



# WLCG Service Report

**Maria.Dimou@cern.ch**  
(for the SCOD team)

~ ~ ~

**WLCG Management Board, 19<sup>th</sup> March 2013**

Thanks to Alexandre Beche, Ivan Dzhunov and Eddie Karavakis for the reliability reports.  
GGUS slides by Maria Dimou

# Summary from all daily meetings

- **Jan 22 KIT general routing problem for LHCOPN.**
- **Jan 29 OSG users recommended to start using the new DigiCert CA for new certificates instead of the DOEGrids CA. WLCG Catch All RA migration completed.**
- **Jan 30 Audioconf connections cut multiple times due to general problem with the CERN tel. lines.**
- **Feb 04 Clarifications needed by some sites about the EMI-2 tar ball and instructions for upgrade.**
- **Feb 04 DN-to-user mapping in EOS broken for 2 days due to the unavailability of a mapping file on AFS, maintained by CMS.**
- **Feb 04 LHCb reported a NAGIOS test problem for swdir at IN2P3. Investigation lasted the whole month. Solution says "problem disappeared".**
- **Feb 25 ASGC, LHCOPN, and LHCONE communication resumed. ASGC network problem solved.**
- **Mar 02 Many failing transfers amongst Canadian sites due to expired CRL.**

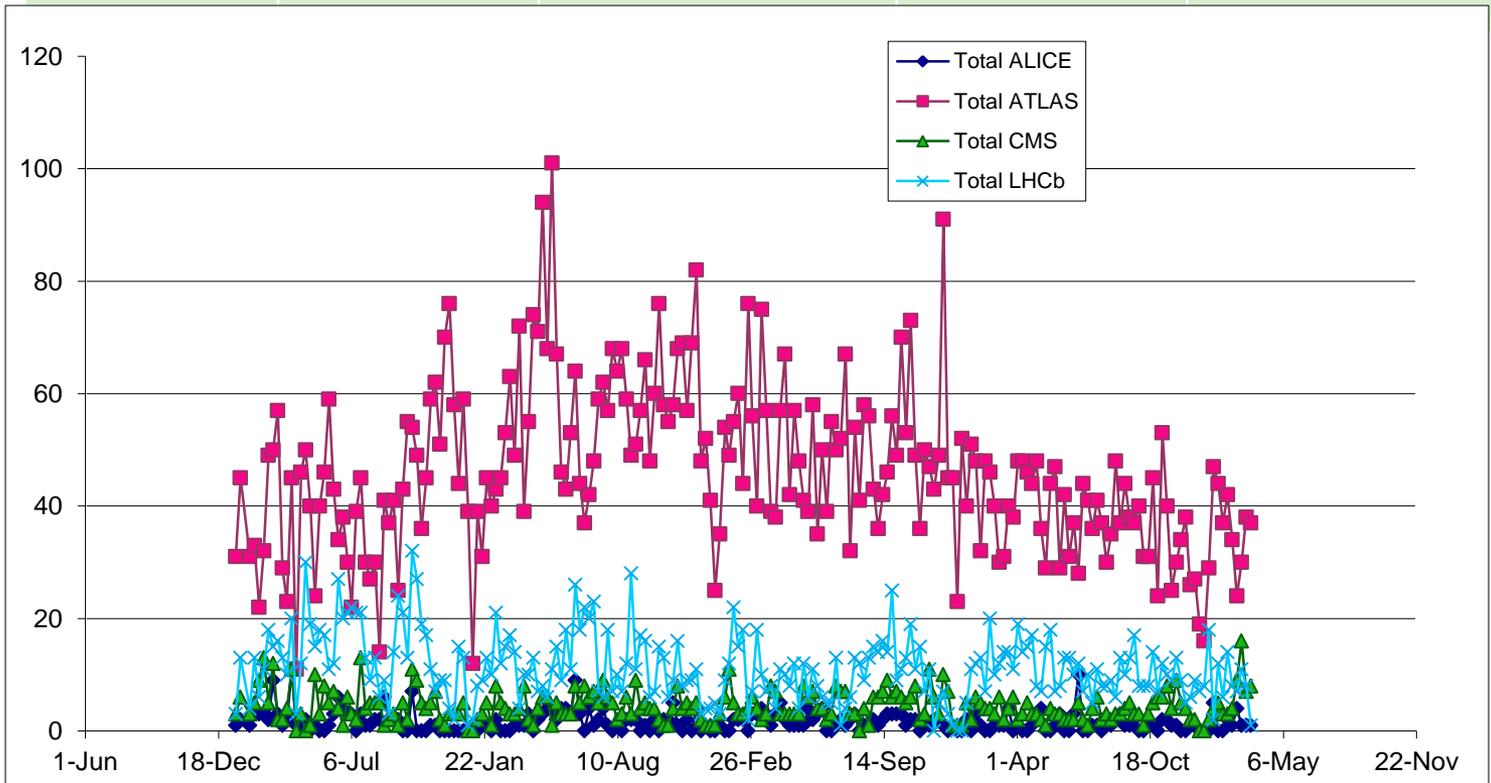
# Summary II

**Three SIRs received, all for CERN:**

- **Jan 22 Duration: 8hrs CASTOR DB overload causing transfer slowness, mainly affecting CMS.**
- **Feb 10 Duration: 8hrs Batch system was down (unavailable for users), then dispatched jobs too slowly due to LSF Master Daemon Crash.**
- **Feb 21-22 Duration: 16hrs The host certificate for voms.cern.ch was renewed with the wrong name, causing a service downtime.**

## GGUS summary ( 8 weeks )

VO	User	Team	Alarm	Total
ALICE	2	7	1	10
ATLAS	43	237	6	286
CMS	37	15	3	55
LHCb	13	53	1	67
Totals	95	312	11	418



# Question:

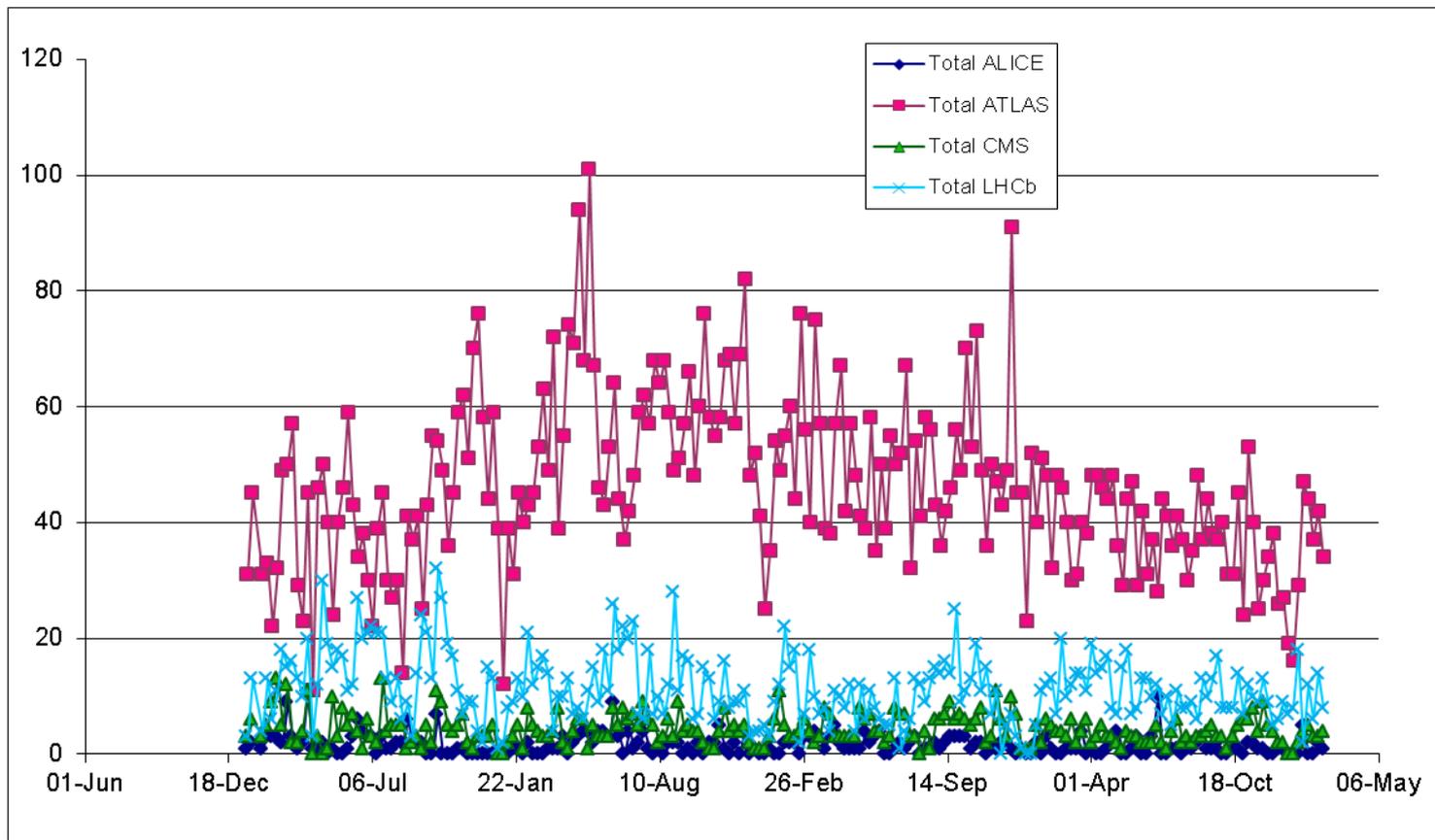
**Does the MB agree with the suggestion to start a new file ggus-tickets.xls since the LS1 start?**

