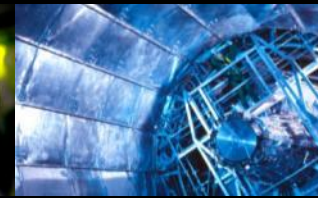


Site notifications with SAM and Dashboards

Marian Babik
SDC/MI Team

12th June 2013 GDB



Introduction

- Recent WLCG reporting proposal
 - Introduces joint T1/T2 reports
 - Removes OPS tests from reports
- This has impact on sites supporting multiple VOs and thus having interest in consolidated monitoring and notifications:
 - Reporting and VO specific tests
 - Notifications of sites in case of failures

Today

- WLCG reporting
 - Overview and recent changes
 - VO specific tests (site vs experiment view)
- Overview of the existing approaches in SAM and Dashboards
 - Notifications/alarming of site failures
- Aim is to start discussion and get feedback from sites and experiments

WLCG Reporting

- **Weekly reports - T0/T1**
 - Presented at the WLCG MB
 - Contain sites availability and list of GGUS tickets
- **Monthly reports - T0/T1/T2**
 - Broadcasted by WLCG office to MB, CB, GDB
 - Current structure
 - T0/T1 summary and detailed reports (OPS and VO specific)
 - T2 detailed report (only OPS)
- **Quarterly reports**
 - Compiled manually from the monthly reports
 - Broadcasted to OB
- **All reports are based on list of critical tests agreed by MB**
 - This set is evolving as technology and experiments needs are changing

WLCG Reporting Changes

- In January changes to current monthly reports were proposed at WLCG MB
 - Introducing joint T1/T2 reports (VO specific)
 - Effectively removing OPS reports
- Update on new reports presented at May WLCG MB
 - Contains review of T2 performance
- T2 reports are based on the list of critical tests defined by each experiment
 - The same tests are used already for some time in T0/T1 reports
 - Performance of T2s based on those tests, although not reported, already followed up by the experiments for some time

Critical Tests Overview

- Reports generated based on list of critical tests contains the following service types:
 - CE, SRM, LFC and WN
- Contains mixture of operational and experiment specific tests
 - All executed with experiments credentials
 - Operational tests supported and developed by the PTs
 - List of critical tests contains mostly operational and only few experiment specific ones (evolving, differ per experiment)
 - Those are easy to understand for sites
- Many tests are shared between experiments
 - CE job submission, SRM
 - Very positive recent development of the CVMFS probe
- List of critical tests used in the new WLCG reports
 - <http://cern.ch/go/H9hj>

