



# ATLAS Suspension and Downtime Procedures

Graeme Stewart  
(for ATLAS Central Operations Team)



# Production



- For ATLAS production we run shift teams who look at the state of sites
  - There's almost 24 hour coverage, but not 24 hour expert support

Cloud Information	Nodes	Jobs	Latest	Pilots (3hrs)	defined	assigned	waiting	activated	sent	running	holding	transferring	finished	failed tot trf other
Overall Production	4783	15536	03-17 19:05	35447	0 / 0	3320 / 0	26 / 0	11819 / 0	0 / 0	14945 / 0	689 / 0	3614 / 396	12759 / 0	8721 / 953 41% 1% 35%
CA ^	346	1060	03-17 19:05	2780	0	145	0	571	0	638	180	492 / 81	1248	540 30% 0% 30%
CERN ^	9	2	03-17 18:54	193	0	0	0	0	0	2	0	0 / 0	10	1 9% 0% 9%
DE ^	760	374	03-17 19:05	1552	0	139	0	69	0	1875	230	36 / 0	1354	187 12% 0% 12%
ES ^	24	344	03-17 19:01	2640	0	0	0	337	0	1	0	0 / 0	0	172 100% 0% 100%
FR ^	398	770	03-17 19:05	138	0	13	0	2338	0	1171	23	797 / 17	1281	385 23% 0% 23%
IT ^	595	934	03-17 19:05	3153	0	0	0	4423	0	2345	3	132 / 0	528	467 47% 0% 47%
IND ^	462	110	03-17 19:05	2676	0	255	0	2003	0	5385	5	0 / 0	1345	55 4% 0% 4%
INDG ^	0	0	0	0	0	0	0	0	0	0	0	0 / 0	0	0
IL ^	539	7308	03-17 19:05	3891	0	866	0	1941	0	1232	34	782 / 27	2899	3654 56% 0% 56%
TW ^	0	0	0	0	0	0	0	0	0	0	0	0 / 0	0	0
UK ^	682	806	03-17 19:05	1132	0	53	0	23	0	932	106	129 / 12	1802	403 20% 0% 20%
US ^	068	3808	03-17 19:05	17091	0	1840	0	114	0	1366	108	1336 / 250	2402	1004 43% 4% 30%

Number 1 problem site

Number 2 problem site

CA Sites	Nodes	Jobs	Latest	Pilots (3hrs)	defined	assigned	waiting	activated	sent	running	holding	transferring	finished	failed tot trf other
Site Name	341	1006	03-19 13:40	2780	0	65	0	1526	0	819	179	943 / 66	2422	503 17% 0% 17%
ALBERTA-LCG2 ^	24	3	03-19 13:40	110	0	0	0	82	0	49	20	71 / 0	29	3 9% 0% 9%
MCGILL-LCG2 ^	0	0	offline	0	0	0	0	0	0	0	0	0 / 0	0	0
SFU-LCG2 ^	32	81	03-19 13:40	576	0	0	0	460	0	233	6	495 / 66	271	9 23% 0% 23%
TORONTO-LCG2 ^	0	0		112	0	0	0	0	0	0	0	0 / 0	0	0
TRIUMF ^	165	104	03-19 13:40	1574	0	65	0	754	0	337	150	0 / 0	1861	104 5% 0% 5%
VICTORIA-LCG2 ^	120	315	03-19 13:39	408	0	0	0	230	0	200	3	387 / 0	261	31 55% 0% 55%



# Offline

- If the problems are diagnosed as site issues then
  - An ATLAS eLog entry is created
  - A GGUS ticket is sent to the site
    - This will be a directly routed ticket
  - The site is set offline if the problems are serious enough

