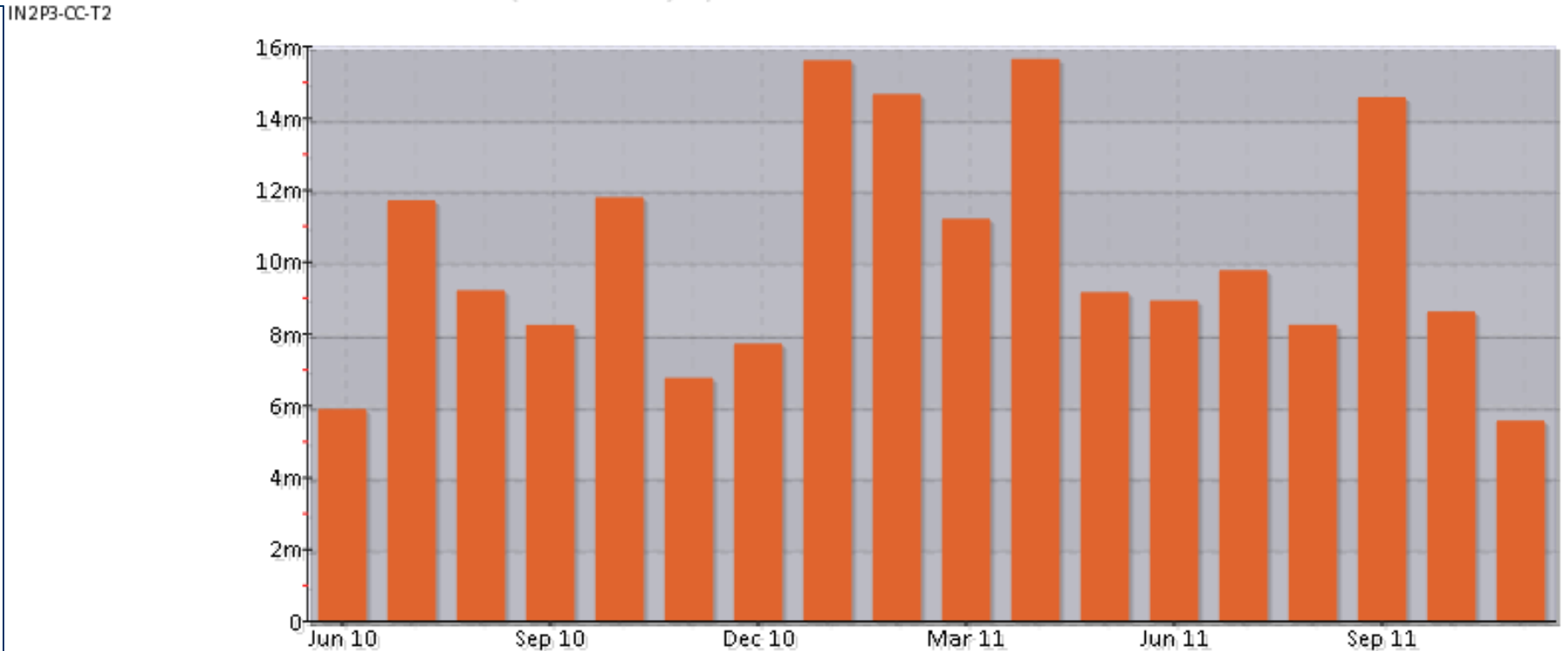
	2009	2010	2011
ALICE	16%	20%	11%
ATLAS	29%	30%	42%
CMS	47%	40%	37%
LHCb	8%	10%	10%



# CCIN2P3 T2 CPU usage



FR-IN2P3-CC-T2 Normalised CPU time (HEPSPEC06) by SITE and DATE



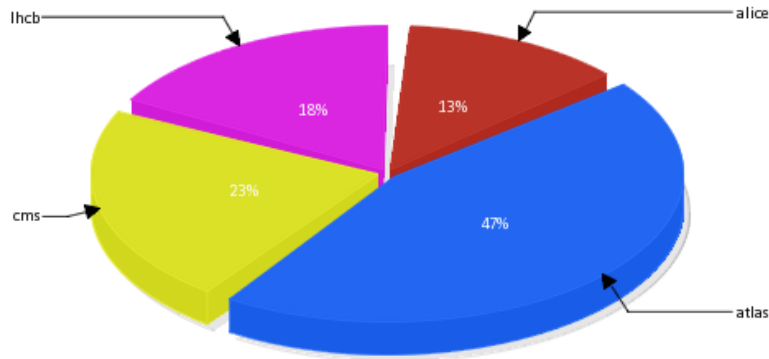
Source: <http://lcg.web.cern.ch/lcg/accounts.htm>