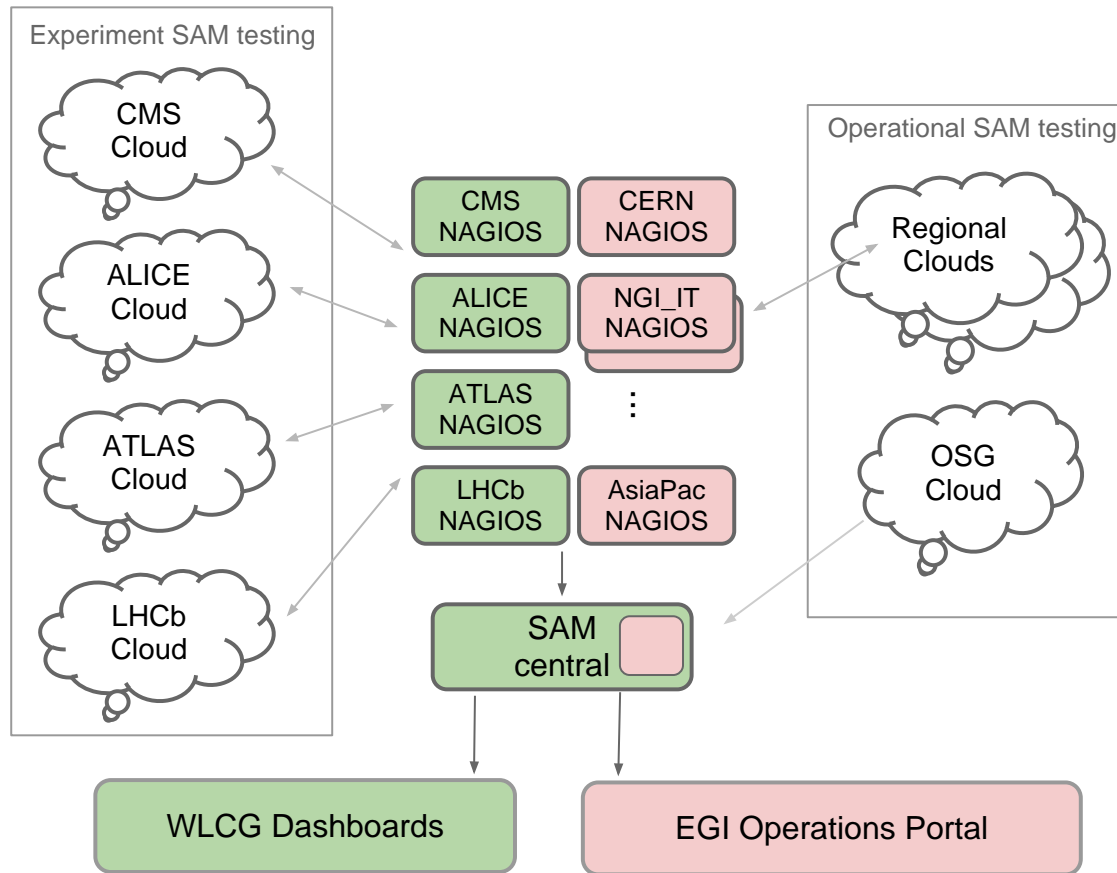
Experiments vs Sites

- Experiment vs site level tests
 - Important to understand if sites failures are due experiment or site level issues
 - Not a new topic – discussed as part of OPS TEG WG1 Monitoring and Metrics
 - Outcome was a recommendation to distinguish site vs experiment specific tests
- For ATLAS, CMS and LHCb this is already the case
 - Major revision of tests was performed in SAM
 - Different metric sets (profiles) are now used to follow up site and experiment specific issues
- ALICE is currently using only site level tests, but plans to include also experiment specific

Site perspective

- Moving from OPS to VO specific reporting has an impact on sites
 - Understand the overall situation across different VOs
 - Get notified in case of site failures affecting reported site's availability/reliability
- SDC/MI team requested to provide an overview of existing approaches in SAM and Dashboards
 - Notification infrastructure
- Aim is to start a discussion and understand what works for sites

