WLCG Service Report
3 weeks Nov 7 – Nov 28

Maarten Litmaath

~~~

WLCG Management Board, 29th November 2011

Executive summary

- BNL switch to dCache Chimera etc. OK
- Smooth switch to HI data taking
  - p-Pb tests OK, but run was prevented by upstream problems → next year

- Various T2 connectivity issues affecting IN2P3-CC and KIT
  - CMS provided links to a CMS monitoring tool to assist in debugging
  - ATLAS affected as well, asked ESnet to look into the matter
  - IN2P3-CC temporarily suppressed LBE flag on outgoing packets

- KIT LFC for ATLAS merged into CERN LFC
  - KIT upgraded ATLAS dCache to new golden release in parallel

- LHCb reprocessed their full 2011 data set in just 7 weeks
- IN2P3-CC completed their batch system migration BQS → SGE

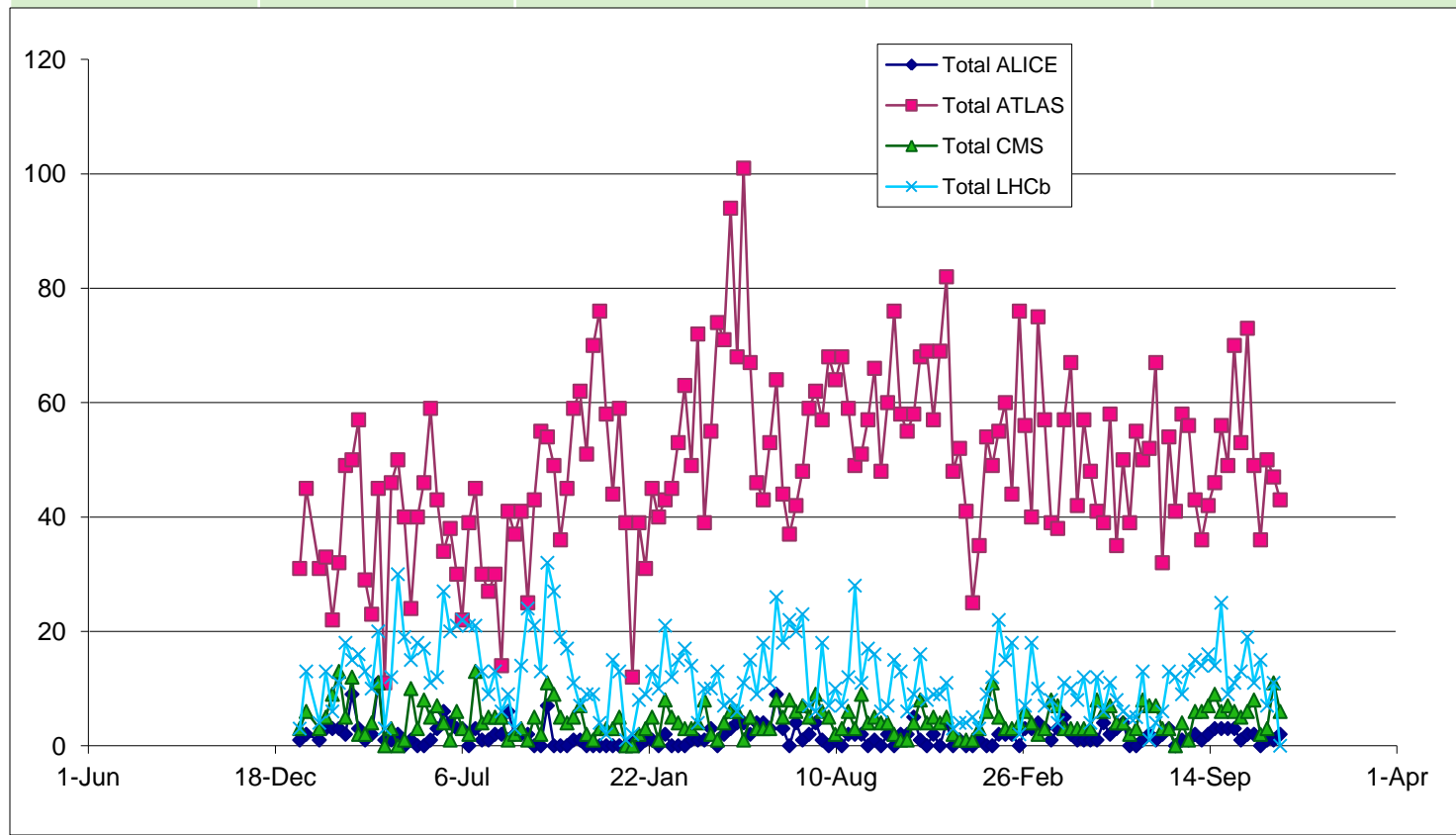- EOS increasingly important for ATLAS and CMS

- And business as usual…

SIR on GGUS peer system updates

- Description
  - Failure of the wrapper tool for interfacing with peer ticketing systems from Nov 4, 16:15 to Nov 7, 11:00 CET.

