

ALICE T2s in France

Outline

- Alice T2(T3) sites in France
- Summary data
- Site reports
- Questions

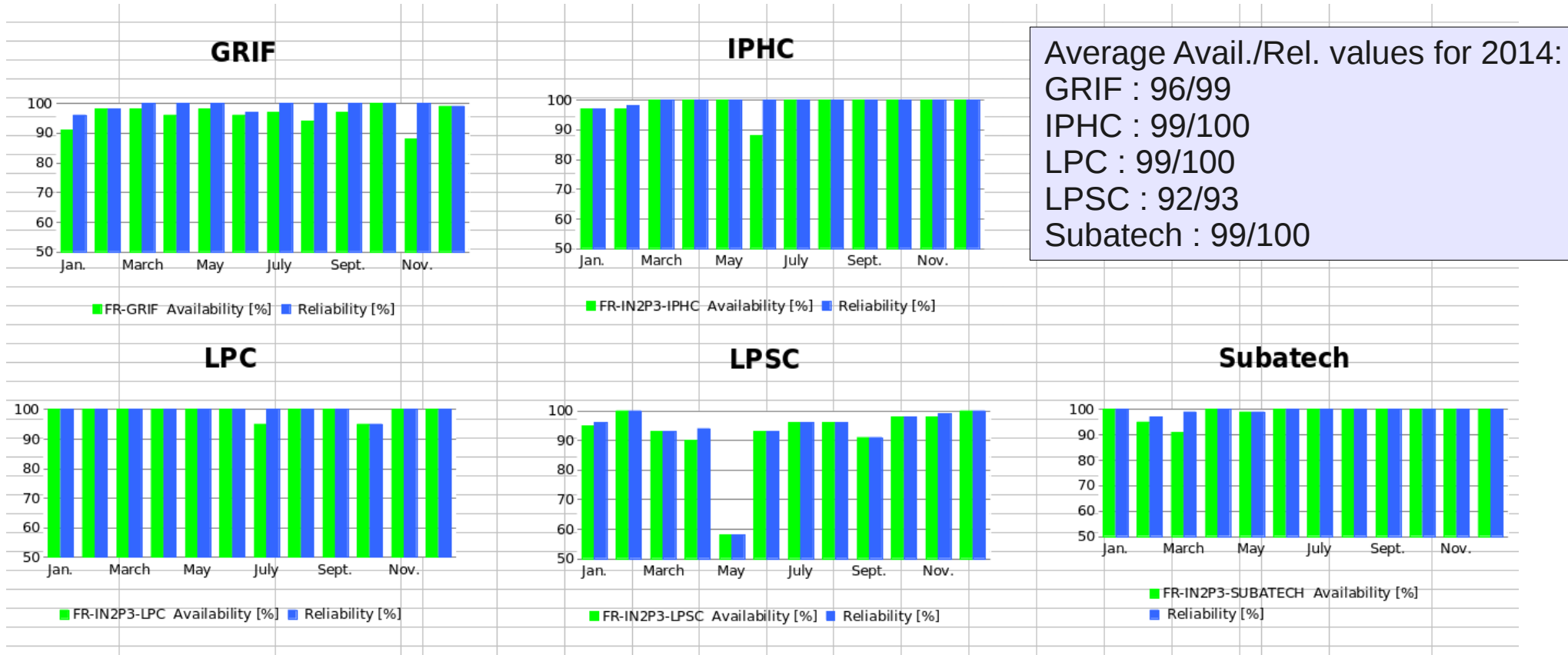
Alice T2/(T3) sites in France

- GRIF : Federated T2 Paris region
 - GRIF IPNO Orsay
 - GRIF IRFU Saclay
- IPHC Strasbourg
- LPC Clermont-Ferrand
- LPSC Grenoble
- IPNL Lyon (T3)
- SUBATECH Nantes