Testing infrastructure

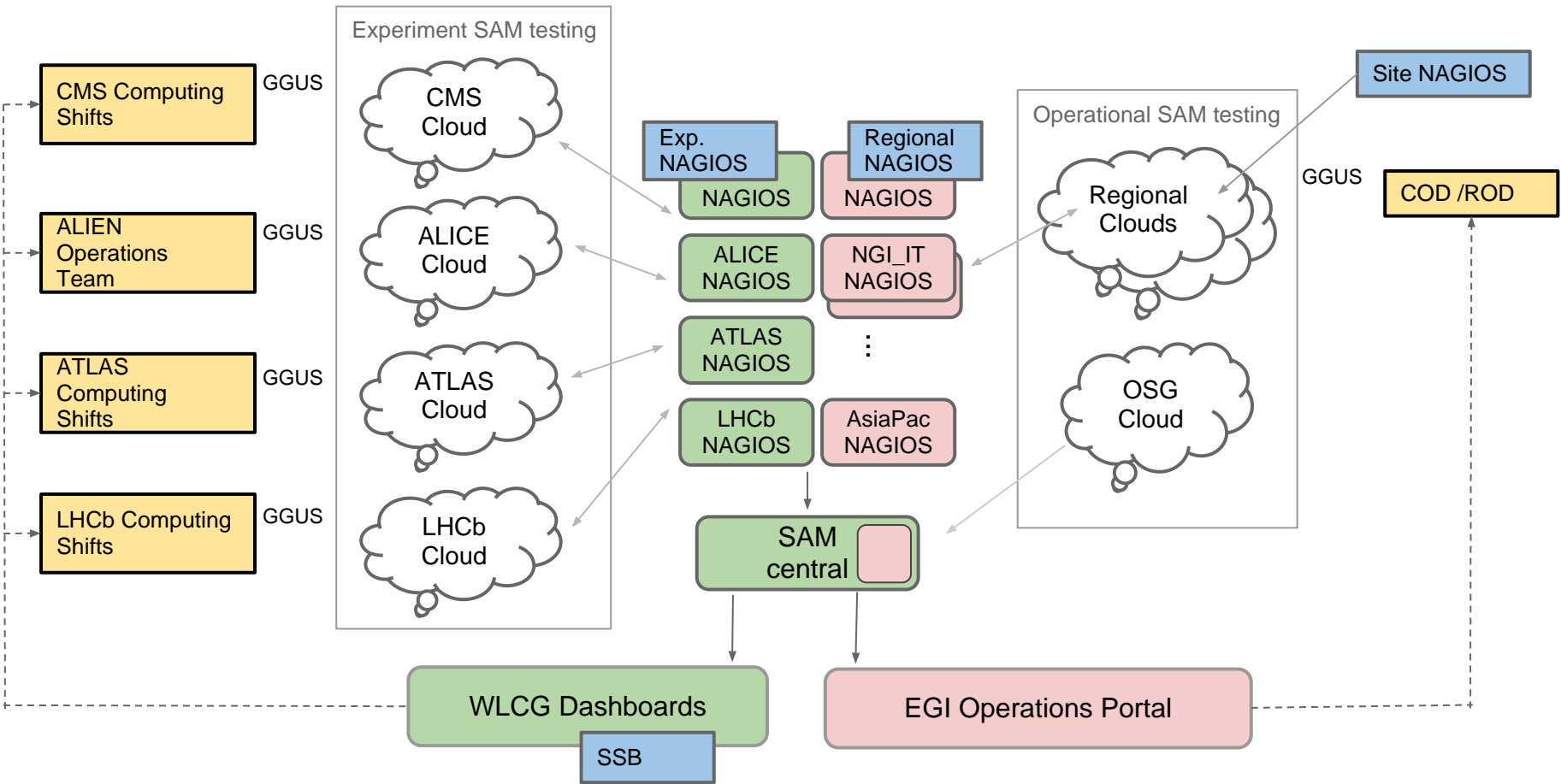


Legend

Exp. testing infrastr.

OPS testing infrast.

Notifications



Experiment Computing Shifts

- Organized to oversee experiments grid operations
- ATLAS – Computing Shifts (24/7) and ATLAS Cloud Squad
 - Actively monitoring T0/T1/T2/T3 sites using SSB (SAM, DDM, etc.)
 - Sites contacted through GGUS
 - Experiment issues tracked through Savannah
- CMS – Computing Shifts (24/7)
 - Actively monitoring T0/T1/T2 sites using SSB
 - Problems reported using Savannah (Savannah-GGUS bridge used for T1s, T2s on request)
- LHCb – Computing Shifts (8/7)
 - Monitoring T1 using SAM interface and cross-checking with DIRAC monitoring
 - If there is a problem shifter contacts GEOC to follow up with sites via GGUS
 - T2 monitoring only through DIRAC (changes are planned)
- ALICE – AliEn Operations Team
 - Active monitoring of sites usually not necessary
 - If there issues T1/T2 sites follow up through GGUS, mailing lists
 - MonALISA notification system in place including status portal

ROD/COD

- **Regional Operator on Duty**
 - Notifications received from regional NAGIOS and processed in the EGI operations portal
 - Follow up and contacts sites through GGUS
- **Central Operator on Duty**
 - Oversees ROD
 - Notified if ROD doesn't handle alarms within 72 hours and for tickets open for more than one month
 - Monthly follow up of site's A/R if below 70%/75%, requesting justification