**See files attached files to page:  
[WLCGOperationsMeetings twiki.](#)**

# GGUS TICKETS 4 YEARS

LHC experiments' activity in GGUS during LHC run

2009/01/13-2013/02/17



3/19/2013

WLCG MB Report WLCG Service Report

# Support-related events since last MB

- There have been **11 real ALARMS** since the **2013/01/22 MB**.
- **1** was submitted by **ALICE**, **1** by **LHCb**, **3** by **CMS** and **6** by **ATLAS**.
- **3** of these **ALARMS** were submitted during the weekend.
- **10** concerned the **CERN** site and **1** was for **KIT**.
- **2 GGUS Releases** took place since the last MB, on **2013/01/23 & 2013/02/27**. All **ALARM** tests were successful (operators received notification, reacted within minutes, interfaces worked, experts closed promptly). Details in [\*\*Savannah:135284\*\*](#) & [\*\*Savannah:136108\*\*](#)

# Support-related events since last MB

## II

**A meeting with EGI, OSG, WLCG security, support and middleware participation was held at CERN on 2013/02/05 on EGI request to investigate additional methods to authenticate to GGUS and other Grid applications.**

**Today we offer a username and passwd login but we require the registration of the user's DN for GGUS.**

**The conclusion was to develop a registration method based on vetting by trusted IDentity Providers. The certificate-based authentication, of course, remains valid.**

**Details in <https://savannah.cern.ch/support/?132872> agenda and minutes here.**

# ATLAS ALARM->CERN FILES INACCESSIBLE ON CASTOR

GGUS:90759

What time UTC	What happened
2013/01/23 15:13	GGUS ALARM ticket opened, automatic email notification to <b>atlas-operator-alarm@cern.ch</b> AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. <b>Type of Problem: File Access.</b>
2013/01/23 15:19	Supporter comments that investigation started.
2013/01/23 15:20	Operator records in the ticket that CASTOR piquet was called.
2013/01/23 16:12	Controller problem found with the host. ITCM ticket opened for the intervention.
2013/01/23 17:28	After controller reset and while still in need to be replaced, the files mentioned by the submitter were found unrecoverable.
2013/01/28 07:40	Ticket 'solved' after exchange of 10 comments. 7 files were unrecoverable but replicas were found, for some on the Grid. Submitters were happy

# ATLAS ALARM->CERN ACRONTAB PROBLEMS GGUS:90928

What time UTC	What happened
2013/01/28 12:01	GGUS ALARM ticket opened, automatic email notification to <b>atlas-operator-alarm@cern.ch</b> AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. <b>Type of Problem: Monitoring.</b>
2013/01/28 12:25	Operator records in the ticket that the service has been contacted.
2013/01/28 16:08	After 5 comments' exchange the submitter confirms the service is back. Ticket was done by the T0 service mgr to the afs experts. Problem was due to the hypervisor which turned the root fs to readonly.
2013/01/30 15:26	Another operator records in the ticket that the acron service has been contacted. Why?
2013/01/30 16:23	Ticket 'solved' & immediately 'verified'. Service migrated to the failover node.

# ATLAS ALARM-> FZK TRANSFERS TO OTHERS FAIL GGUS:91316

What time UTC	What happened
2013/02/08 20:16	GGUS TEAM ticket opened, automatic email notification to <b>lcg-admin@lists.kit.edu</b> AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. <b>Type of Problem: Monitoring.</b>
2013/02/09 02:28 SATURDAY	3 more comments by more shifters show the problem persists. Raising ticket priority & pressing the GGUS escalation button to Support Unit NGI_DE.
2013/02/09 07:50	TEAM-to-ALARM ticket upgrade. Email sent to kit-alarm@scc.kit.edu .
2013/02/09 21:05	Site mgrs contributed 16 comments to keep the submitter informed. A network problem was the culprit in the end. Routing misbehaviour not obvious why...
2013/02/09 21:05	Ticket 'solved' & on Monday 'verified'. All shifters agree the problem is gone.

# CMS ALARM-> CERN MAIN TIER0 NODE DOWN GGUS:91325

What time UTC	What happened
2013/02/09 19:38 SATURDAY	GGUS TEAM ticket opened, automatic email notification to <a href="mailto:grid-cern-prod-admins@cern.ch">grid-cern-prod-admins@cern.ch</a> AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. Ticket immediately upgraded to ALARM. Email sent to cms-operator-alarm@cern.ch. <b>Type of Problem: Other.</b>
2013/02/09 19:49	Operator comments in the ticket that no Lemon alarm was received for this node vocms15! It is pingable but doesn't accept connects. It shows status 'maintenance'.
2013/02/09 19:56	Operator comments the machine is back in prod.
2013/02/11 08:45	Ticket 'solved' on Monday by a grid service mgr. Operator's action was enough.

# CMS ALARM-> CERN LSF DOWN

## GGUS:91328

What time UTC	What happened
2013/02/10 08:53 SUNDAY	GGUS TEAM ticket opened, automatic email notification to <a href="mailto:grid-cern-prod-admins@cern.ch">grid-cern-prod-admins@cern.ch</a> AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. <b>Type of Problem: Local Batch System.</b>
2013/02/10 09:18	Ticket upgraded to ALARM. Email sent to <a href="mailto:cms-operator-alarm@cern.ch">cms-operator-alarm@cern.ch</a> . Operator comments "the ticket was sent to the service able to help".
2013/02/10 19:05	Grid service mgr, after 5 comments' exchange with the submitter to check service quality with various tests, including Hammercloud and CAF set the ticket to 'solved'.
2013/02/11 16:02	To make these drills, we checked with the service mgr and learnt that an LSF reconfiguration achieved this solution.

# ATLAS ALARM->CERN LSF DOWN

## GGUS:91329

What time UTC	What happened
2013/02/10 09:02 SUNDAY	GGUS ALARM ticket opened, automatic email notification to <b>atlas-operator-alarm@cern.ch</b> AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. <b>Type of Problem: Local Batch System.</b>
2013/02/10 09:02	Operator records in the ticket that the service has been contacted, while the submitter added attachments with submission rates.
2013/02/10 14:10	After 9 comments' exchange with the same service mgr who handled the 'simultaneous' CMS ALARM the submitter confirms the service is back. In the meantime, LSF was back but several queues remained stuck for some time and scheduling was not working.
2013/02/10 19:07	Ticket 'solved' & soon afterwards 'verified'. LSF re-configuration fixed the problem.