Note : No more CCIN2P3 T2 since 2014

Availability

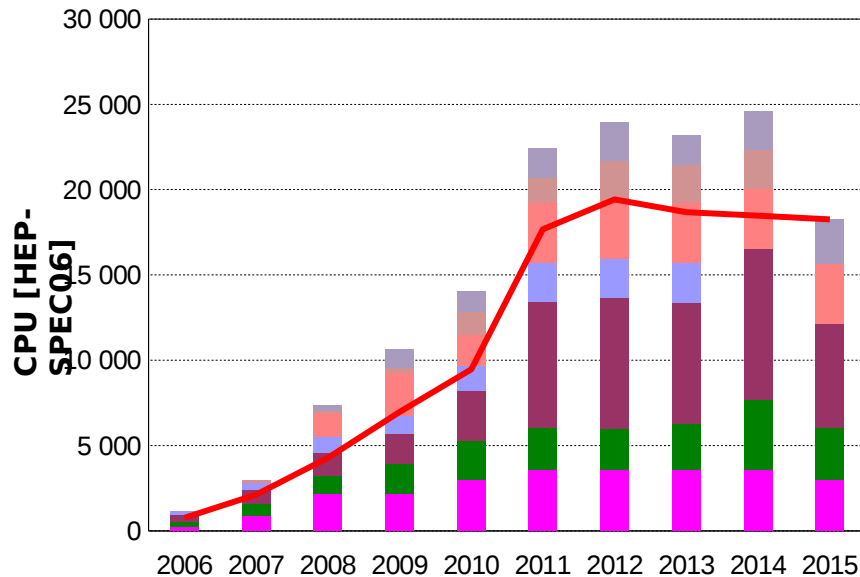


Notes :

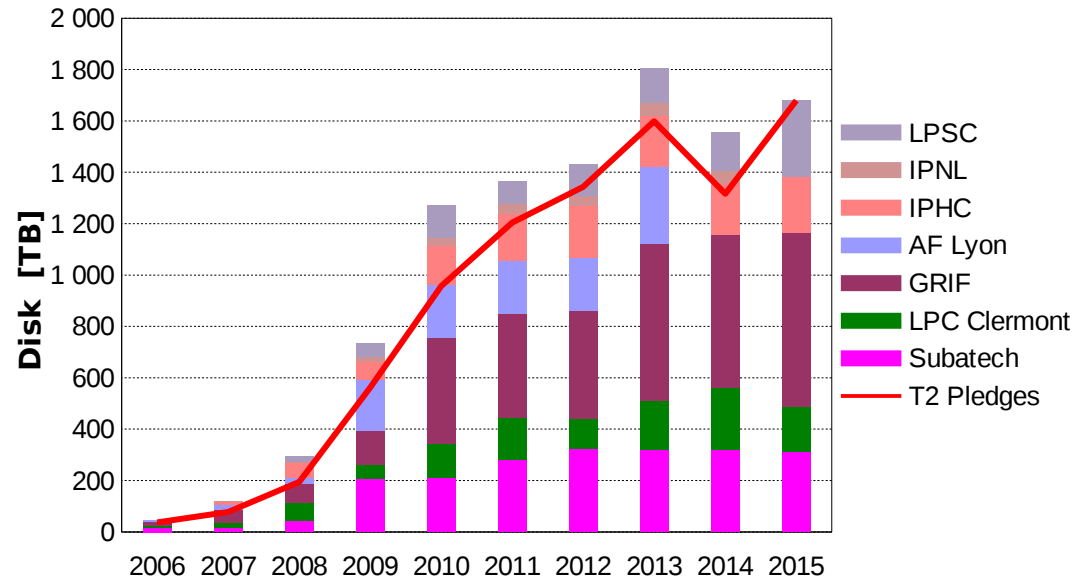
- LPSC low values explained by an issue with site BDII impacting JA submission plus an issue with the scheduling of Alice SAM jobs in the local cluster
- IPNL not shown here since, as a T3, it's availability is not measured

Resources

LCG-France sites
CPU capacity (T2+T3) available for ALICE VO



LCG-France sites
Disk capacity (T2+T3) available for ALICE VO



Comments :