© CESGA 'EGI View': FR-IN2P3-CC-T2 / normcpu-HEPSPEC06 / 2010:6-2011:11 / SITE-DATE / Lhc (x) / ACCBAR-LIN / i 2011-11-23 00:55

# Compared CPU usages at CCIN2P3-T2

Source: <http://lcg.web.cern.ch/lcg/accounts.htm>

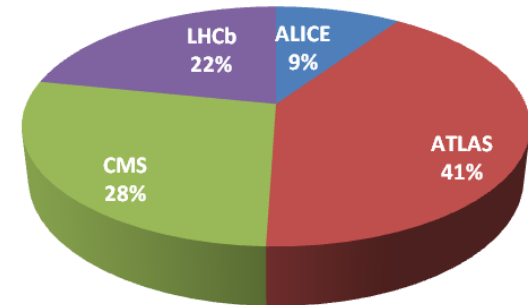
FR-IN2P3-CC-T2 Normalised CPU time (HEPSPEC06) per VO



© CERN. EGI View: FR-IN2P3-CC-T2 / normcpu-HEPSPEC06 / 20106-2011:11 / SITE-VO / lhcb (x) / ACCBAR-LIN / i

2011-11-23 00:55

## IN2P3-CC-T2 CPU Share (Pledges 2011)

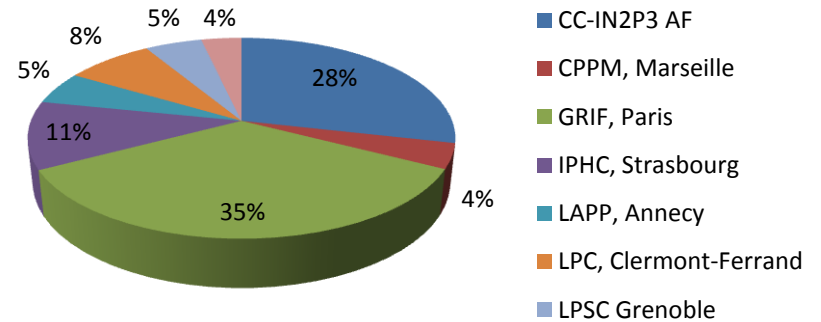


Source: <http://wlcg-rebus.cern.ch/apps/topology/>

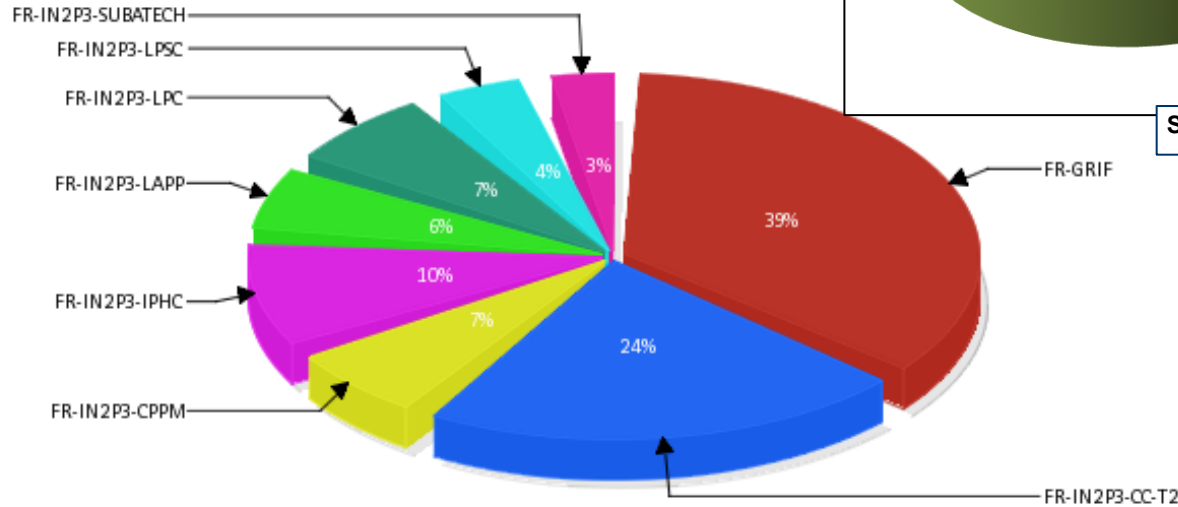
# CCIN2P3 T2 compared to French T2s



## CPU share of French T2s (Pledges 2011)



France (Tier2) Normalised CPU time (HEPSPEC06) per TIER2



Source: <http://wlcg-rebus.cern.ch/apps/topology/>

Source: <http://lcg.web.cern.ch/lcg/accounts.htm>