Site NAGIOS

- Standalone solution for integrated monitoring and notification
 - Designed as an extension of an already existing NAGIOS run by the site admins
 - Offered as part of SAM (supported on best effort basis)
- Capabilities
 - Importing metric results from regional or experiment NAGIOSes
 - Re-using existing NAGIOS web interface to show consolidated view
 - Supports native NAGIOS notifications
 - e-mails sent to sysadmins with basic soft/hard state filtering
- Very few sites have shown actual interest (PIC, NIKHEF)

Site Nagios







org.atlas.CE-Pilot-submission-atlas-vo		OK	2013-06-11 10:02:35	19d 20h 28m 18s	1/2	samnag035.cern.ch: 2013-06-11 07:40:32 juk.nikhef.nl (380/420) yellow .
org.atlas.WN-FrontierSquid- /atlas/Role=lcgadmin-vo		OK	2013-06-11 08:49:17	0d 1h 21m 30s	1/2	samnag035.cern.ch: wn-val-036.farm.nikhef.nl: OK
org.atlas.WN-swspace- /atlas/Role=lcgadmin-vo		OK	2013-06-11 08:49:17	0d 1h 21m 30s	1/2	samnag035.cern.ch: wn-val-036.farm.nikhef.nl: OK
org.atlas.WN-swspace- /atlas/Role=plot-vo		OK	2013-06-11 09:26:13	20d 14h 32m 38s	1/2	samnag041.cern.ch: wn-car-016.farm.nikhef.nl: OK
org.atlas.WN-swtag- /atlas/Role=lcgadmin-vo		OK	2013-06-11 08:49:17	0d 1h 21m 30s	1/2	samnag035.cern.ch: wn-val-036.farm.nikhef.nl: OK: VO- atlas-AtlasPhysics-17.0.6.2.1-i686-slc5-gcc43-opt VO- atlas-AtlasPhysics-17.0.6.2.2-i686-slc5-gcc43-opt VO- atlas-AtlasPhysics-17.0.6.2.3-i686-slc5-gcc43-opt VO- atlas-AtlasPhysics-17.0.6.4.10-i686-slc5-gcc43-opt VO- atlas-AtlasPhysics-17.0.6.4.11-i686-slc
org.gstat.CE		CRITICAL	2013-06-11 09:57:51	208d 15h 57m 38s	4/4	CRITICAL - errors 2, warnings 0, info 0
org.lhcb.WN-sft-brokerinfo- /lhcb/Role=production-vo		OK	2013-06-11 09:14:40	53d 16h 15m 17s	1/2	samnag038.cern.ch: wn-car-067.farm.nikhef.nl: Test passed!
org.lhcb.WN-sft-csh- /lhcb/Role=production-vo		OK	2013-06-11 09:14:40	53d 16h 15m 17s	1/2	samnag038.cern.ch: wn-car-067.farm.nikhef.nl: Test CSH succeeding
org.lhcb.WN-sft-lcg-rm-gfal- /lhcb/Role=production-vo		OK	2013-06-11 09:14:40	53d 16h 15m 17s	1/2	samnag038.cern.ch: wn-car-067.farm.nikhef.nl: Test succeeded!
org.lhcb.WN-sft-vo-swdir- /lhcb/Role=production-vo		OK	2013-06-11 09:19:37	53d 16h 10m 21s	1/2	samnag038.cern.ch: wn-car-067.farm.nikhef.nl: Test fully successful
org.lhcb.WN-sft-voms- /lhcb/Role=production-vo		OK	2013-06-11 09:14:40	13d 17h 5m 57s	1/2	samnag038.cern.ch: wn-car-067.farm.nikhef.nl: Test succeeding w/o major problem
org.nagios.BDII-Check		OK	2013-06-11 09:58:08	11d 19h 8m 25s	1/4	LDAP OK - 0.005 seconds response time