The red line represents the pledges

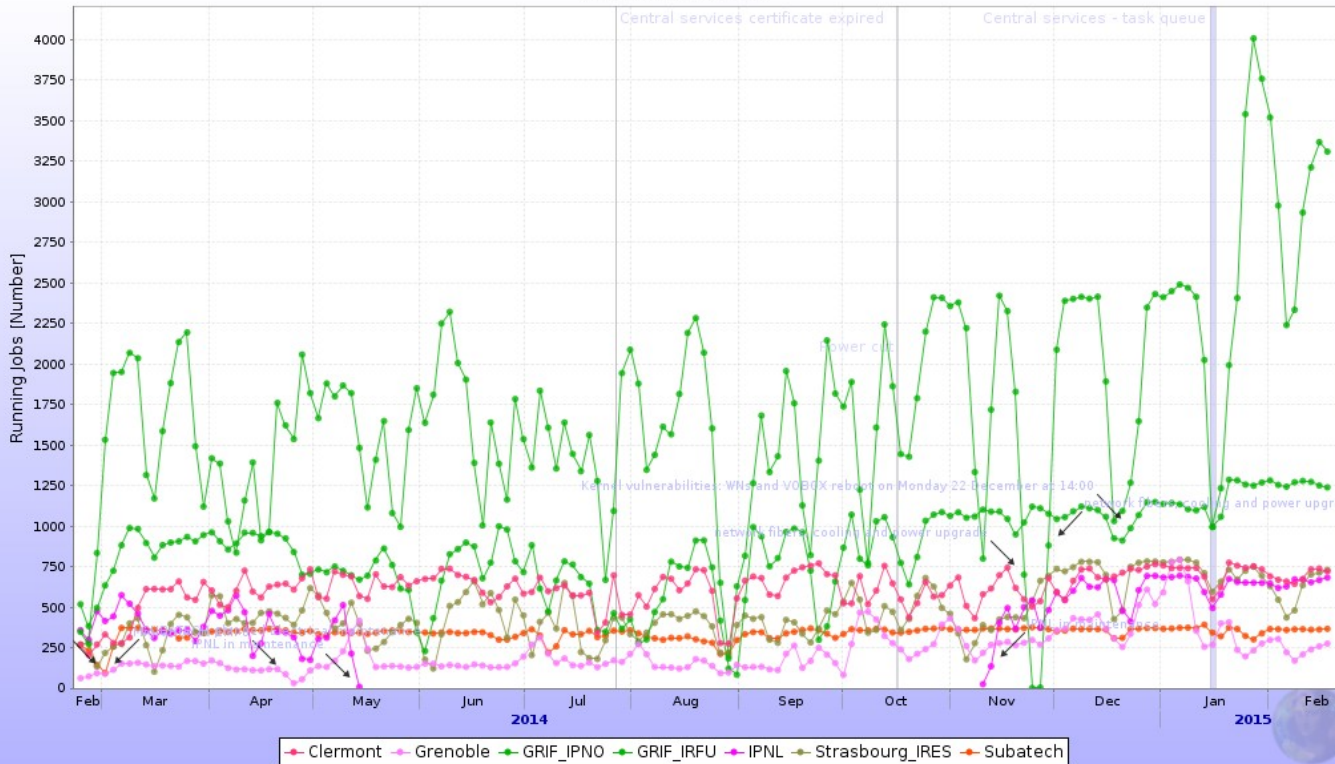
French T2s account for 9% of the CPU over all alice T2s and 7% of the disk