- Impact
  - Peer tickets in SNOW (CERN), NGI_France and ROC_Russia could neither be created nor updated.

- Analysis
  - An unrecognized issue with the KIT mail server led to a restart of the Remedy system to try and cure the problem. The Remedy server ended up with an incorrect environment due to the restart being done from the wrong account (should have been root). This led to various failures in scripts used by GGUS, including the wrapper tool.
  - The startup script of the Remedy server has been made more robust to avoid such problems.

GGUS summary *(3 weeks)*   by Maria Dimou

VO	User	Team	Alarm	Total
ALICE	4	0	0	4
ATLAS	17	118	5	140
CMS	11	8	1	20
LHCb	1	16	1	18
Totals	33	142	7	182

4

Support-related events since last MB

- There were 7 real ALARM tickets since the 2011/11/08 MB (3 weeks), 5 by ATLAS, 1 by CMS, 1 by LHCb.

- 4 ALARM tickets concerned CERN, 2 were for SARA and 1 for CNAF.

- All of them are in status 'solved', most are also 'verified'.

  Details follow…

ATLAS ALARM->CERN slow LSF response
GGUS:76039

What time UTC	What happened
2011/11/07 08:01	GGUS ALARM ticket, automatic email notification to atlas-operator-alarm@cern.ch AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation took place 2hrs 28 mins later! This is the 1st time we see this. Maybe SNOW was in the middle of its weekly release (every Monday) and didn't accept connections? Type of Problem = **ToP: Local Batch System**.
2011/11/07 08:02	The operator records in the ticket that it-dep-pes-ps-sms was informed.
2011/11/07 08:21	Service expert records a hardware problem with the LSF master node. The service ran on a secondary node for the rest of the day with slow performance. 5 comments exchanged with the submitter along these lines.
2011/11/08 07:18	Service expert sets the ticket to 'solved' as hardware issues were addressed and a reconfiguration solved the slow performance. A ticket was opened to Platform to investigate the root cause of the problem.
2011/11/08 07:27	Submitter sets the ticket to 'verified'. Thus the root cause will never be recorded…!!!???

CMS ALARM-> CERN LSF problem
GGUS:76045

What time UTC	What happened
2011/11/07 08:54	GGUS TEAM ticket, automatic email notification to grid-cern-prod-admins@cern.ch AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. **ToP: Local Batch System.**
2011/11/07 10:31	CERN SNOW 2nd Line Supporter assigns the ticket to LXBATCH 3rd Line support group.
2011/11/07 15:58	Expert sets the ticket in status 'in progress'. It is interesting that the same expert had been much more responsive to the similar ticket from ATLAS (previous slide) which was opened as an ALARM and not a TEAM ticket, although, at this point in time, a solution was not yet found.
