

Monitoring at GRIF

Frédéric SCHAER

Frederic.schaer .@. cea.fr

Hepix Workshop, April 27. 2012

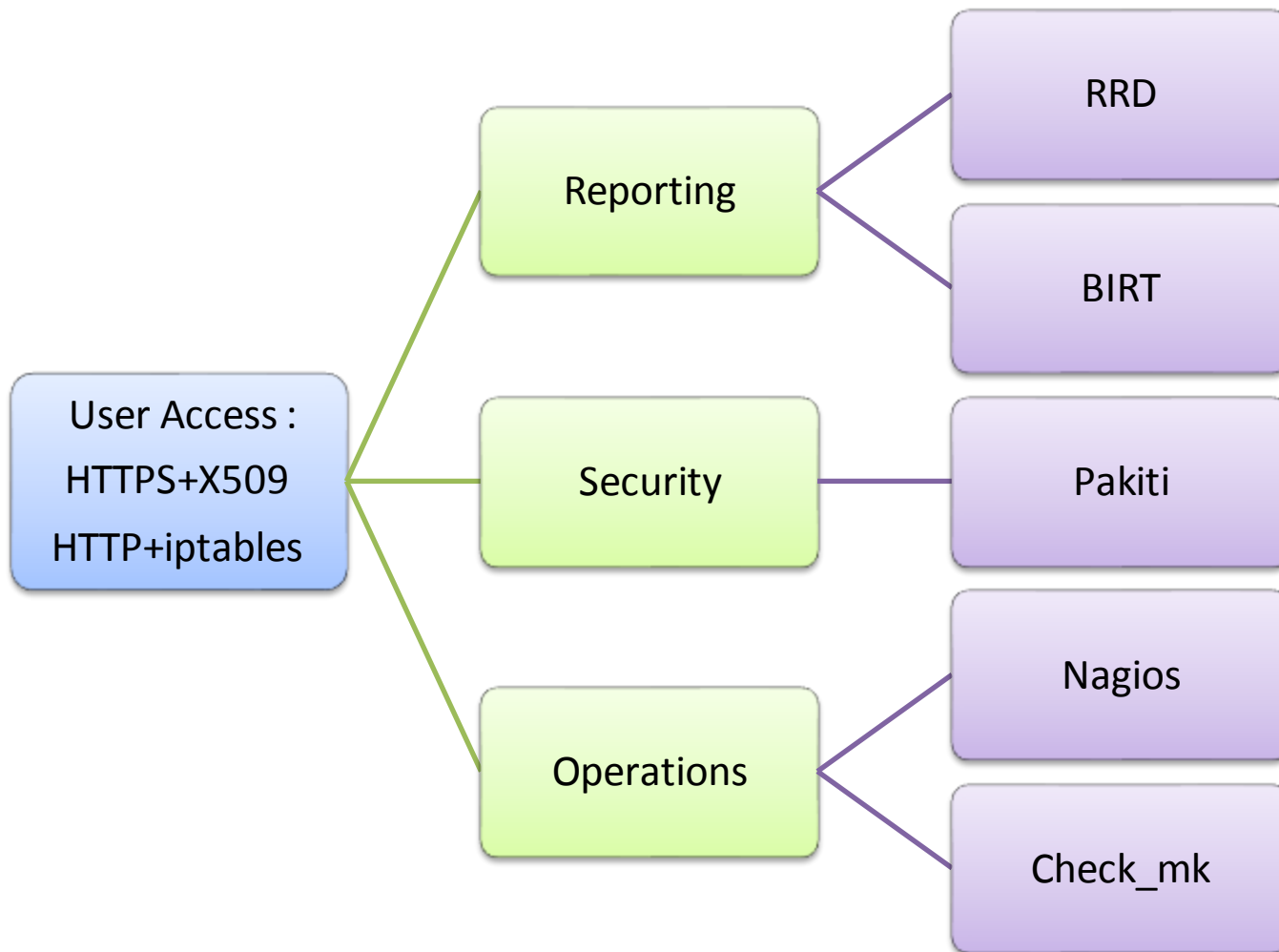
- Monitoring Software
- Monitoring Architecture
- Lessons learnt
- Foreseen evolutions