org.nagios.GridFTP-Check		OK	2013-06-11 09:58:24	125d 0h 2m 41s	1/4	FTP OK - 0.019 second response time on port 2811 [220 juk.nikhef.nl GridFTP Server 6.16 (gcc64, 1352462503-83) [Globus Toolkit 5.2.3] ready.]
wierde.nikhef.nl						
Ganglia gmond		OK	2013-06-11 09:59:29	138d 18h 45m 0s	1/1	TCP OK - 0.005 second response time on port 8649
NCD result		OK	2013-06-11 10:01:40	138d 18h 32m 30s	1/1	OK: [/var/log/ncm/ncd.log.1.gz:2013/06/03-16:11:14]
Quattor daemons		OK	2013-06-11 10:03:47	138d 18h 39m 58s	1/1	OK: ncm-cdispd,cdp-listend:
free disk space /		OK	2013-06-11 10:05:57	138d 18h 47m 28s	1/1	DISK OK - free space: / 6865 MB (59% inode=72%):
free disk space /tmp		OK	2013-06-11 10:08:07	138d 18h 44m 56s	1/1	DISK OK - free space: /tmp 3690 MB (96% inode=99%):
free disk space /var		OK	2013-06-11 10:00:15	138d 18h 42m 23s	1/1	DISK OK - free space: / 6865 MB (59% inode=72%):
free swap		OK	2013-06-11 10:02:25	138d 18h 39m 54s	1/1	SWAP OK - 100% free (0 MB out of 0 MB)
hr.srce.MyProxy-CertLifetime		OK	2013-06-11 07:27:10	138d 18h 20m 51s	1/2	CERT LIFETIME OK - Certificate will expire in 276.77 days (Mar 14 23:59:59 2014 GMT)
hr.srce.MyProxy-CertLifetime-ngi		OK	2013-06-10 13:55:38	0d 20h 13m 23s	1/2	sam-ngi: CERT LIFETIME OK - Certificate will expire in 277.50 days (Mar 14 23:59:59 2014 GMT)
hr.srce.MyProxy-ProxyLifetime-ops.biggrid.nl		OK	2013-06-11 08:27:26	138d 18h 26m 45s	1/3	MyProxy OK - Certificate will expire in 6.87 days (Jun 18 03:15:04 2013 GMT)
hr.srce.MyProxy-Store- /ops/Role=lcgadmin-ngi		OK	2013-06-11 09:53:40	49d 2h 59m 40s	1/4	sam-ngi: MyProxy OK - MyProxy credential created. Querying stored credential succeeded. MyProxy credential destroyed.
hr.srce.MyProxy-Store-ops.biggrid.nl		OK	2013-06-11 09:12:43	138d 18h 33m 6s	1/4	MyProxy OK - MyProxy credential created. Querying stored credential succeeded. MyProxy credential destroyed.
kernel version		OK	2013-06-11 10:04:36	17d 22h 53m 46s	1/1	OK: kernel version = 2.6.32-358.6.2.el6.x86_64
mounted file systems		OK	2013-06-11 10:06:35	138d 18h 47m 52s	1/1	OK
org.nagios.BDII-Check		OK	2013-06-11 09:12:59	138d 18h 24m 28s	1/4	LDAP OK - 0.006 seconds response time
org.nmap.MyProxy		OK	2013-06-11 09:13:20	111d 19h 2m 26s	1/2	NMAP OK - open
system load		OK	2013-06-11 09:58:45	138d 18h 47m 20s	1/1	OK - load average: 0.00, 0.00, 0.00
time offset		OK	2013-06-11 04:30:53	431d 20h 37m 24s	1/1	OK: time offset is -5.092829s
used memory		OK	2013-06-11 10:03:03	138d 18h 47m 18s	1/1	OK - 14.9% (284612 kB) used.