2014: Disk: CCIN2P3 AF/T2 resources reintegrated in T1 resources

2015: CPU: T3 resources not yet included in these graphs

Running Jobs

Running Jobs

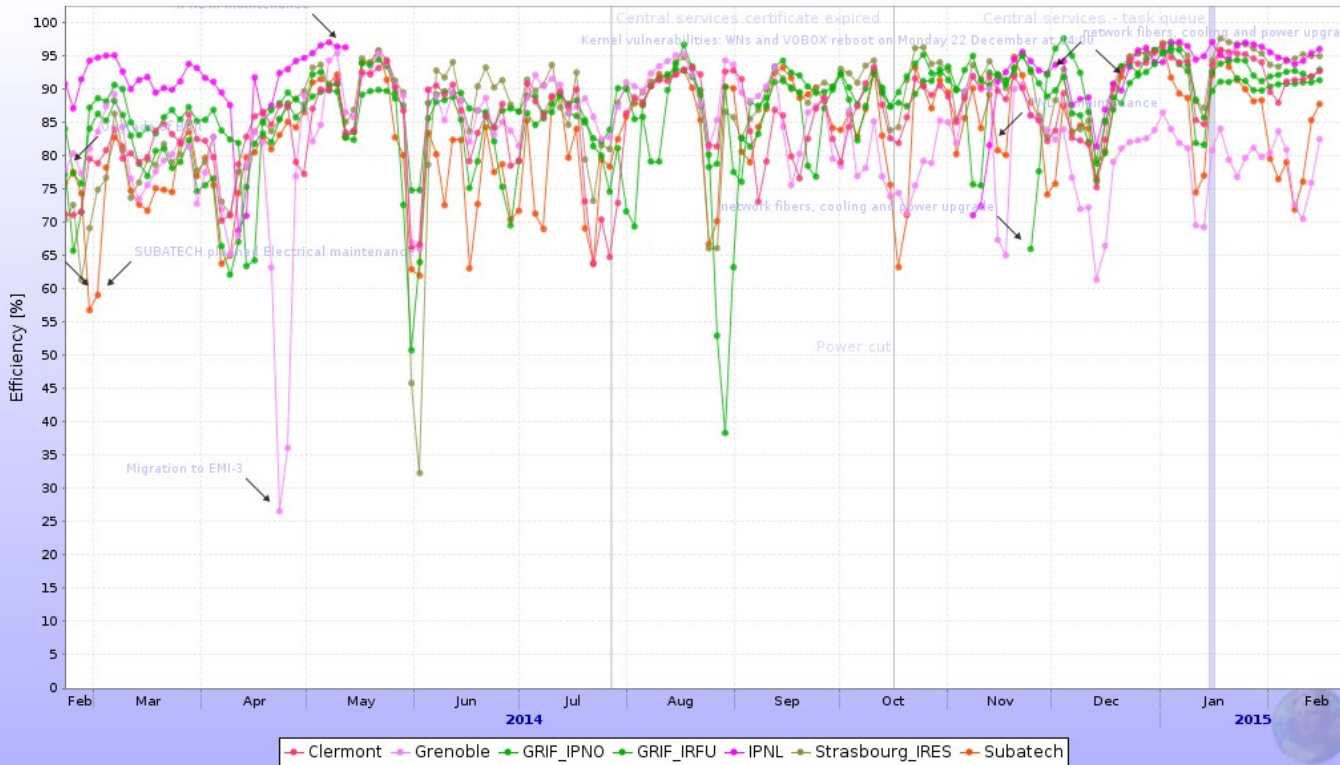


Running Jobs

	Series	Last value	Min	Avg	Max
1.	Clermont	721	0	615.7	1095
2.	Grenoble	275	0	225.4	1021
3.	GRIF_IPNO	1239	0	858.2	2392
4.	GRIF_IRFU	3310	0	1724	5750
5.	IPNL	682	0	482.1	847
6.	Strasbourg_IRES	730	0	469.8	1257
7.	Subatech	366	0	339	938
Total		7323		4714	

CPU efficiency

Jobs efficiency (cpu time / wall time)