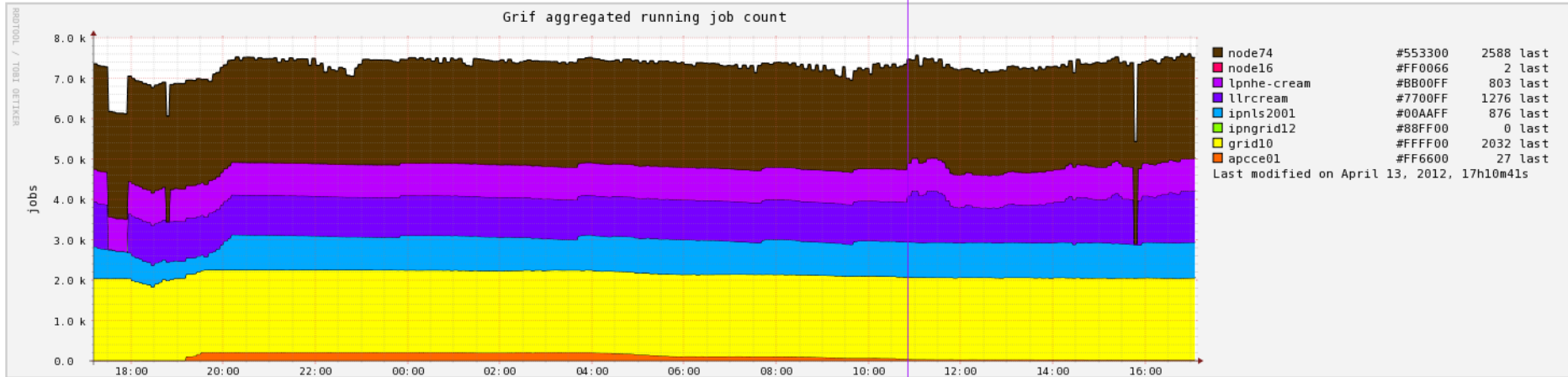
- Used in the past
 - Lemon, nagiosgraph
- In use now
 - Nagios+nrpe+pnpp+rrdcached, ipmi, pakiti,cacti, custom graphing scripts, and since 2011 : parts of check_mk
 - Vendor HW monitoring sensors : using nagios checks



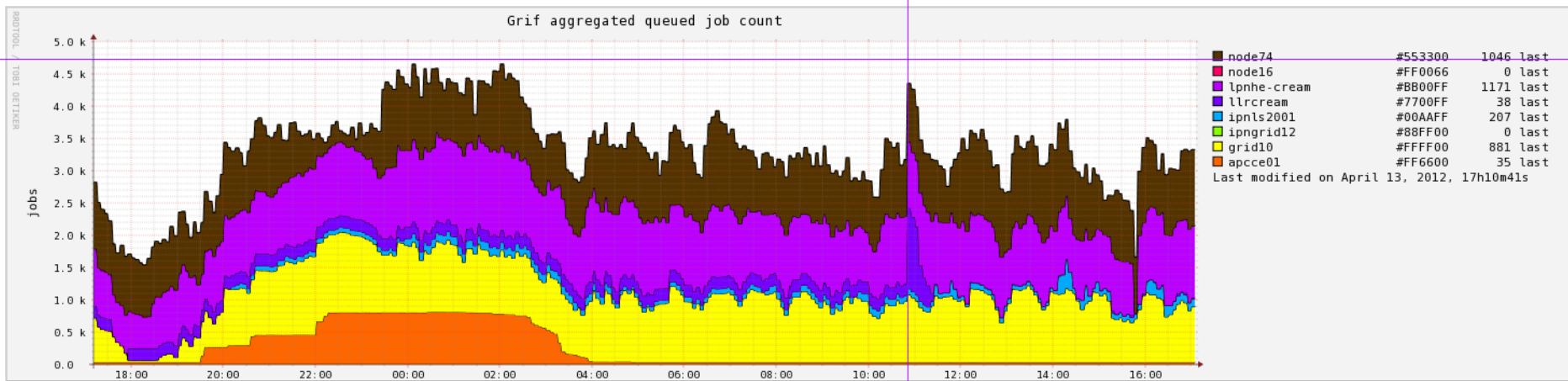
- Constraints
 - Multiple physical locations : CEA, LAL, LLR, APC, LPNHE
 - Protocol restrictions : no SSH, firewalls
 - Variety of monitored systems and configurations
 - Until recently : sl4/sl5/sl6, x86_64, i686
 - Various hardware, different capabilities
 - Number of admins
 - 1000 hosts + 16 admins = many config changes/day
 - Quattor