Experiment NAGIOS

- Monitoring and notification at the experiment level
 - Offered from the four experiment NAGIOSes
 - Offers equivalent functionality as the Regional NAGIOS
- Capabilities
 - Native Nagios notifications can be enabled to send e-mails to site contacts (from GOCDDB/OIM or custom list)
 - Contains soft/hard state filtering before notification is sent
 - Can be fine-tuned per site, service type, service
 - The same information can be also sent via message bus
 - Re-using the way Regional NAGIOS informs EGI Operations Portal
- Notifications already enabled for some CMS sites

Experiment Nagios

- Filtering quite complex - usually many notifications are sent
 - Currently nothing is sent if tests are not executed – freshness mechanism can be enabled to resolve this, but caused flapping in the past
 - Filtering notifications only for critical metrics currently not implemented

ce1.ipgp.fr		org.lhcb.WN-sft-brokerinfo- /lhcb/Role=production	?	OK	01-22-2013 02:12:34	402d 15h 46m 13s	1/2	wn102.cluster.dmz8-grid-net.ipgp.fr: Test passed!
		org.lhcb.WN-sft-csh- /lhcb/Role=production	?	OK	01-22-2013 02:12:35	329d 20h 10m 50s	1/2	wn102.cluster.dmz8-grid-net.ipgp.fr: Test CSH succeeding
		org.lhcb.WN-sft-lcg-rm-gfal- /lhcb/Role=production	?	OK	01-22-2013 02:12:36	402d 15h 46m 11s	1/2	wn102.cluster.dmz8-grid-net.ipgp.fr: Test succeeded!
		org.lhcb.WN-sft-vo-swdir- /lhcb/Role=production	?	OK	01-22-2013 02:16:07	402d 15h 42m 40s	1/2	wn102.cluster.dmz8-grid-net.ipgp.fr: Test fully successful
		org.lhcb.WN-sft-voms- /lhcb/Role=production	?	OK	01-22-2013 02:12:38	329d 20h 10m 47s	1/2	wn102.cluster.dmz8-grid-net.ipgp.fr: Test succeeding w/o major problem
		org.sam.CREAMCE-JobState- /lhcb/Role=production		OK	06-11-2013 10:28:55	0d 9h 11m 4s	1/2	OK: [Running] https://wms305.cern.ch:9000/Hm2MCTxj-ZjrTzaog8zehQ
		org.sam.CREAMCE- JobSubmit- /lhcb/Role=production	? 	OK	06-11-2013 09:53:55	0d 8h 20m 51s	1/2	OK: success.
		org.sam.glxec.CE-JobState- /lhcb/Role=pilot		OK	06-11-2013 10:30:44	2d 12h 10m 32s	1/2	OK: [Running] https://wms306.cern.ch:9000/e2WmphMtpYZuyPvzkTgjyg
		org.sam.glxec.CE- JobSubmit-/lhcb/Role=pilot	? 	OK	06-11-2013 09:05:45	2d 11h 52m 49s	1/2	OK: success.
		org.sam.glxec.WN-gLExec- /lhcb/Role=pilot	? 	WARNING	06-11-2013 09:46:45	213d 21h 32m 30s	2/2	wn2011.plgrid.icm.edu.pl: WARNING: authorization error (203)

Site Status Board Notifications

- Monitoring and notification at the experiments level
 - Offered by the experiment SSBs deployed at CERN
- Capabilities
 - Multiple profiles, defined by experiment experts
 - Includes SAM tests, Job Information, DDM, ...
 - Site admins can subscribe to desired information
- Notifications sent when site changes status
 - Monitoring part heavily used by ATLAS & CMS
 - ~20 CMS site admins already registered

Site Status Board Notifications

Help Login Register for site notifications Site Status for the CMS sites, v0.8.0_rc1