CA Sites	Hodes	Jobs	Latest	Pilots (3hrs)	defined	assigned	waiting	activated	sent	running	holding	transferring	finished	failed tot	trf	other	
Site Name	341	1006	03-19 13:40	<u>2780</u>	<u>0</u>	<u>65</u>	<u>0</u>	<u>1526</u>	<u>0</u>	<u>819</u>	<u>179</u>	<u>943 / 66</u>	<u>2422</u>	<u>503</u>	<u>17%</u>	<u>0%</u>	<u>17%</u>
<a href="#">ALBERTA-LCG2</a>	24	3	03-19 13:40	<u>110</u>	0	0	0	<u>82</u>	0	<u>49</u>	<u>20</u>	<u>71 / 0</u>	<u>29</u>	<u>3</u>	<u>9%</u>	<u>0%</u>	<u>9%</u>
<a href="#">MCGILL-LCG2</a>	0	0	offline	<u>0</u>	0	0	0	0	0	0	0	0 / 0	0	0			
<a href="#">SFU-LCG2</a>	32	81	03-19 13:40	<u>576</u>	0	0	0	<u>460</u>	0	<u>233</u>	<u>6</u>	<u>485 / 66</u>	<u>271</u>	<u>81</u>	<u>23%</u>	<u>0%</u>	<u>23%</u>
<a href="#">TORONTO-LCG2</a>	0	0		<u>112</u>	0	0	0	0	0	0	0	0 / 0	0	0			
<a href="#">TRIUMF</a>	165	104	03-19 13:40	<u>1574</u>	0	<u>65</u>	0	<u>754</u>	0	<u>337</u>	<u>150</u>	0 / 0	<u>1861</u>	<u>104</u>	<u>5%</u>	<u>0%</u>	<u>5%</u>
<a href="#">VICTORIA-LCG2</a>	120	315	03-19 13:39	<u>408</u>	0	0	0	<u>230</u>	0	<u>200</u>	<u>3</u>	<u>387 / 0</u>	<u>261</u>	<u>315</u>	<u>55%</u>	<u>0%</u>	<u>55%</u>

- If the problem is well understood and resolved quickly by the site then the site will usually be set directly online again



# Suspension and Revalidation



- If the problem is extended, not well explained or the site have done, e.g., a major upgrade, then their queues can be put into test status while they are revalidated
  - Test status queues can pull only test jobs from PanDA!
  - Test jobs are usually small event generations, but do do a full chain test:

The screenshot shows the PanDA monitor interface. At the top, there are navigation links: [Production](#), [Clouds](#), [DDM](#), [PandaMover](#), [AutoPilot](#), [Sites](#), [Analysis](#), [Physics data](#), [Usage](#), [Plots](#), [ProdDash](#), [DDMDash](#). A user is logged in as [Not logged in](#) with a [List users](#) link.

The main heading is "Panda job information for test jobs". Below this, it shows "type:test Site:UKI-SCOTGRID-GLASGOW" and a [Click for help](#) link.

There is a search bar with "Summary of test" selected, "jobs for the last 1 days in any state at UKI-SCOTGRID-GLASGOW site" and buttons for "Go" and "Retrieve All".

Summary statistics: 7 jobs. Click job number to see details. States: defined:0 assigned:0 waiting:0 activated:0 running:0 transferring:4 holding:3 finished:0 failed:0. Users: [graeme.stewart?Z](#). Releases: [Atlas-14.2.10:Z](#). Sites: [UKI-SCOTGRID-GLASGOW:Z](#).