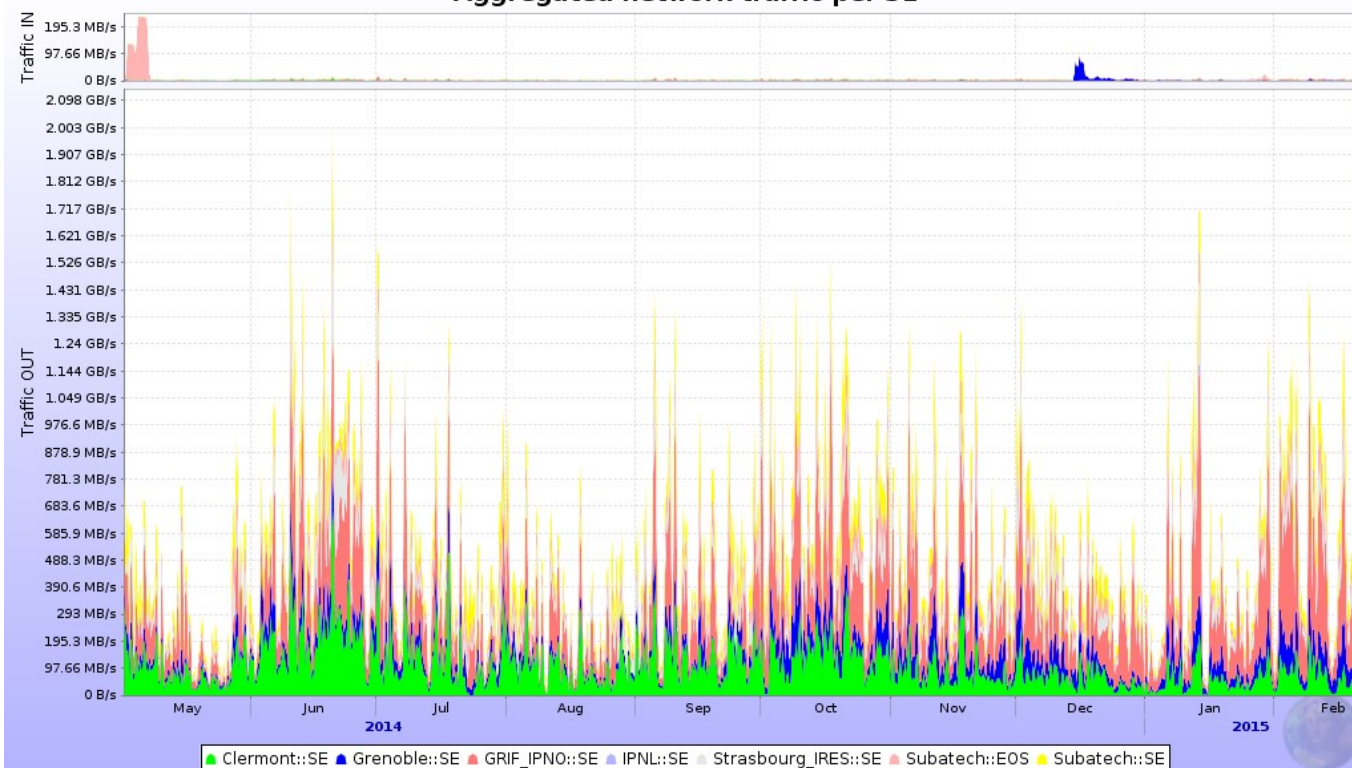


Jobs efficiency (cpu time / wall time)

	Series	Last value	Min	Avg	Max
1.	■ Clermont	92.65	0	85.82	100
2.	▲ Grenoble	82.46	0	82.17	100
3.	▲ GRIF_IPNO	92.87	0	85.26	100
4.	▲ GRIF_IRFU	91.3	0	87.43	100
5.	▲ IPNL	95.98	0	91.68	100
6.	▲ Strasbourg_IRES	94.93	0	87.33	100
7.	▲ Subatech	87.74	0	82.71	100
Total		91.13		86.06	

Storage

Aggregated network traffic per SE



From May 2014 - Feb 2015
(big spike just before May)

Comparable to the T1
(avg 670MB/s tot.19PB)

Traffic IN

	Series	Last value	Min	Avg	Max	Total
1.	Clermont::SE	314.8 KB/s	2.897 KB/s	870.4 KB/s	137.9 MB/s	20.62 TB
2.	Grenoble::SE	1.272 MB/s	0.118 KB/s	1.318 MB/s	378.9 MB/s	31.98 TB
3.	GRIF_IPNO::SE	1.438 MB/s	1.887 KB/s	1.649 MB/s	95.74 MB/s	40 TB
4.	IPNL::SE	106.7 KB/s	0 B/s	59.49 KB/s	5.788 MB/s	1.409 TB
5.	Strasbourg_IRES::SE	568.1 KB/s	2.303 KB/s	489.9 KB/s	53.48 MB/s	11.6 TB
6.	Subatech::EOS	158.6 KB/s	77.52 B/s	3.402 MB/s	284.6 MB/s	82.52 TB
7.	Subatech::SE	399.6 KB/s	0.142 KB/s	474.8 KB/s	13.25 MB/s	11.25 TB
Total		4.222 MB/s		8.219 MB/s		199.4 TB

Traffic OUT

	Series	Last value	Min	Avg	Max	Total
1.	Clermont::SE	54.26 MB/s	78.67 B/s	112.2 MB/s	1.447 GB/s	2.659 PB
2.	Grenoble::SE	34.67 MB/s	0.225 KB/s	52.28 MB/s	724.4 MB/s	1.238 PB
3.	GRIF_IPNO::SE	214.2 MB/s	44.03 B/s	175.8 MB/s	2.519 GB/s	4.165 PB
4.	IPNL::SE	6.149 MB/s	0 B/s	3.365 MB/s	344 MB/s	81.63 TB
5.	Strasbourg_IRES::SE	71.39 MB/s	0.239 KB/s	54.4 MB/s	892.3 MB/s	1.289 PB
6.	Subatech::EOS	37.76 MB/s	0 B/s	50.77 MB/s	521.4 MB/s	1.203 PB
7.	Subatech::SE	56.12 MB/s	0.116 KB/s	65.92 MB/s	673.8 MB/s	1.562 PB
Total		474.5 MB/s		514.8 MB/s		12.19 PB