Index Expanded Table

Existing views

View Id	View	Columns	Visible		
3	site_readiness	Commissioned Links,Readiness last week_3month,HC glidein,Site SAM availability,Topology/Maintenances,HC,IsSiteInSiteDB?,Good links combined,SiteReadiness: Status	ON	modify	remove
4	expert	Pledged, TB, SiteSize, Pledged slots	ON	modify	remove
6	maint	Maint Info,Topology/Maintenances,BDII Topology,All maintenances with topology,Topology/Maintenance in SAM	ON	modify	remove
7	transfers	Transfer_rate_phedex,To_T3_phedex,To_T2_phedex,To_T1_phedex,To_T0_phedex,Out_rate_phedex,In_rate_phedex	ON	modify	remove
8	jobrobot_debug	JobRobot_1h,JobRobot_24h	ON	modify	remove
9	storage	Custodial (TB),Installed tape (TB),Installed Disk (TB),Total tape (TB),Total disk (TB),Free tape (TB),Used tape (TB),Tape Pledge (TB),Free disk (TB),Used disk (TB),Disk pledge (TB),Incoming (TB),Non-custodial (TB)	ON	modify	remove
10	test	Commissioned T2	ON	modify	remove
11	default	Visible,glidein HammerCloud,HammerCloud,SAM SRM,SAM CE,Topology/Maintenances,Savannah CMS,Out_rate_phedex,In_rate_phedex,SiteIssues,Under investigation,Commissioned Links,Pending,Running,Analysis/Production	ON	modify	remove
12	sam_sum	CE,SAM SRM,SAM CE,SRM	ON	modify	remove
13	production	Topology/Maintenances,avg 24h app proc jobSuccess,Avg 24h grid proc jobSuccess,Avg 24h Running jobs,Avg 24h Pending jobs,Overall slots,24h Alarm,Instant Alarm,8h Alarm,Glidein Alarm,RunJobs divided by Pledge,Prod status,Pending jobs,Running jobs,Prod Pledged slots	ON	modify	remove
14	Pledges	Prod Pledged slots,Federation,Overall slots,RunJobs divided by Pledge	ON	modify	remove

Notifications Summary

Notifications	Pros/Cons
Site Nagios	<ul style="list-style-type: none">• Hosted locally at sites• Native NAGIOS notifications and filtering• Possibility to add custom site-only metrics 😊• Doesn't contain detection of missing results or critical metrics filter 😞
Experiment Nagios	<ul style="list-style-type: none">• Hosted centrally• Native NAGIOS notifications and filtering• Notifications accessible via message bus• Doesn't contain detection of missing results or critical metrics filter 😞
Site Status Board	<ul style="list-style-type: none">• Hosted centrally• Provides both dashboard and notifications• Metrics defined by Experiment expert 😞• Contains detection of missing results and can support filtering of notifications only for critical metrics 😊

Plans

- Many existing monitoring and notification endpoints
 - The actual use varies depending on endpoint
- Feedback from sites needed to understand what is their preference
- Common WLCG monitoring project aiming at critical analysis of what is monitored, technologies used, deployment and support models
 - Reduce complexity of the system
 - Ensure simplified but effective operations and support
 - Unify deployment strategy, dev. process and components implementation (where possible)

References and Credits

- Credits to WLCG Operations Coordination and Commissioning Team
- WLCG Reports
 - New T1/T2 monthly reports: <http://cern.ch/go/CK6w>
- SAM References
 - SAM Documentation: <http://cern.ch/go/c7s9>
 - SAM Reporting Web: <http://cern.ch/go/hq9f>
- Dashboard Links
 - ALICE SSB: <http://dashb-alice-ssb.cern.ch>
 - ATLAS SSB: <http://dashb-atlas-ssb.cern.ch>
 - CMS SSB: <http://dashb-ssb.cern.ch>
 - LHCb SSB: <http://dashb-lhcb-ssb.cern.ch>

Backup

SAM/Dashboard APIs

- Possibility to access raw results, service/site status and service/site availability
- Capabilities:
 - Web API (JSON, XML)
 - Includes logic to detect missing results
 - Metric sets filtering is possible
- Not a subscription service (notification system), but rather point of access
 - Requires post-processing (development) on the client
 - Quite popular according to Web API statistics
- Hosted centrally
 - SAM Web API specification - <http://cern.ch/go/c6zv>