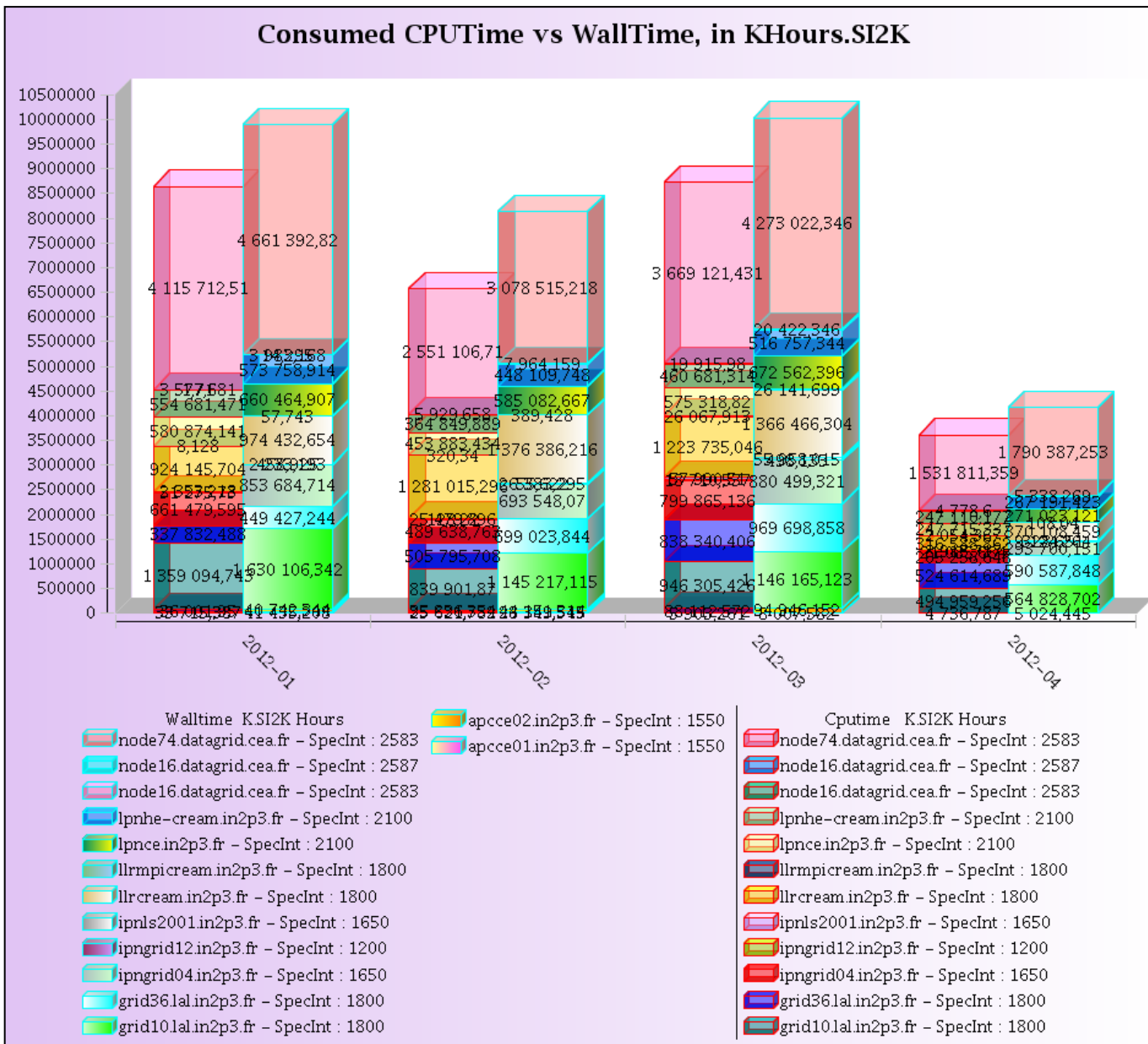


GRIF detailed running graphs



GRIF detailed queued graphs





- Pakiti setup
 - Every GRIF node runs a cron job, and reports back to the server at some random time, using SSL
 - The server
 - periodically downloads RHEL Oval patch definitions
 - Computes vulnerabilities on node report
 - Displays result to trusted users

Pakiti - Patching Status System

Navigation: [Select hosts by CVE](#) | [package](#), [Display all hosts](#) | [domains](#) [Settings](#)


[Click to select host](#)

[Click to select package](#)

[Click to select CVE](#)

Tag:

View:


Selected host:  package: all CVE: all

Host/Package name	Installed version	Required version (<i>Security repository, Main repository</i>)	CVEs (<i>Critical, Important, Moderate, Low</i>) Show/Hide CVEs
<p><i>Domain:</i> datagrid.cea.fr <i>Os:</i> Scientific Linux 6.1 (x86_64) <i>Kernel:</i> 2.6.32-220.4.1.el6.x86_64.debug</p>			<p>CVE-2011-1162 CVE-2011-1577 CVE-2011-2494 CVE-2011-2699 CVE-2011-2905 CVE-2011-3188 CVE-2011-3191 CVE-2011-3353 CVE-2011-3359 CVE-2011-3363 CVE-2011-3593 CVE-2011-4326 CVE-2011-1160 CVE-2011-1745 CVE-2011-1746 CVE-2011-1833 CVE-2011-2022 CVE-2011-2484 CVE-2011-2496 CVE-2011-2521 CVE-2011-2723 CVE-2011-2898 CVE-2011-2918 CVE-2011-4127</p>
kernel-firmware	0:2.6.32/131.12.1.el6		
perl-CGI	0:3.51/119.el6_1.1		CVE-2010-2761 CVE-2010-4410 CVE-2011-1487 CVE-2011-2939 CVE-2011-3597
perl-Compress-Raw-Zlib	0:2.023/119.el6_1.1		CVE-2010-2761 CVE-2010-4410 CVE-2011-1487 CVE-2011-2939 CVE-2011-3597
perl-Compress-Zlib	0:2.020/119.el6_1.1		CVE-2010-2761 CVE-2010-4410 CVE-2011-1487 CVE-2011-2939 CVE-2011-3597
perl-ExtUtils-MakeMaker	0:6.55/119.el6_1.1		CVE-2010-2761 CVE-2010-4410 CVE-2011-1487 CVE-2011-2939 CVE-2011-3597
perl-ExtUtils-ParseXS	1:2.2003.0/119.el6_1.1		CVE-2010-2761 CVE-2010-4410 CVE-2011-1487 CVE-2011-2939 CVE-2011-3597