2011/11/07 16:48	As the situation didn't improve, the ticket was upgraded to ALARM. Email was sent to cms-operator-alarm@cern.ch.
2011/11/07 16:59	Operator records in the ticket that it-dep-pes-ps-sms@cern.ch received a copy of the notification. The ticket was solved the next day at the same time as the ATLAS one.

ATLAS ALARM->CERN slow AFS
GGUS:76519

What time UTC	What happened
2011/11/16 16:29	GGUS ALARM ticket, automatic email notification to atlas-operator-alarm@cern.ch AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. **ToP: File Access.**
2011/11/16 16:35	Operator records in the ticket that AFS support was contacted.
2011/11/16 17:24	5 comments exchanged between the submitter and the service managers led to setting the ticket to 'solved' because performance gradually improved without needing to move data to another server.
2011/11/16 21:51	Submitter sets the ticket to 'verified'.

LHCb ALARM-> SARA SE down
GGUS:76629

What time UTC	What happened
2011/11/20 08:34 SUNDAY	GGUS TEAM ticket, automatic email notification to eugrid.support@sara.nl AND automatic assignment to NGI_NL **ToP: Storage Systems.**
2011/11/20 10:56	Ticket upgrade to ALARM. Email sent to nlt1-alarms@biggrid.nl.
2011/11/20 11:03	Expert on call recorded in the ticket that the problem is known, ATLAS reported the same and being investigated.
2011/11/20 11:10	Submitter bans SARA to prevent jobs from being submitted there while problem not yet solved.
2011/11/20 20:36	SARA service mgr traces the problem down to a full partition of the dCache namespace node. Cleared up and restarted dCache cluster. Ticket set to 'solved' at 21:32hrs. During the daily WLCG ops meeting the next day the site reminded that during weekends they are outside SLA hours and they respond to ALARMs on best effort.
2011/11/21 09:10	An acknowledgement of ALARM reception sent only now from a NIKHEF address (the content seems to be a standard text).

ATLAS ALARM-> unable to contact SARA SRM
GGUS:76628

What time UTC	What happened
2011/11/20 06:17 SUNDAY	GGUS TEAM ticket, automatic email notification to eugrid.support@sara.nl AND automatic assignment to NGI_NL **ToP: File Transfer.**
2011/11/20 08:03	Ticket upgrade to ALARM after a reminder that data exports to SARA are failing during data taking. Email sent to nlt1-alarms@biggrid.nl.
2011/11/20 08:24	Expert on call recorded in the ticket the standard text of ALARM receipt acknowledgement.
2011/11/20 11:07	Submitter takes SARA out of exports while the problem is not yet solved.
2011/11/20 20:32	SARA service mgr traces the problem down to a full partition of the dCache namespace node. Cleared up and restarted dCache cluster. Ticket set to 'solved' at 21:33hrs. Logging level was planned to be increased for debugging during the week but as the submitter 'verified' the ticket on 2011/11/21 00:44hrs no more updates are possible.

ATLAS ALARM-> data exports to CNAF fail
GGUS:76663

What time UTC	What happened
2011/11/21 13:17	GGUS TEAM ticket, automatic email notification to t1-admin@lists.cnaf.infn.it AND automatic assignment to NGI_IT. **ToP: File Transfer.**
2011/11/21 13:54	Ticket upgrade to ALARM. Email sent to t1-alarms@cnaf.infn.it
2011/11/21 14:09	Site mgr comments that un unscheduled downtime of 2011/11/18 (Friday) had left these instabilities. Things started getting better.
2011/11/21 15:29	Site mgr sets the ticket in status 'solved' after confirmation from the ALARMer that the errors went down.
2011/11/21 19:04	Another shifter re-opens the ticket finding the batch queue closed and wondering if the problem persists or the queue re-opening was simply forgotten.
2011/11/22 12:55	Site mgr re-sets the ticket to 'solved' pasting the last shifters question about the batch queue found closed without any comment (?!)