GRIF Paris



- GRIF resources available for the 1/04/2015:
 - IPNO T2 3060HS06 and 338TB probably available probably earlier
T3 out of warranty material extra 2500 HS06 and 200 TB maintained as long as possible
 - IRFU (T2+T3) 3825HS06 and 423TB in production
- No specific production problem with the ALICE VO in the last months
- Foreseen evolution during the year:
 - removal of some old file server => 20 to 40 TB less on the T3 part at IPNO
 - second VO box at IRFU for the use of the ARC-CE (node16). We plan to move some worker nodes for the cream CE (node74) to the ARC-CE.
 - IPv6 deployment at IRFU

LPSC Grenoble 1/2

- Resources available on the 04/01/2015:
 - Entire resources pledged: 296 TB (native xrootd) + 2611 HSPEC06
- No specific production problem with the ALICE VO in the last months
 - Renewal of storage difficult with native xrootd
 - no native tools: rsync+link by hand -> error prone
 - Since the LHC VOs are used to measure the site reliability/availability:
 - Cron to push the alibs jobs in the queue (if not the probes fail)
- Evolutions this year:
 - 34 TB of storage out of production
 - 120 TB brand new (5 years guaranty)
- Non grid resources recently explored by ALICE physicists (acces with login/pwd):
 - Opportunistic computing resources on a local HPC-T2 centre

LPSC Grenoble 2/2

- Questions for ALICE
 - Jobs memory: ALICE jobs are high memory consumers:
 - Multi-core jobs (simple fork) coming soon ?
 - Native xrootd storage: not well suited for managing data at sites (for example drain) and to share disk servers between VOs (this is important for small sites)
 - Is there still a problem if sites use DPM-xroot ?
 - It is sometimes difficult to keep track of the useful web pages to monitor on the site; could ALICE provide a kind of site dashboard where site admins could see at a glance the status of one given site ?

LPC Clermont

Resources available on the 04/01/2015:

Entire resources pledged:

Disk : 270 TB (+50%)

CPU : 4430 HSPEC06 (+50%)

Storage Upgraded to xrootd v4.4

IPHC Strasbourg

Available resources for ALICE at the beginning of April 2015 :

- 3500 HEP-SPEC06 (already available)
- 200 TB

Evolution of the infrastructure for ALICE at IPHC in 2015:

- Removal of old disk servers (5 x 20 TB)
- Installation of new disk servers to provide 200 TB disk capacity
- IPv6 deployment (second semester)

Comments:

- Alice's jobs are running fine. No problem observed since a long time ago.
- For the new storage, should we go to EOS? What is the status of EOS for T2? Is EOS mandatory?

IPNL Lyon

T3 site => no pledges, no commitment

Resources available on the 04/01/2015:

2469 HEP-SPEC06 (26% of site capacity)

Limitation of 700 slots maximum

40TB of storage

Plans : upgrade storage from 40TB to 80TB

Problems : jobs (coming by waves) exceeding the 2GB memory

SUBATECH Nantes

- Resources available 04/01/2015 (AF not taken into account) :
 - Pledged to WLCG : 310 TB / 3000 HSPEC06
 - Currently available : 342 TB (xrootd+EOS) / 4976 HSPEC06 [1]
 - April 2015 : 434 TB (all EOS) / 4976 HSPEC06 [1]
- Issues :
- Evolutions this year :
 - Replacement of 118TB of storage (xrootd) by 210TB under EOS
 - Replacement of 10 worker nodes (flat CPU resources)
- Analysis Facility : SAF-II (see next slide)
- Joint application with CCIPL (regional HPC center) to enlarge computing and storage resources. (Cf. Laurent's talk)