# CMS ALARM->CERN XRDCP FAILS FOR CASTOR FILES GGUS:91380

What time UTC	What happened
2013/02/11 23:05	GGUS TEAM ticket opened, automatic email notification to <b>grid-cern-prod-admins@cern.ch</b> AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. <b>Type of Problem: File Access.</b>
2013/02/11 23:38	Ticket upgraded to ALARM because of imminent data taking. Email sent to <a href="mailto:cms-operator-alarm@cern.ch">cms-operator-alarm@cern.ch</a>
2013/02/11 00:00	Operator records in the ticket that the CASTOR piquet has been contacted.
2013/02/12 00:12	Expert comments that SLS stats won't be reliable until the morning, a fix is foreseen. Concerning true production situation, one specific diskserver is much too busy, a problem seen again in the recent past
2013/02/13 13:49	Ticket was re-opened twice and finally set to 'solved' after 12 comments' exchange. CMS has to revisit their file access workflow. A single test file accessed 450 times within 1h40mins.

# ATLAS ALARM->CERN LSF NO JOBS DISPATCHED GGUS:91501

What time UTC	What happened
2013/02/13 21:03	GGUS ALARM ticket opened, automatic email notification to <b>atlas-operator-alarm@cern.ch</b> AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. <b>Type of Problem: Local Batch System.</b>
2013/02/13 21:14	Operator records in the ticket that the service has been contacted, while the submitter added attachments with submission rates.
2013/02/13 22:33	After 4 comments' exchange with the service manager the submitter agreed that the job dispatching time was never more than 5mins and that no action was needed.
2013/02/14 10:50	Ticket 'solved' & soon afterwards 'verified'. Nothing was needed to be done. No real problem was found.

# LHCB ALARM->CERN AFS-HOSTED WEB SERVICE DOWN GGUS:91690

What time UTC	What happened
2013/02/21 12:36	GGUS ALARM ticket opened, automatic email notification to <b>lhcb-operator-alarm@cern.ch</b> AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. <b>Type of Problem: File Transfer.</b>
2013/02/21 12:53	Operator records in the ticket that the 'right' service has been contacted.
2013/02/21 12:54	AFS service expert finds no problem with the server and re-assigns to web support.
2013/02/21 13:25	T0 site mgr prompts the escalation to 3 <sup>rd</sup> Line web support as this is an ALARM.
2013/02/21 16:32	4 comments were exchanged, most of the debugging was done offline between submitter and web service expert. Reason was a given machine doing too many downloads and slowing down the service. Ticket 'solved' & 'verified' the next day.

# ALICE ALARM->CERN VOMS SERVER

## WRONG CERT GGUS:91706

What time UTC	What happened
2013/02/21 18:56	GGUS TEAM ticket opened, automatic email notification to <b>grid-cern-prod-admins@cern.ch</b> AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. <b>Type of Problem: Other.</b> Ticket upgraded to ALARM 6 mins later. Email sent to <a href="mailto:alice-operator-alarm@cern.ch">alice-operator-alarm@cern.ch</a>
2013/02/21 19:23	Operator records in the ticket that the 'service in charge' has been contacted.
2013/02/22 07:10	Grid service manager updated the certificates with the right (generic) CN.
2013/02/22 09:33	Ticket set to 'solved' and, later on, 'verified'. No further comments added.

# ATLAS ALARM->CERN TRANSFER

## FAILURES FROM T2 GGUS:92166

What time UTC	What happened
2013/03/06 02:56	GGUS TEAM ticket opened, automatic email notification to <b>grid-cern-prod-admins@cern.ch</b> AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. <b>Type of Problem: File Transfer.</b>
2013/03/08 16:21	After 5 comments of additional information and proof from the fts server and the dashboard that remained without any answer, the ticket is upgraded into ALARM due to the problem's persistence for >2 days. Email was sent to <a href="mailto:atlas-operator-alarm@cern.ch">atlas-operator-alarm@cern.ch</a> .
2013/03/08 16:50	Grid service expert starts investigation. Operator records in the ticket that the 'right' service has been contacted.
2013/03/11 13:59	Ticket set to 'solved' after an fts agents' restart. Reasons still not understood. <b>Re-opened on 2013/03/14 06:42hrs related to TEAM <a href="#">GGUS:92487</a></b>

# Analysis of the reliability plots :

## Week 21/01/2013

- ALICE: Nothing to report

- ATLAS:

- 2.1: NIKHEF. The following CREAM-CE tests were timing out: JobSubmit, swspace and swtag.
- 2.2: SARA. The following CREAM-CE tests were timing out: JobSubmit, swspace and swtag.

- CMS:

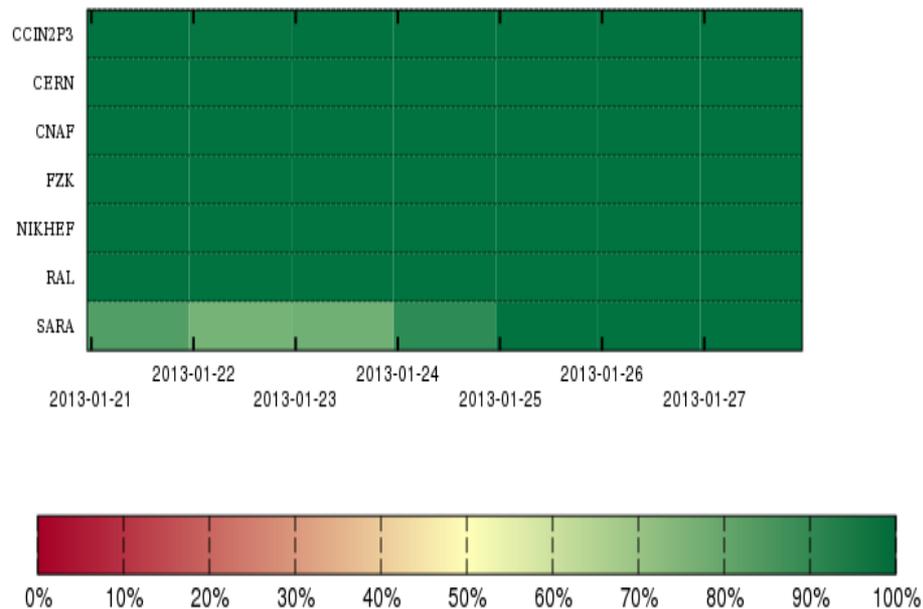
- 3.1: PIC. The WN-gLExec CREAM-CE test was timing out and the xrootd-access and xrootd-fallback tests were failing with the cmsRun exitcode: 21504.
- 3.2: TAIWAN. The WN-gLExec CREAM-CE test was timing out and the xrootd-access and xrootd-fallback tests were failing with the cmsRun exitcode: 21504.

- LHCb:

- 4.1: GRIDKA. After the change of the GRIDKA SRM endpoint also monitoring (SLS,SUM) has been updated accordingly but the new endpoint was not set as 'production' in gocdb: [GGUS:90634](#).

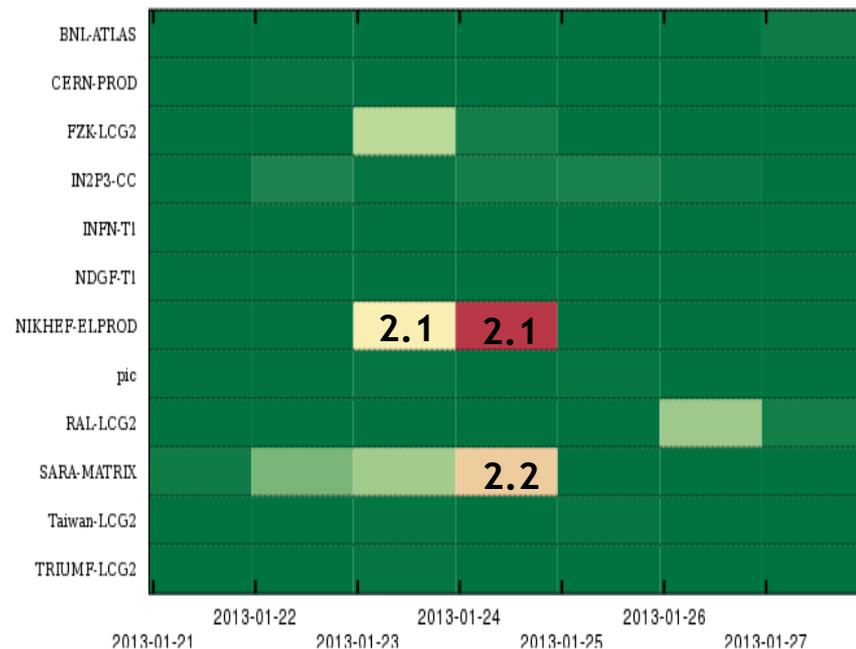
## Site reliability using ALICE\_CRITICAL