perl-IO-Compress-Base	0:2.020/119.el6_1.1		CVE-2010-2761 CVE-2010-4410 CVE-2011-1487 CVE-2011-2939 CVE-2011-3597
perl-IO-Compress-Zlib	0:2.020/119.el6_1.1		CVE-2010-2761 CVE-2010-4410 CVE-2011-1487 CVE-2011-2939 CVE-2011-3597
perl-Module-Pluggable	1:3.90/119.el6_1.1		CVE-2010-2761 CVE-2010-4410 CVE-2011-1487 CVE-2011-2939 CVE-2011-3597
perl-Pod-Escapes	1:1.04/119.el6_1.1		CVE-2010-2761 CVE-2010-4410 CVE-2011-1487 CVE-2011-2939 CVE-2011-3597
perl-Pod-Simple	1:3.13/119.el6_1.1		CVE-2010-2761 CVE-2010-4410 CVE-2011-1487 CVE-2011-2939 CVE-2011-3597
perl-Test-Harness	0:3.17/119.el6_1.1		CVE-2010-2761 CVE-2010-4410 CVE-2011-1487 CVE-2011-2939 CVE-2011-3597
perl-Test-Simple	0:0.92/119.el6_1.1		CVE-2010-2761 CVE-2010-4410 CVE-2011-1487 CVE-2011-2939 CVE-2011-3597
perl-version	3:0.77/119.el6_1.1		CVE-2010-2761 CVE-2010-4410 CVE-2011-1487 CVE-2011-2939 CVE-2011-3597
phonon-backend-gstreamer	1:4.6.2/17.el6_1.1		CVE-2011-3193 CVE-2011-3194
php	0:5.3.3/3.el6		CVE-2011-0708 CVE-2011-1148 CVE-2011-1466 CVE-2011-1468 CVE-2011-1469 CVE-2011-1471 CVE-2011-1938 CVE-2011-2202 CVE-2011-2483
php-cli	0:5.3.3/3.el6		CVE-2011-0708 CVE-2011-1148 CVE-2011-1466 CVE-2011-1468 CVE-2011-1469 CVE-2011-1471 CVE-2011-1938 CVE-2011-2202 CVE-2011-2483
php-common	0:5.3.3/3.el6		CVE-2011-0708 CVE-2011-1148 CVE-2011-1466 CVE-2011-1468 CVE-2011-1469 CVE-2011-1471 CVE-2011-1938 CVE-2011-2202 CVE-2011-2483
php-gd	0:5.3.3/3.el6		CVE-2011-0708 CVE-2011-1148 CVE-2011-1466 CVE-2011-1468 CVE-2011-1469 CVE-2011-1471 CVE-2011-1938 CVE-2011-2202 CVE-2011-2483
php-ldap	0:5.3.3/3.el6		CVE-2011-0708 CVE-2011-1148 CVE-2011-1466 CVE-2011-1468 CVE-2011-1469 CVE-2011-1471 CVE-2011-1938 CVE-2011-2202 CVE-2011-2483
php-mysql	0:5.3.3/3.el6		CVE-2011-0708 CVE-2011-1148 CVE-2011-1466 CVE-2011-1468 CVE-2011-1469 CVE-2011-1471 CVE-2011-1938 CVE-2011-2202 CVE-2011-2483