[1] Based on HEP-Spec benchmark on site

SUBATECH SAF-II

- Managed by Puppet
- CVMFS based (AliEn, ROOT, AliROOT, AliPhysics)
- Simplified Staging with custom bash scripts
- Dynamic datasets (registered queries to AliEN-FC)
- « Filtering » of data
- Basic information system for loaded datasets
- Uses a Robot-certificate for authentication to AliEn

nansafmaster2.in2p3.fr

[User manual](#)

[Pie charts of disk usage by file type, data type, pass, etc...](#)

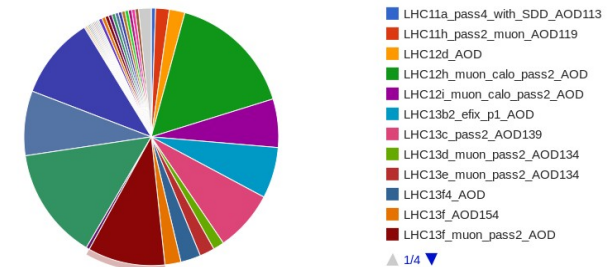
[Tree map of disk usage](#)

[Raw list of dataset groups](#)

[Data repartition by server](#)

Disk space usage details on nansafmaster2.in2p3.fr

Disk space by AOD



Common plans for all French T2s (part of LCG-France effort)

- Network : IPV6 migration of sites :
 - IPv6 deployment in France (double stack perfonar foreseen this spring, some sites already setup with IPV6). This is coordinated by the CCIN2P3 network team (Jérôme Bernier & Co)

Summary of Questions

- Related to Monalisa monitoring/reporting
 - GRIF-IRFU job efficiency [slide 19]
 - Subatech KSI2K [slide 19]
 - Site status at a glance : a dashboard ?
 - Accounting differences between Monalisa and CESGA [slide 20]
- Computing
 - Multicore jobs : foreseen ? When ?
 - Still a few memory-eating jobs
- Storage
 - DPM-xrootd acceptable ?
 - EOS still pushed ?

Questions : Alice Report

January 2015

ALICE reports

Report on ALICE sites' activity (01.01.2015 - 31.01.2015)											
Site	Group		Pledged	Delivered		Occupancy	Missing KSI2K	Efficiency	Job statistics		
		Tier	KSI2K	CPU	Wall	Wall/Pledged	Pledged - Wall	CPU/Wall	Assigned	Completed	Efficiency
11. CCIN2P3	IN2P3	T1	-	7314	8269	-	-	88.45%	666497	574383	86.18%
19. Clermont	IN2P3	T2	-	1688	1806	-	-	93.46%	155253	132553	85.38%
27. Grenoble	IN2P3	T2	-	904.3	1107	-	-	81.65%	72959	61577	84.4%
28. GRIF_IPNO	IN2P3	T2	-	2461	2657	-	-	92.63%	230499	198072	85.93%
29. GRIF_IRFU	IN2P3	T2	-	6947	7499	-	-	92.64%	300834	122887	40.85%
38. IPNL	IN2P3	T2	-	1313	1369	-	-	95.94%	60816	51540	84.75%
79. Strasbourg_IRES	IN2P3	T2	-	2068	2170	-	-	95.31%	110468	93778	84.89%
80. Subatech	IN2P3	T2	-	150.2	171.5	-	-	87.54%	90219	77232	85.61%

Why is GRIF-IRFU only 40% efficient as for the job statistics ?

Why is Subatech delivered CPU so low ? (wrong KSI2K value for the site?)