168 hours from 2013-01-21 00:00 to 2013-01-28 00:00



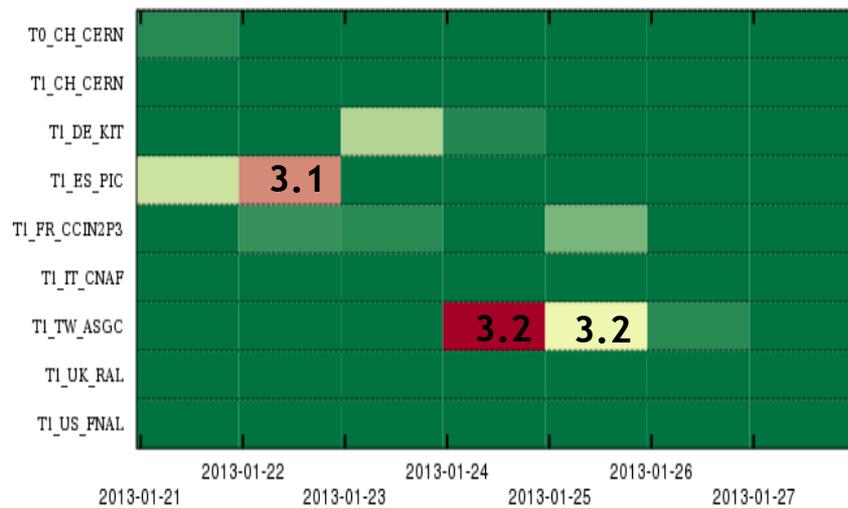
## Site reliability using ATLAS\_CRITICAL

168 hours from 2013-01-21 00:00 to 2013-01-28 00:00



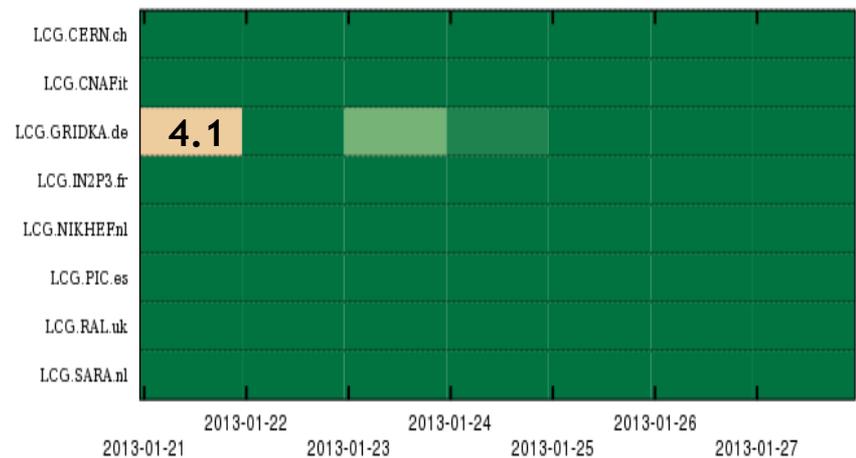
## Site reliability using CMS\_CRITICAL\_FULL

168 hours from 2013-01-21 00:00 to 2013-01-28 00:00



## Site reliability using LHCb\_CRITICAL

168 hours from 2013-01-21 00:00 to 2013-01-28 00:00



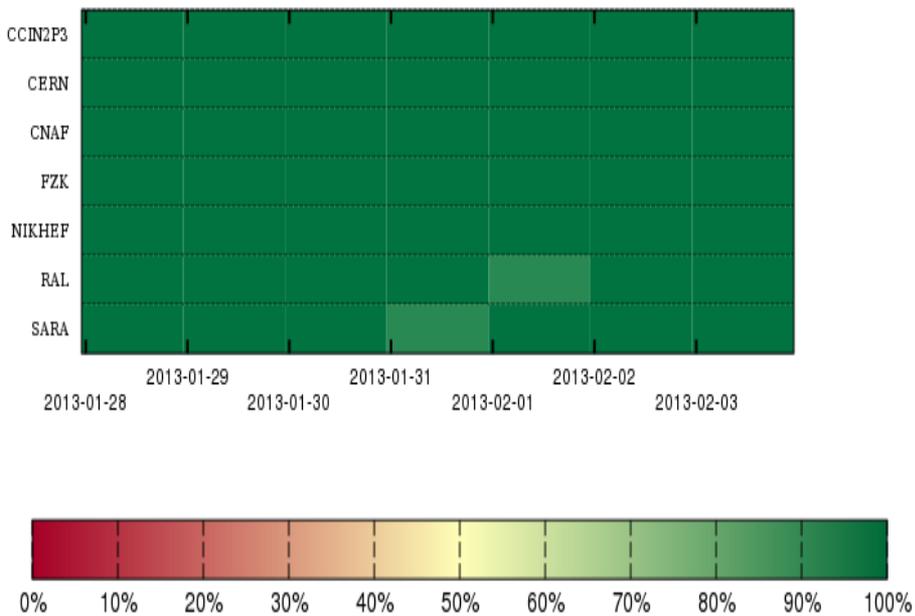
# Analysis of the reliability plots :

## Week 28/01/2013

- ALICE: Nothing to report
- ATLAS: Nothing to report
- CMS: Nothing to report
- LHCb: Nothing to report

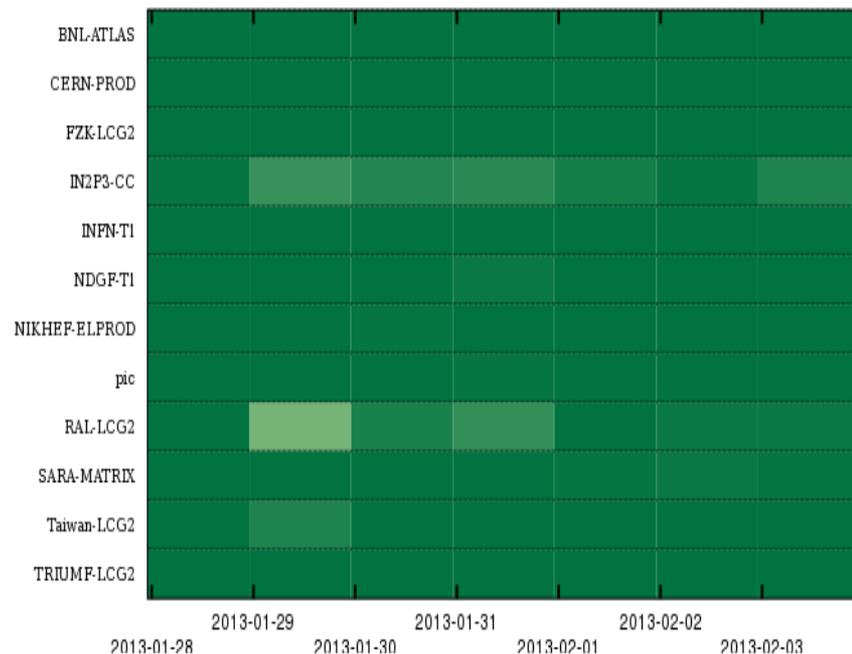
# Site reliability using ALICE\_CRITICAL