ATLAS ALARM->CERN LFC down
GGUS:76770

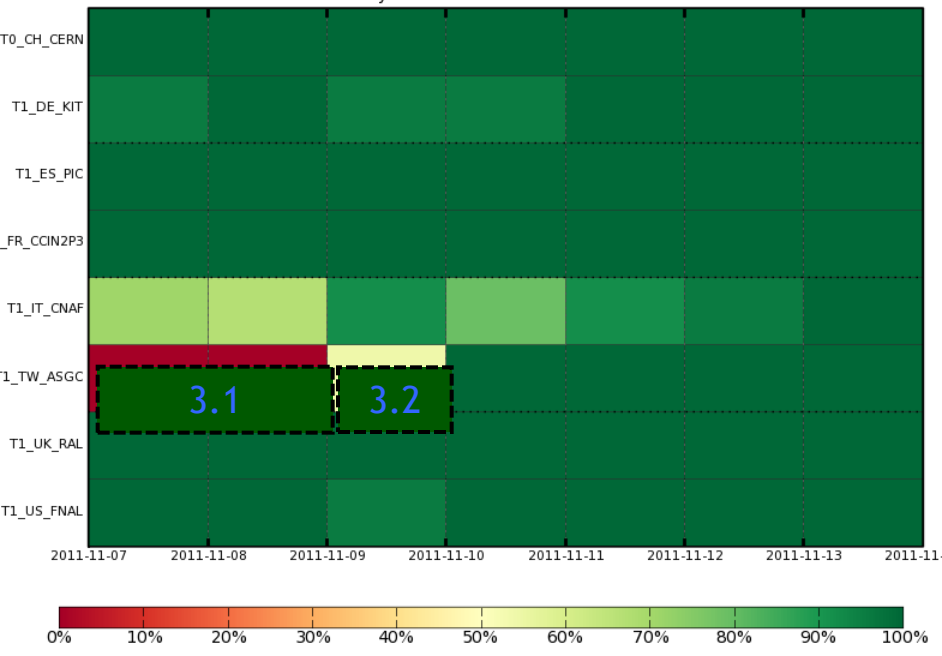| What time UTC	What happened
2011/11/24 04:37	GGUS TEAM ticket, automatic email notification to grid-cern-prod-admins@cern.c h AND automatic assignment to ROC_CERN. Automatic SNOW ticket creation successful. **ToP: Databases.**
2011/11/24 04:51	Operator records in the ticket that it-dep-pes-ps-sms@cern.ch was contacted. This is strange because the operators are not included in the e-group notified (!?)
2011/11/24 05:52	Authorised ALARMer upgrades the ticket and offers a possible incident reason related to ADCR db. Email sent to atlas-operator-alarm@cern.ch Maybe the timestamp in the ticket is incorrect…
2011/11/24 07:31	CERN 2nd Line supporter assigns the ticket to "LFC 2nd Line Support". This is not the right level given that it is now an ALARM.
2011/11/24 09:37	Grid services' expert started working on the problem immediately. After a few exchanges this was 'solved' as related to adcr_lfc Oracle DB not available.