	hosts	services	dependencies	Hardware state
2007-2008 : LEMON + Nagios @GRIF/IRFU only	~100	~1000	~1000	KVM Virtual Machine - Idle
2008-2009 : LEMON + Nagios @ GRIF	~600	~8000	~10000	Bi-cpu KVM VM, quite loaded
2012 : Nagios only @ GRIF	990	19259	35744	4 Cores Opteron, 8GB Ram – Dying
2012 :				
Nagios @ GRIF/LLR	169	3108	5844	8 cores Xeon, Idle
Nagios @ GRIF except GRIF/LLR	821	16151	29900	4CPU, busy

- 2 Nagios servers in GRIF for now
 - Running nagios, with MK's mklivestatus broker
 - Independant nagioses : no complex setup
 - Generated configuration : quick reinstall
 - Mainly using active checks using nrpe
 - Mklivestatus exporting data
 - using xinetd : firewalled ports : only nagios hosts allowed
 - Using unix sockets
 - MK's multisite (mod_python)
 - transparently aggregates data
 - Efficiently displays data
 - Is **fast**, amongst many other things

multisite
check  v1.1.10

Tactical Overview X

Hosts	Problems	Unhandled
990	3	3
Services	Problems	Unhandled
19259	39	36

Quicksearch X

Views X

- Hosts
 - All hosts
 - All hosts (Mini)
 - All hosts (tiled)
 - Host search
- Hostgroups
 - Hostgroups
 - Hostgroups (Grid)
 - Hostgroups (Summary)
- Services
 - All services
 - Recently changed services
 - Serv. by host groups
 - Service search
- Servicegroups
 - Servicegroups (Grid)
 - Servicegroups (Summary)
 - Services by group
- Problems
 - Alert Statistics
 - Host problems
 - Pending Services
 - Service problems
 - Unchecked services
- Addons
 - Search PNP graphs