168 hours from 2013-01-28 00:00 to 2013-02-04 00:00



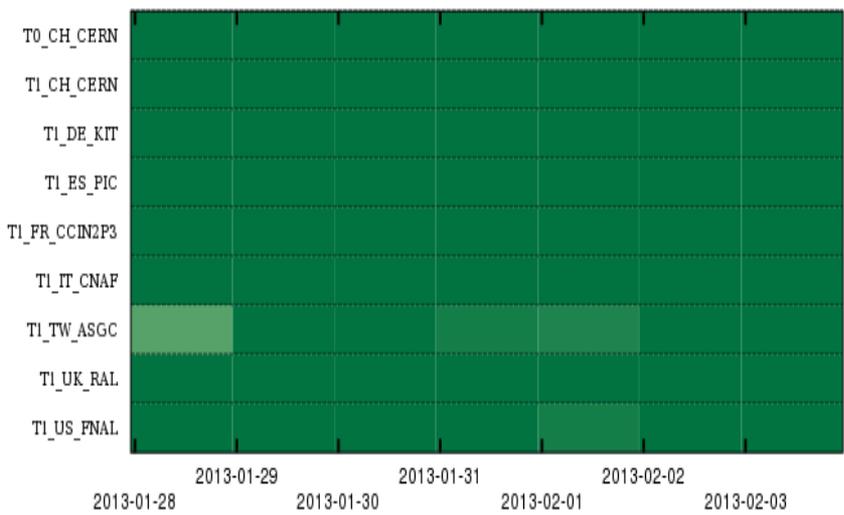
# Site reliability using ATLAS\_CRITICAL

168 hours from 2013-01-28 00:00 to 2013-02-04 00:00



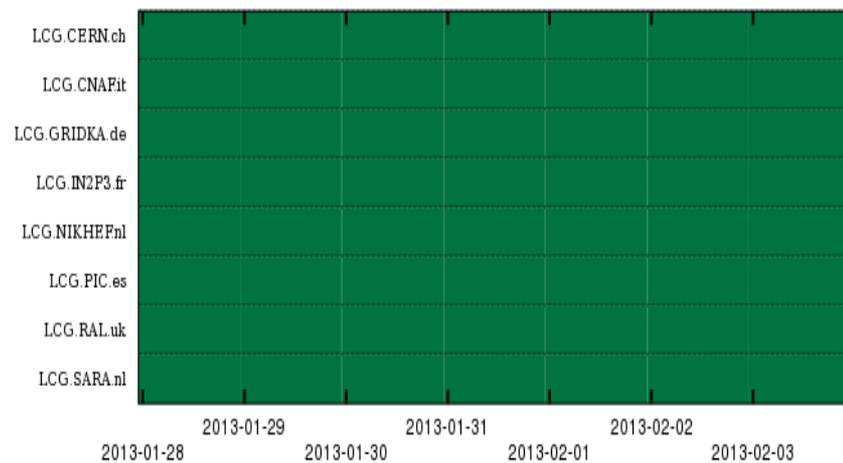
# Site reliability using CMS\_CRITICAL\_FULL

168 hours from 2013-01-28 00:00 to 2013-02-04 00:00



# Site reliability using LHCb\_CRITICAL

168 hours from 2013-01-28 00:00 to 2013-02-04 00:00



# Analysis of the reliability plots :

## Week 04/02/2013

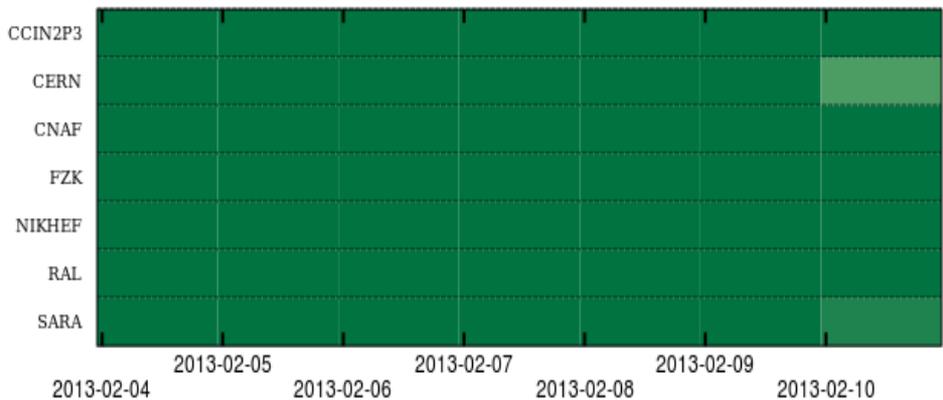
- **ALICE:** Nothing to report

- **COMMON:**

- **0.1:** A network problem at KIT causes SRM failures (see GGUs ticket 91316). ATLAS, CMS and LCHb were impacted.

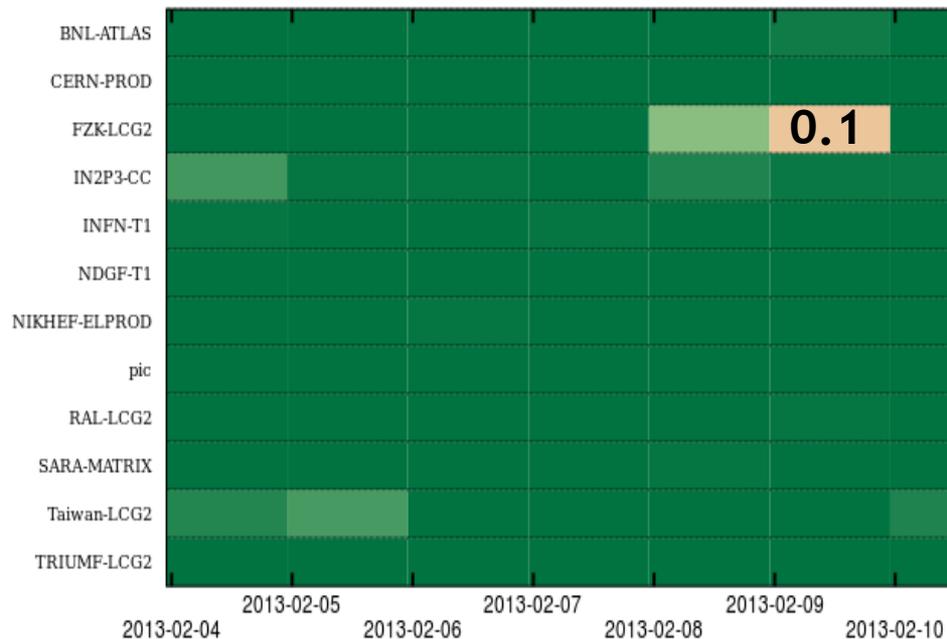
# Site reliability using ALICE\_CRITICAL

168 hours from 2013-02-04 00:00 to 2013-02-11 00:00



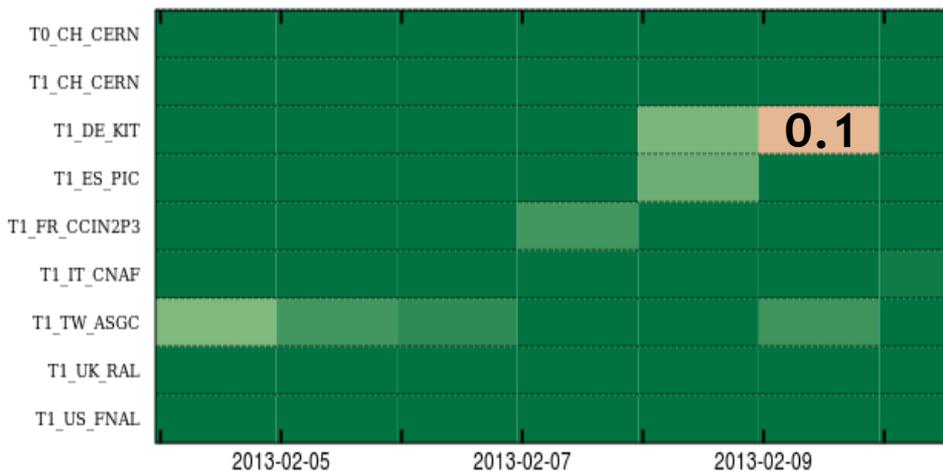
# Site reliability using ATLAS\_CRITICAL