ALICE
7 Days from 2011-11-07 to 2011-11-14
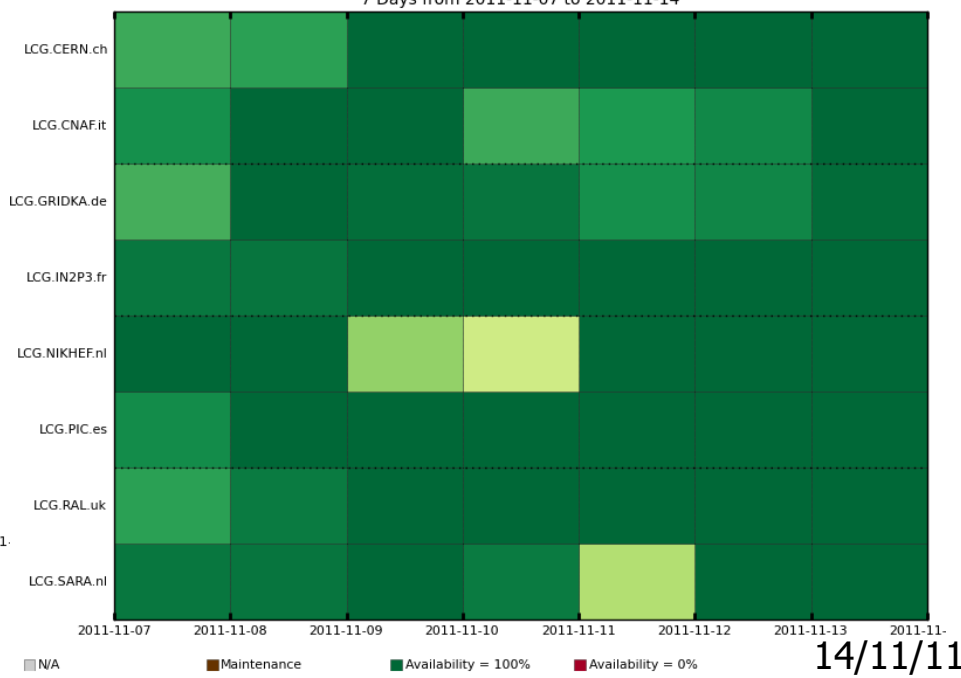
ATLAS
7 Days from 2011-11-07 to 2011-11-14

CMS
7 Days from 2011-11-07 to 2011-11-14

LHCb
7 Days from 2011-11-07 to 2011-11-14

14/11/11

Analysis of the availability plots: Week of 07/11/2011 – 13/11/2011   by David Tuckett

ATLAS

2.1 BNL-ATLAS: Scheduled downtime 07/11-09/11 - Three days of facility maintenance for all services.

2.2 BNL-ATLAS: Unscheduled downtime 09/11 20:00 - 10/11 15:00 - dCache downtime extended.

2.3 INFN-T1: CE/CREAMCE JobSubmit tests failing from 08/11 02:00 to 10/11 01:00 and from 10/11 19:00 to 12/11 18:00 – "330 min timeout for the job exceeded. Cancelling the job." Understood to be a network issue; the network groups from CNAF and CERN have maybe found and fixed the problem.

2.4 RAL-LCG2: Known monitoring issue with current SAM visualization. The upcoming SUM visualization resolves this issue.

CMS

3.1 T1_TW_ASGC: Scheduled downtime 07/11-08/11 - Two days maintenance for Castor upgrade.

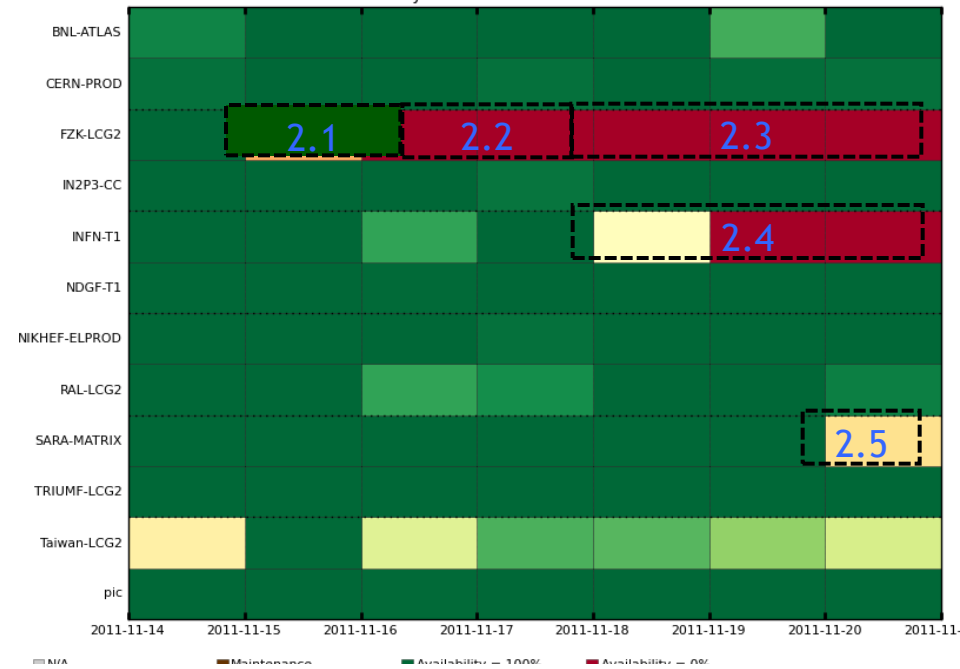**3.2 T1_TW_ASGC:** Monitoring issue. Dashboard SAM Visualization developers are investigating.