Add snapin © Mathias Kettner

Service problems

grif (admin) 16:30 

Filter Commands Display 1 2 3 4 5 6 8 30 s 60 s 90 s ∞ Edit

CRIT						
Host	Service	Status detail	Age	Checked	Icons	Perf-O-Meter
05.in2p3.fr	SMART health	CRITICAL :	2012-03-14 02:55:09	2012-03-16 14:56:09		
dataqid.cea.fr	quattor daemons	ncm-cdispd is not running	2012-04-02 02:55:27	2 min		
2.in2p3.fr	read only filesystems	Errors detected ! This node requires urgent attention !	12 hrs	3 min		
UNKN						
Host	Service	Status detail	Age	Checked	Icons	Perf-O-Meter
in2p3.fr	SMART health	CHECK_NRPE: Socket timeout after 120 seconds.	6 hrs	6 hrs		
in2p3.fr	SPMA logfiles	CHECK_NRPE: Socket timeout after 20 seconds.	6 hrs	6 hrs		
in2p3.fr	nrpe daemon	CHECK_NRPE: Socket timeout after 10 seconds.	6 hrs	108 sec		
in2p3.fr	System checks	CHECK_NRPE: Socket timeout after 20 seconds.	6 hrs	6 hrs		
in2p3.fr	System checks	CHECK_NRPE: Socket timeout after 20 seconds.	5 hrs	5 hrs		
in2p3.fr	quattor daemons	CHECK_NRPE: Socket timeout after 20 seconds.	2 min	53 sec		
WARN						
Host	Service	Status detail	Age	Checked	Icons	Perf-O-Meter
dataqid.cea.fr	SMART health	WARNING :	2012-02-25 14:46:51	3 min		
dataqid.cea.fr	SMART health	WARNING :	2012-03-01 09:54:23	5 hrs		
dataqid.cea.fr	SMART health	WARNING :	2012-03-01 21:44:01	5 hrs		
dataqid.cea.fr	SMART health	WARNING :	2012-03-08 16:19:37	5 hrs		
024.in2p3.fr	hung nrpe processes	PROCS WARNING: 81 processes with command name 'nrpe'	2012-03-11 16:45:11	72 min		
019.in2p3.fr	hung nrpe processes	PROCS WARNING: 82 processes with command name 'nrpe'	2012-03-14 14:51:28	65 min		
135.in2p3.fr	hung nrpe processes	PROCS WARNING: 9 processes with command name 'nrpe'	2012-03-15 20:52:18	63 min		
118.in2p3.fr	hung nrpe processes	PROCS WARNING: 19 processes with command name 'nrpe'	2012-03-15 22:50:21	66 min		