168 hours from 2013-02-04 00:00 to 2013-02-11 00:00



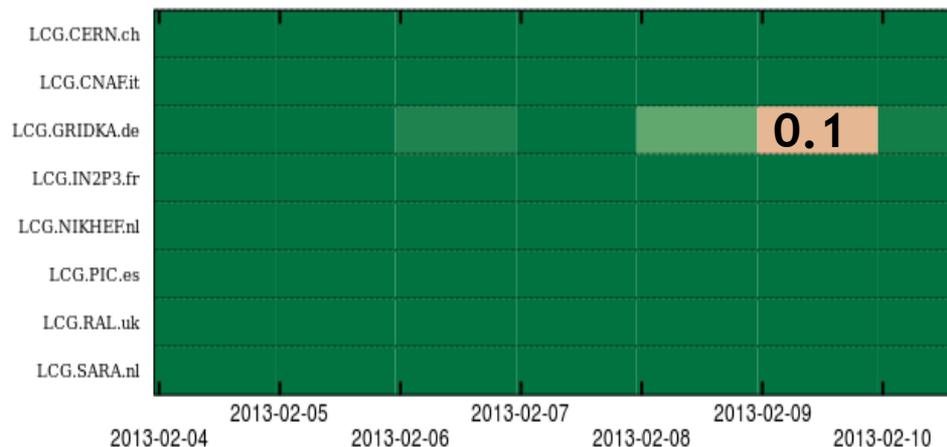
# Site reliability using CMS\_CRITICAL\_FULL

168 hours from 2013-02-04 00:00 to 2013-02-11 00:00



# Site reliability using LHCb\_CRITICAL

168 hours from 2013-02-04 00:00 to 2013-02-11 00:00



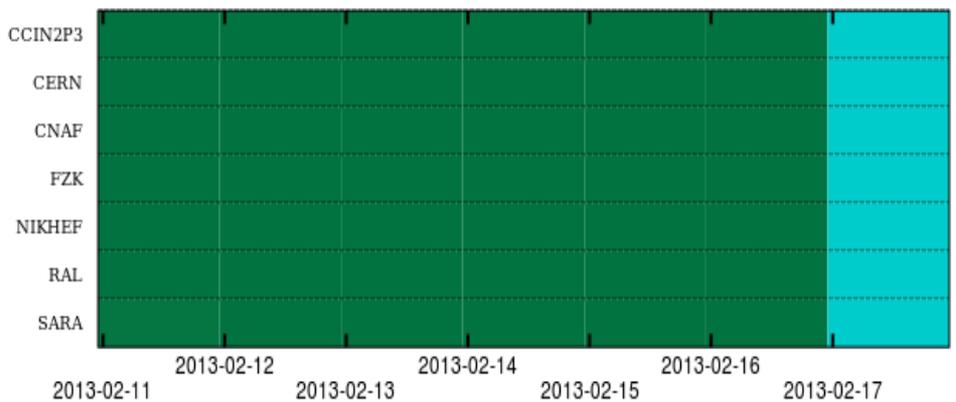
# Analysis of the reliability plots :

## Week 11/02/2013

- ATLAS:** Nothing to report
- CMS:** Nothing to report
- LHCb:** Nothing to report
  
- ALICE:** Availability is missing for February 17, The machine where the SAM tests are running has been rebooted at 11.00 am (February 18) and everything is back to normal.

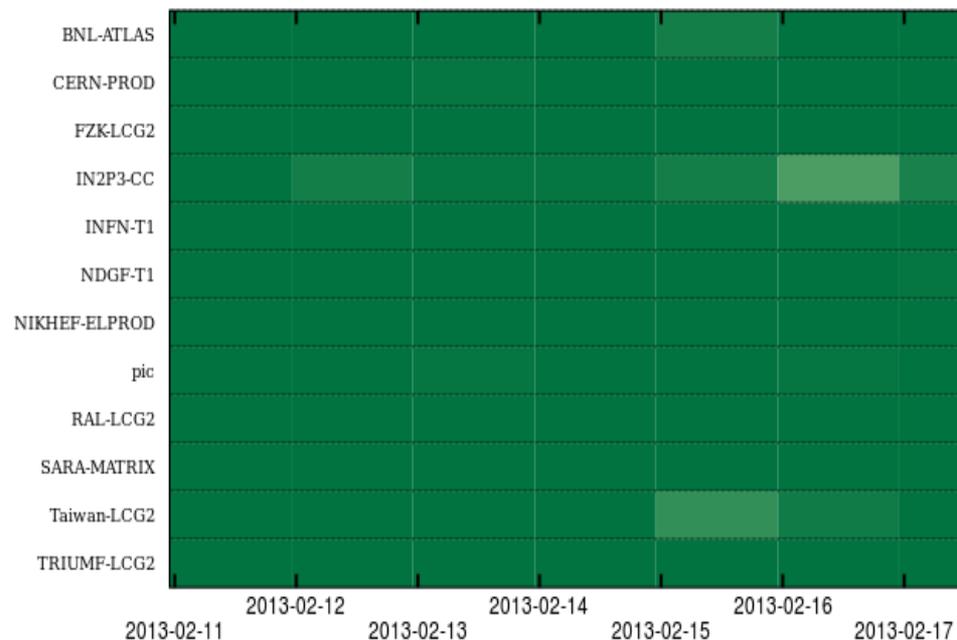
# Site reliability using ALICE\_CRITICAL

168 hours from 2013-02-11 00:00 to 2013-02-18 00:00



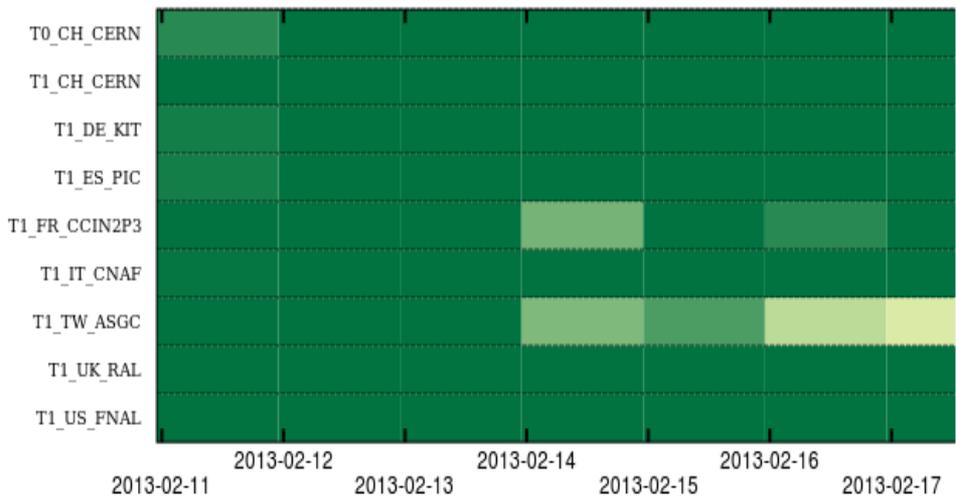
# Site reliability using ATLAS\_CRITICAL

168 hours from 2013-02-11 00:00 to 2013-02-18 00:00



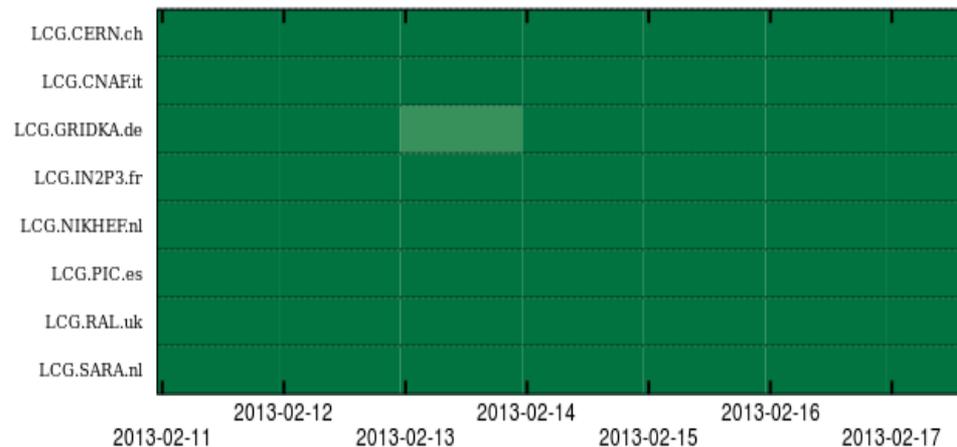
# Site reliability using CMS\_CRITICAL\_FL

168 hours from 2013-02-11 00:00 to 2013-02-18 00:00



# Site reliability using LHCb\_CRITICAL

168 hours from 2013-02-11 00:00 to 2013-02-18 00:00



# Analysis of the reliability plots :

## Week 18/02/2013

- ATLAS:** Nothing to report

- LHCb:** Nothing to report

- ALICE:**

- 1.1:** Job timed out in the CREAMCE-JobSubmit: 330 min timeout for the job exceeded. Cancelling the job.