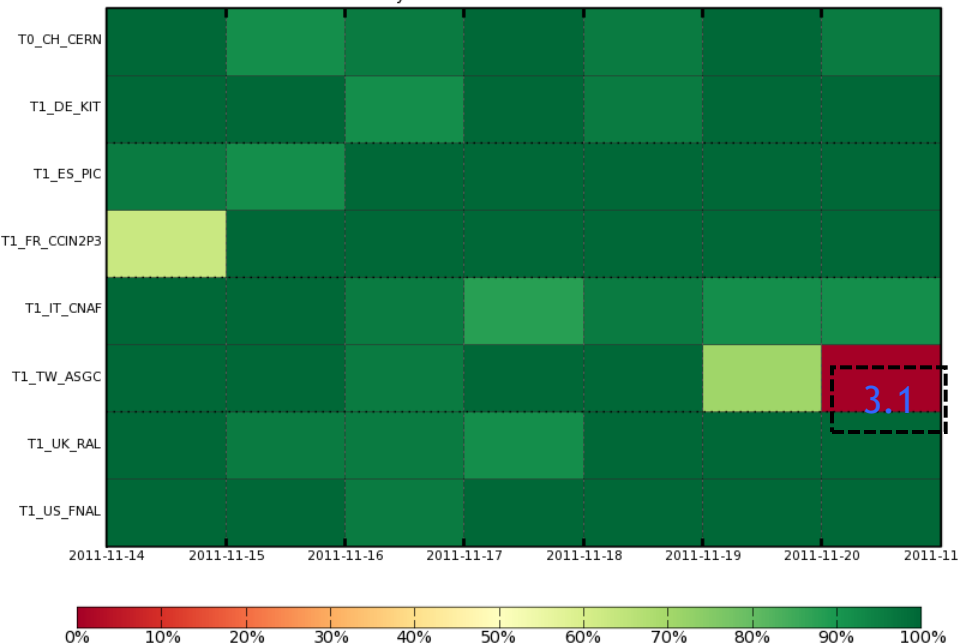
14/11/11

ALICE
7 Days from 2011-11-14 to 2011-11-21
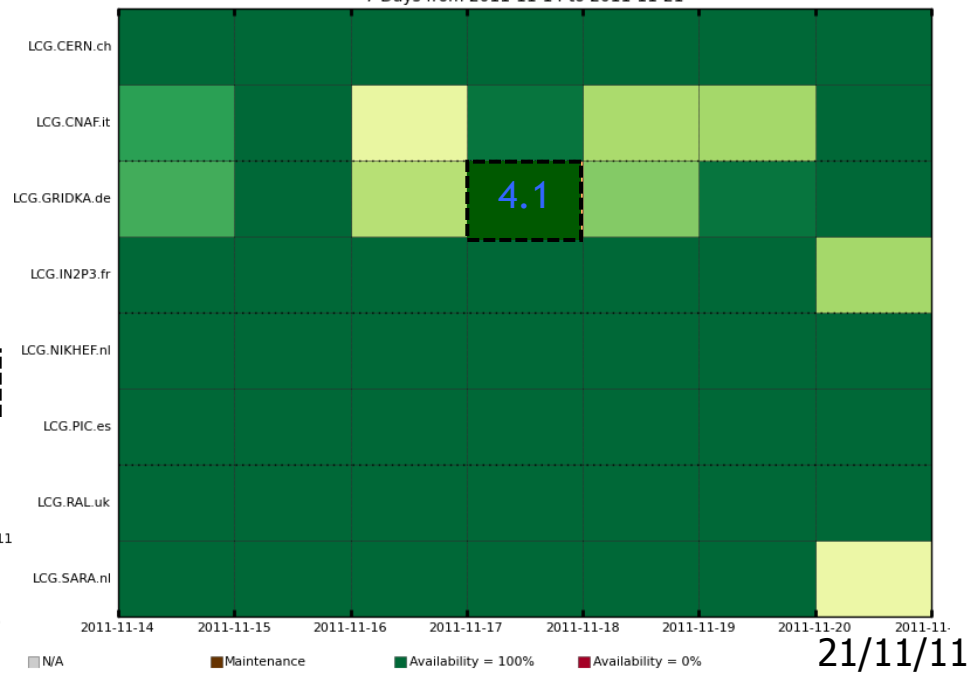
ATLAS
7 Days from 2011-11-14 to 2011-11-21

CMS
7 Days from 2011-11-14 to 2011-11-21

LHCb
7 Days from 2011-11-14 to 2011-11-21

N/A   Maintenance   Availability = 100%   Availability = 0%

21/11/11

Analysis of the availability plots: Week of 14/11/2011 – 20/11/2011   by David Tuckett

ATLAS

2.1 FZK-LCG2: Scheduled outage from 15/11 07:00 to 16/11 12:00.