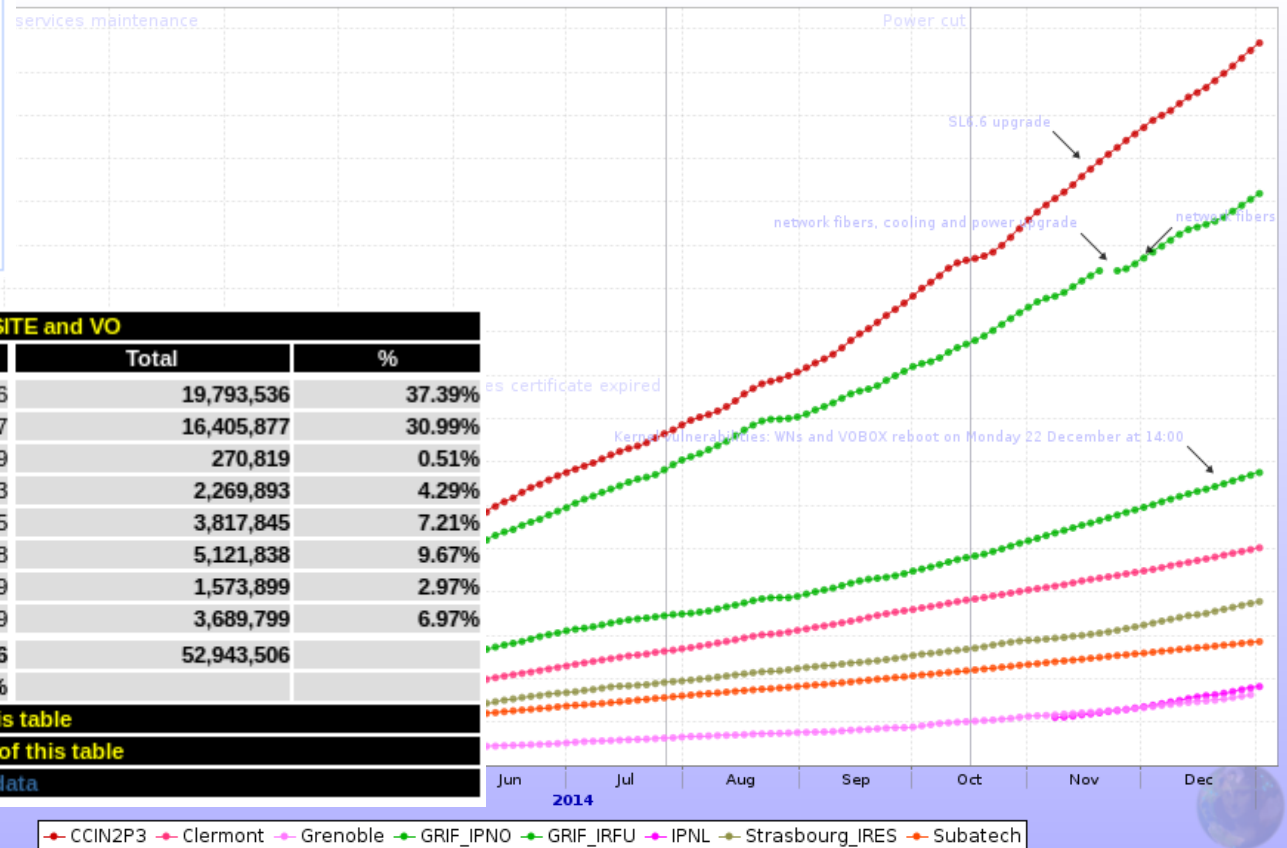
Accounting ?

January-December 2014 :

Total wall time for ALICE jobs

	Series	Last value	Min	Avg	Max
1.	CCIN2P3	16669571	11630	7185663	16669571
2.	Clermont	5020760	2075	2340447	5020760
3.	Grenoble	1628569	805	594126	1628569
4.	GRIF_IPNO	6756651	7481	3105786	6756651
5.	GRIF_IRFU	13194956	449	6116157	13194956
6.	IPNL	1818350	369	829594	1818350
7.	Strasbourg_IRES	3776766	3222	1720433	3776766
8.	Subatech	2854142	5172	1383907	2854142
Total		51719765		23276116	

Total wall time for ALICE jobs



Total elapsed time [units Hours] by SITE and VO

SITE	alice	Total	%
GRIF	19,793,536	19,793,536	37.39%
IN2P3-CC	16,405,877	16,405,877	30.99%
IN2P3-CC-T2	270,819	270,819	0.51%
IN2P3-IPNL	2,269,893	2,269,893	4.29%
IN2P3-IRES	3,817,845	3,817,845	7.21%
IN2P3-LPC	5,121,838	5,121,838	9.67%
IN2P3-LPSC	1,573,899	1,573,899	2.97%
IN2P3-SUBATECH	3,689,799	3,689,799	6.97%
Total	52,943,506	52,943,506	
Percentage	100.00%		

[Click here for a CSV dump of this table](#)

[Click here for a Extended CSV dump of this table](#)

[Click here for XML encoded data](#)

Thank you

Credits :

Thank you to all my colleagues at ALICE T2s in France for their inputs and comments