- CMS:**

- 3.1:** CREAMCE-JobSubmit failures : Transfer to CREAM failed due to exception: Failed to create a delegation id for job

<https://wms305.cern.ch:9000/HKwVvfRoNmZVtCwlbem-2g>: reason is Received NULL fault



# Analysis of the reliability plots :

## Week 25/02/2013

### •Common:

- **0.1 (ALICE, ATLAS, LHCB):** Registered downtime at SARA : Major network maintenance and upgrade of Worker nodes to Centos6 and UMD2. Initial :25-26FEB, extended to 27 FEB.
- **0.2 (ATLAS, CMS):** Both IPLC 10G link and 2.5G link were disconnected from 04.47am to 19.45 (UTC) caused by an electric power down after a fire disaster event of the IPLC city pop.

### •ALICE:

- 1.1:** org.sam.CREAMCE-JobSubmit : CRITICAL: Job was aborted.
- 1.2:** The lack of availability informations for ALICE since 28 of February is under investigation.

### •CMS:

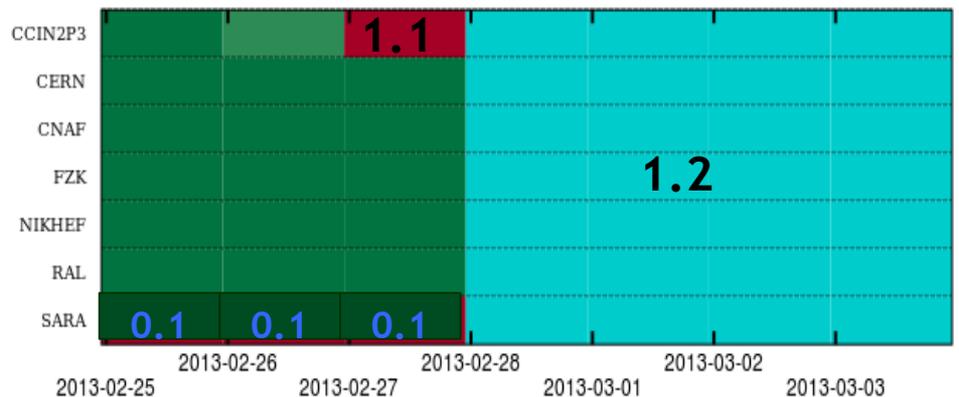
- 3.1:** org.sam.CREAMCE-JobSubmit : CRITICAL: Job was aborted. Xrootd access and fallback tests failed too: Input file root://... was not found.

### •LHCB:

- 4.1:** No data available for metric org.lhcb.WN-sft-vo-swdir.

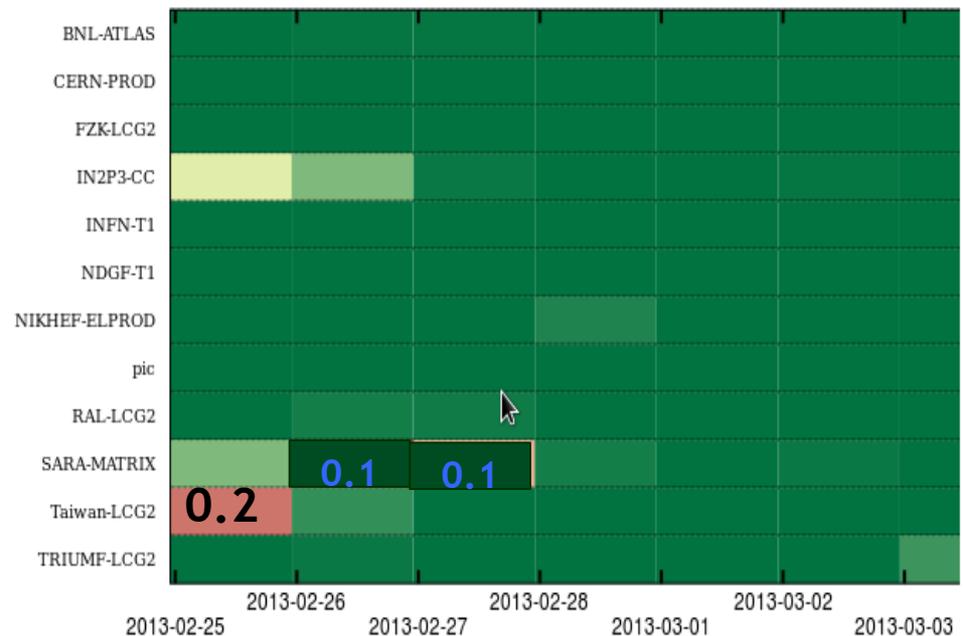
## Site reliability using ALICE\_CRITICAL

168 hours from 2013-02-25 00:00 to 2013-03-04 00:00



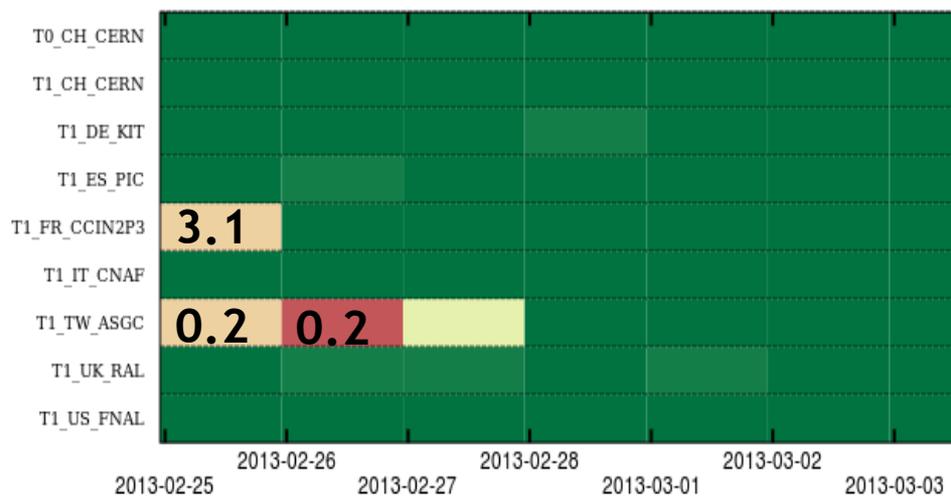
## Site reliability using ATLAS\_CRITICAL

168 hours from 2013-02-25 00:00 to 2013-03-04 00:00



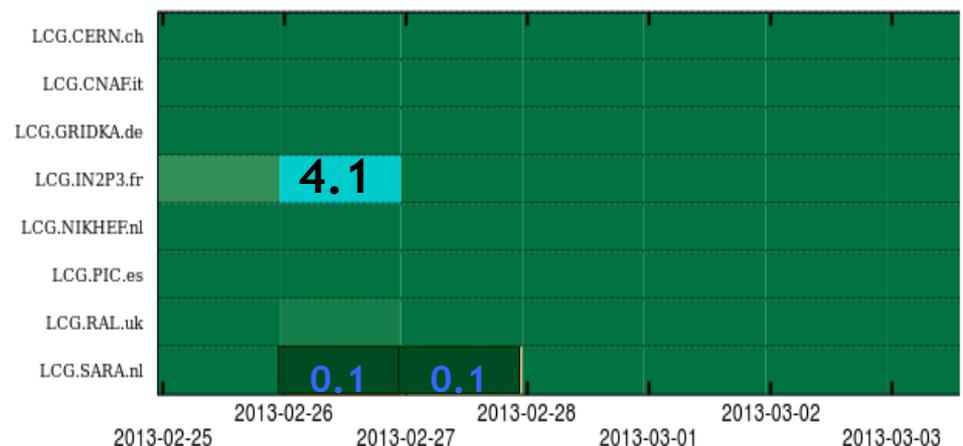
## Site reliability using CMS\_CRITICAL\_FULL