2.2 FZK-LCG2: Unscheduled outage from 16/11 12:00 to 17/11 17:00.

2.3 FZK-LCG2: LFC/LFC_L Ping and Read tests failing since outage until 18/11 11:00 - "Name server not active" - then not running. GGUS:76605

2.4 INFN-T1: CE/CREAMCE JobSubmit tests failing from 18/11 16:00 - "45 min timeout for status [Waiting] exceeded. Cancelling the job." Possibly related to GGUS:76597

2.5 SARA-MATRIX: SRMv2 VOLsDir and VOPut tests failing from 20/11 04:00 to 20/11 20:00 - "User timeout over." GGUS:76628.
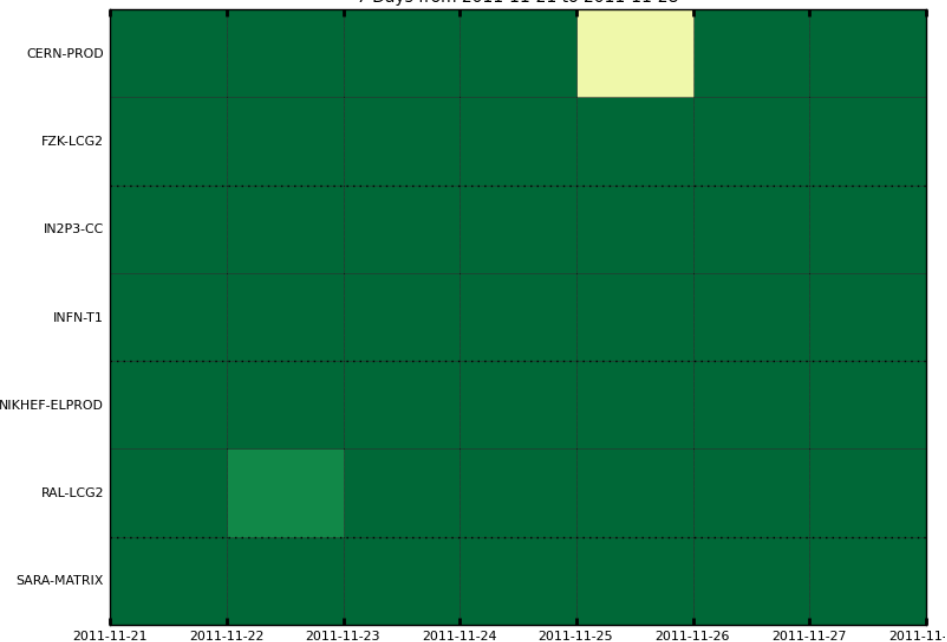
CMS