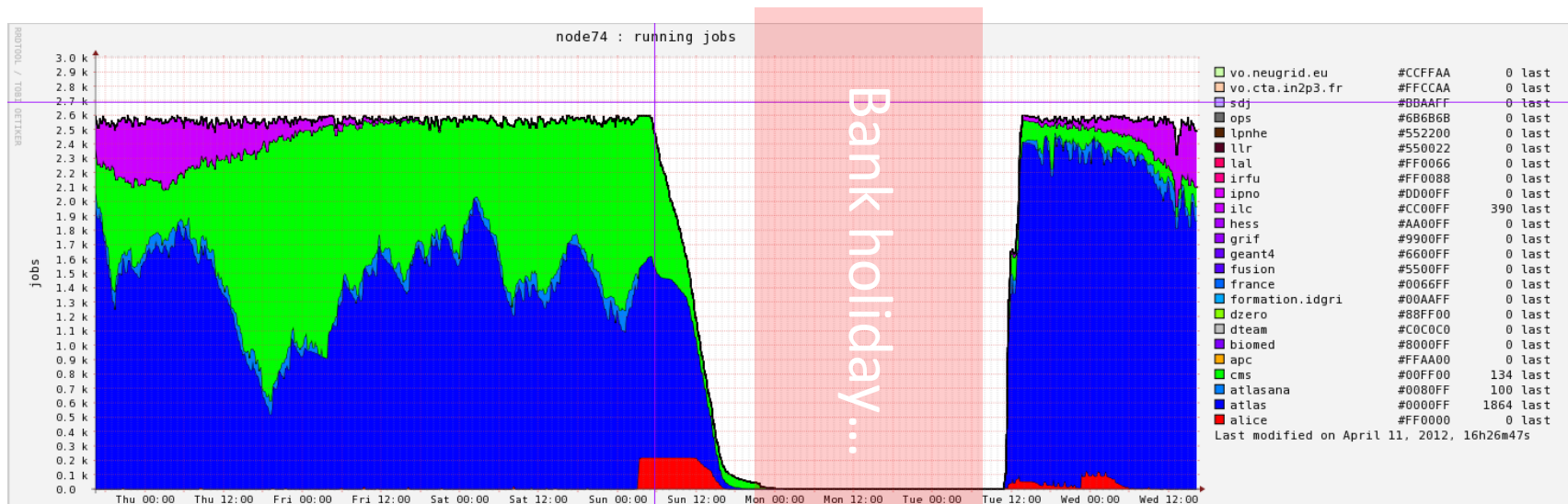
- Nagios/check_mk config
 - Automatically generated using quattor information
 - No {warning,recovery,flapping,unknown} email
 - Uses hosts dependencies (downtimes, network outages)
 - Uses services dependencies (mainly on nrpe)
 - Uses the « large setups tweaks » nagios option
 - Use rrdcached daemon for PNP
 - Required many optimizations (and always will)
- NRPE
 - Arguments forbidden
 - firewalled

- Compression :
 - `Nrpe_check output | bzip2 | uuencode`
 - `Nagios check_nrpe | bunzip2 | uudecode`
- Compressing check output allows to retrieve up to ~500K data (with op5 patch)
 - Still not enough under high load for `check_mk` agents, because of `ps` output

- Patched/package software (SL5/6 32/64):

Soft	Patch
Nagios	startup script : 15min -> 5 min Allow SELinux
Check_mk	Fix X509 SSL+FakeBasicAuth usage Change instant check retry button timeout Agents patches (disable ipmi, reduce data size, fix \$PATH) fix hardcoded configuration paths, wrong links
Nrpe	Allow up to 64K output (op5 patch) Create system user/group
Nagios plugins, GRIF plugins	
rrdtool	

- Nagios recovery actions must be temporary, and bugs must be fixed asap.
- What should not happen
 - when the almighty nagios crashed)
 - Or when a /dev/null becomes a regular file, or...



- Nagios is powerful, but has its limits :
 - with no patch, restart takes 15 minutes !
 - Excessive number of active checks are limited by the fork capabilities of the server
 - more expensive hardware (see evolutions)?
- Monitoring is an ongoing process
 - Tests output must constantly be interpreted/correlated
 - Too much monitoring kills monitoring (admins)
 - But there's always a need to monitor more things

- Divide nagios load on multiple and distributed hardware in GRIF
- Fork issue ?
 - Big nagios performance boost when using Hugepages memory (man libhugetlbf) ... at the expense of nagios crashes.
 - But Hugetlb seems worth investigation
- reduce load with “active” passive checks (using check_multi ? MK’s check_mrpe ?)
- Many things aren’t monitored yet :
 - Memory and cpu cores that just disappear, exotic filesystems configurations, new services... and more to come

- Check_mk
 - Using check_mk agents and service auto discovery might help reduce load (passive checks)
 - but agents large output is incompatible with huge server loads
 - But some MK's probes must be tweaked (autofs, virtual NICs ...)
- Rework nagios config generation, so that it can be shared with the quattor community ?