168 hours from 2013-02-25 00:00 to 2013-03-04 00:00



## Site reliability using LHCb\_CRITICAL

168 hours from 2013-02-25 00:00 to 2013-03-04 00:00

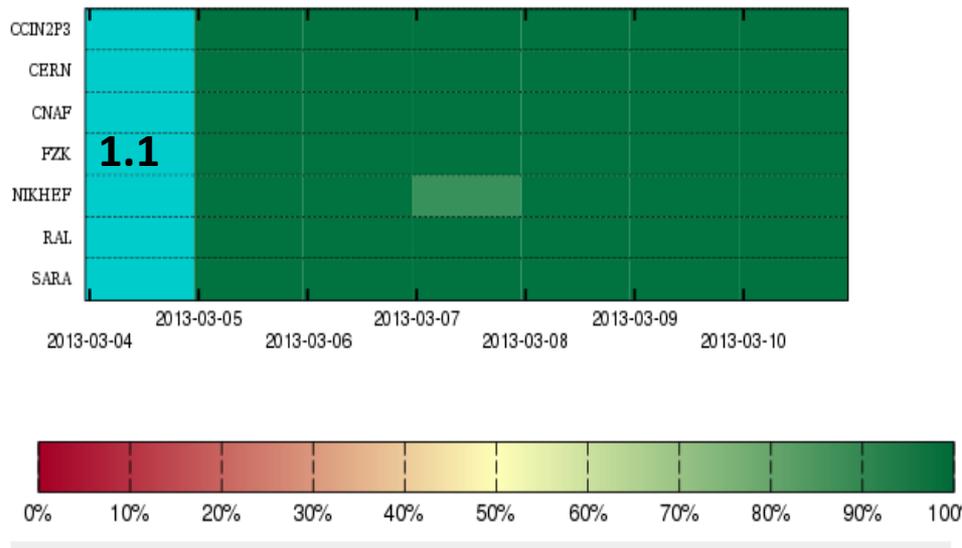


# Analysis of the reliability plots: Week 04/03/2013

- **ALICE :**
- **1.1:** No SAM test results available, because of a problem with the machine that sends them. After reinstall of the machine everything is back to normal.
- **ATLAS :** NTR
- **CMS :** NTR
- **LHCB :** NTR

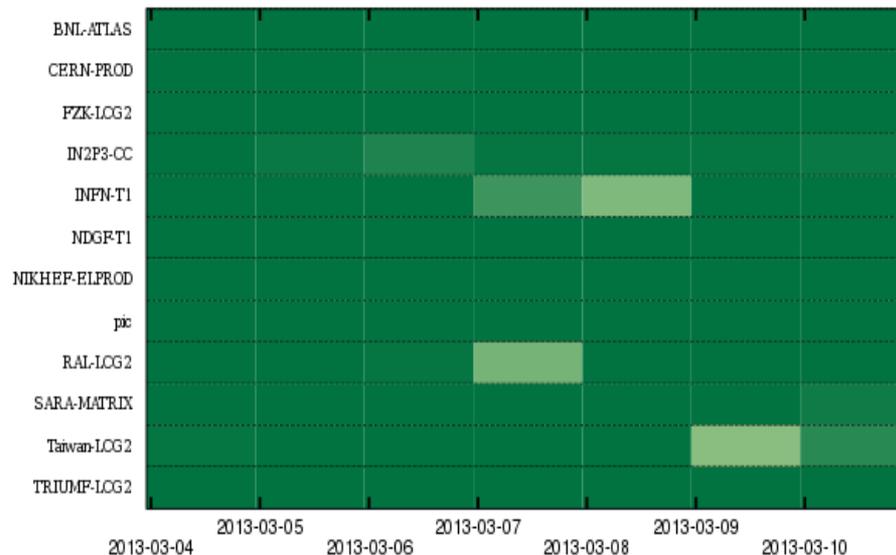
## Site reliability using ALICE\_CRITICAL

168 hours from 2013-03-04 00:00 to 2013-03-11 00:00



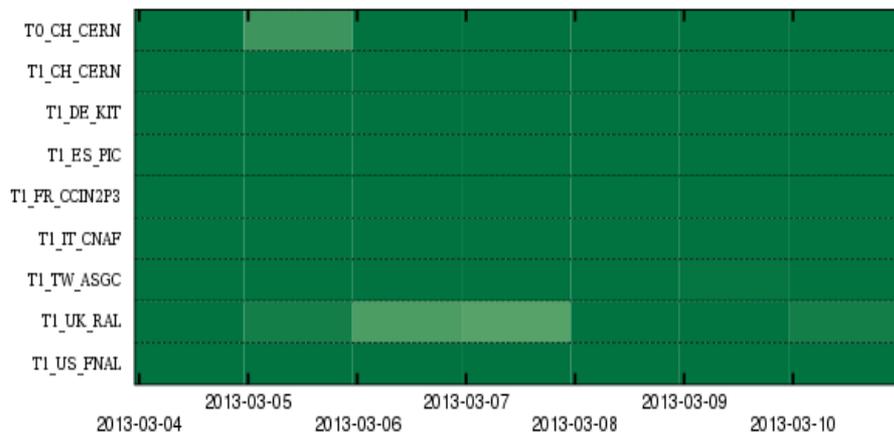
## Site reliability using ATLAS\_CRITICAL

168 hours from 2013-03-04 00:00 to 2013-03-11 00:00



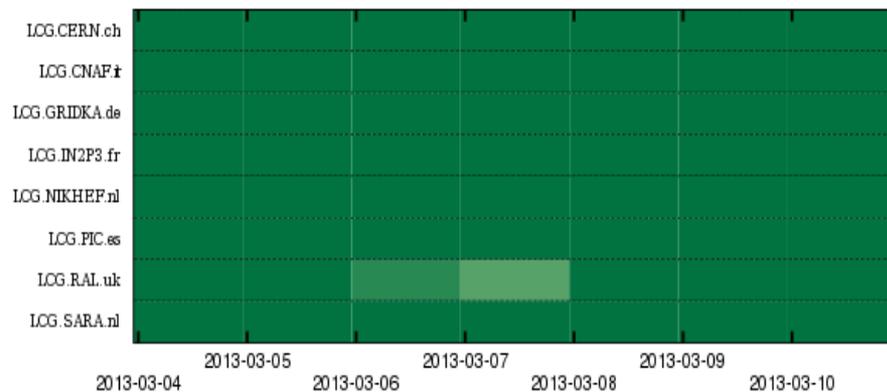
## Site reliability using CMS\_CRITICAL\_FULL

168 hours from 2013-03-04 00:00 to 2013-03-11 00:00



## Site reliability using LHCb\_CRITICAL

168 hours from 2013-03-04 00:00 to 2013-03-11 00:00



# Analysis of the reliability plots :

## Week 11/03/2013

- Common:

- 0.1: Schedule downtime 12/03/2013 08:45 - 12/03/2013 15:30 , extended until 12/03/2013 21:00.
- Reconfiguration of core network within the RAL Tier1. Storage (Castor) services will be stopped. LFC stopped. FTS and Batch drained of active transfers/jobs. Other services (e.g. BDII) may see some short breaks in connectivity.

- ALICE: Nothing to report

- ATLAS: Nothing to report

- CMS: Nothing to report

- LHCb: Nothing to report



# Conclusions

- 8 weeks of activity since the last MB on Jan 22<sup>nd</sup>.
- The LHC shutdown took place in about the middle of this period Feb 14<sup>th</sup>.
- The daily meeting now takes place on Mondays and Thursdays. Almost no change in attendance. Items solve themselves via GGUS tickets' follow-up so the volume of reporting, hence the meeting duration remains the same (even drops).
- Having said that, if the meeting were still daily would [GGUS:92166](#) remain still open?