3.1 T1_TW_ASGC: SRMv2 VOPut tests failing from 19/11 17:00 to 21/11 08:30 - "zero number of replicas." GGUS:75377,76648.
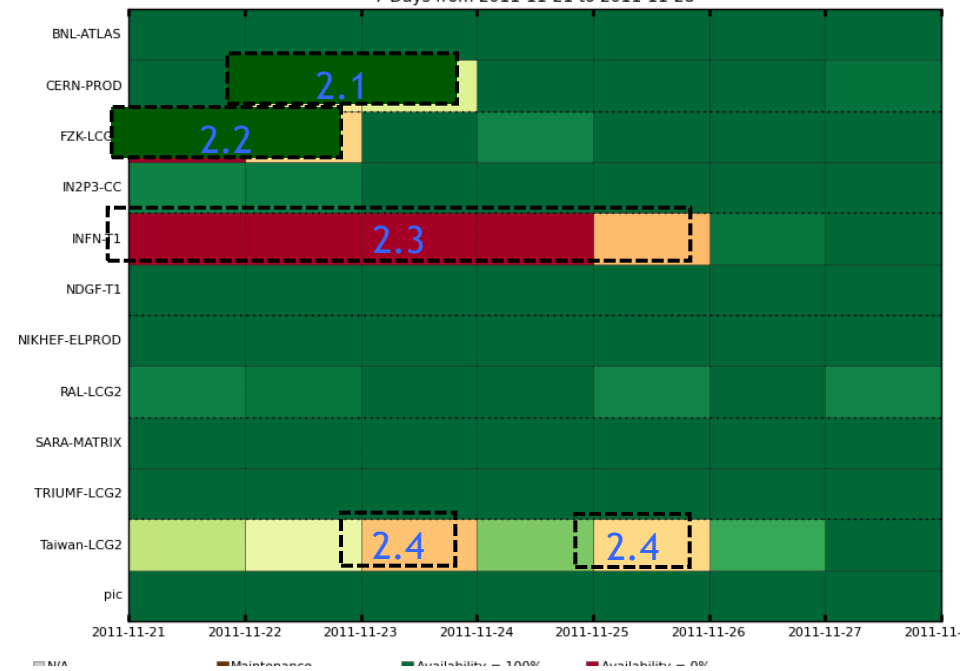
LHCb

4.1 LCG.GRIDKA.de: Known monitoring issue. Resolved in upcoming SUM visualization.

28/11/11

Analysis of the availability plots: Week of 21/11/2011 – 27/11/2011   by David Tuckett

ATLAS

2.1 CERN-PROD: An SRM endpoint had been decommissioned but tests were still running against it. The tests have been updated.

2.2 FZK-LCG2: Following migration of LFC to CERN, LFC tests for FZK should no longer be considered in availability calculation. The availability criteria has been updated 22/11 and the reported availability returned to normal 23/11.

2.3 INFN-T1: CE/CREAMCE JobSubmit tests failing from 18/11 16:00 to 25/11 17:00 - "45 min timeout for status [Waiting] exceeded. Cancelling the job." Cause unknown. Under investigation.

2.4 Taiwan-LCG2: CE/CREAMCE JobSubmit tests failing intermittently 23/11 and 25/11. CEs had problem trusting BNL VOMS server but no problem trusting CERN VOMS server causing unstable SAM test results. Fixed 25/11.