Showing 7 jobs modified from 2009-03-19 22:58:57 to 2009-03-19 22:42:57

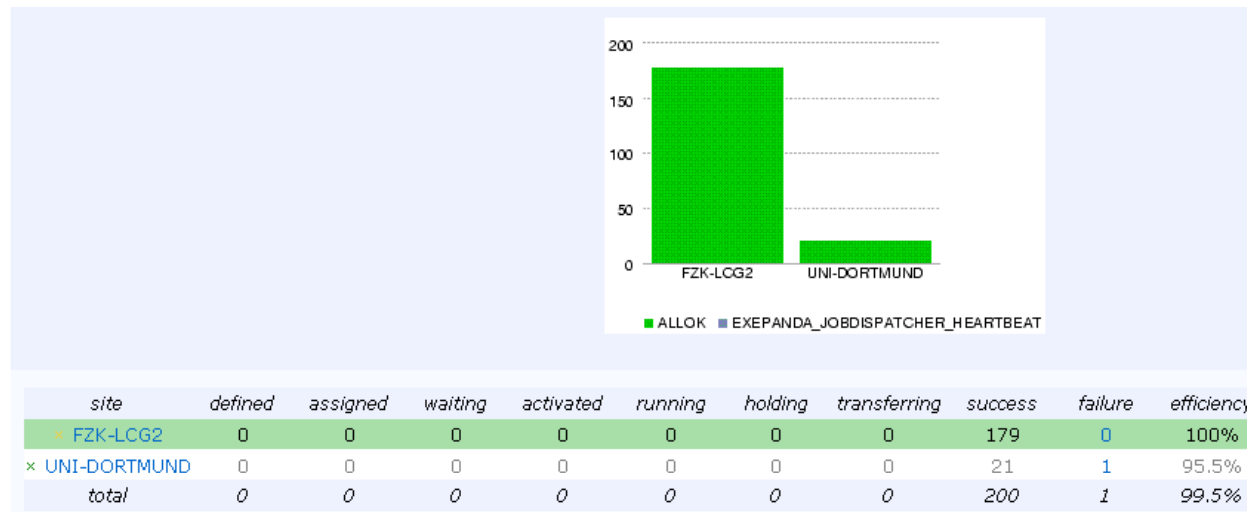
PandaID, Owner	Job	Status	Created	Time to start	Duration	Ended/Modified	Cloud Site, Type	Priority
<a href="#">1002281930</a> <a href="#">graeme.stewart</a>	trans=csc_simul_trf.py, pkg=AtlasProduction/14.2.10.1 In: <a href="#">mc08.108001.PythiaPhotonJet1_evgen.EVNT.e344_tid023729</a> Out: <a href="#">testpanda.graeme.6325fcd8-79a1-4488-a5a7-643ce0d8e8ac_tid999990</a>	transferring	03-19 14:20	8:09:42	0:16:38	03-19 22:58	<a href="#">UKI/UKI-SCOTGRID-GLASGOW</a> , test	1001
<a href="#">1002281929</a> <a href="#">graeme.stewart</a>	trans=csc_simul_trf.py, pkg=AtlasProduction/14.2.10.1 In: <a href="#">mc08.108001.PythiaPhotonJet1_evgen.EVNT.e344_tid023729</a> Out: <a href="#">testpanda.graeme.352e5d0cc-1ae6-4adc-8e0a-8c0dc9274891_tid999990</a>	holding	03-19 14:20	8:09:43	0:16:32	03-19 22:46	<a href="#">UKI/UKI-SCOTGRID-GLASGOW</a> , test	1001
<a href="#">1002281928</a> <a href="#">graeme.stewart</a>	trans=csc_simul_trf.py, pkg=AtlasProduction/14.2.10.1 In: <a href="#">mc08.108001.PythiaPhotonJet1_evgen.EVNT.e344_tid023729</a> Out: <a href="#">testpanda.graeme.c1d8b2ba-b311-4e01-9269-2786a1ac9886_tid999990</a>	transferring	03-19 14:20	8:09:44	0:16:32	03-19 22:58	<a href="#">UKI/UKI-SCOTGRID-GLASGOW</a> , test	1001
<a href="#">1002281927</a> <a href="#">graeme.stewart</a>	trans=csc_simul_trf.py, pkg=AtlasProduction/14.2.10.1	transferring	03-19 14:20	8:09:21	0:14:32	03-19 22:58	<a href="#">UKI/UKI-SCOTGRID-GLASGOW</a> , test	1001

- If the site runs all it's test jobs successfully it will be set online  
Graeme Stewart: ATLAS Computing



# Production System Functional Tests

- In addition to these very targeted site specific tests we run weekly 'production system functional tests' for the whole cloud
  - Jobs are similar short evgens to the site targeted tests
- These are very useful, e.g., after a Tier-1 downtime



- Good test of whole cloud, especially if there is little or no production
  - These tests are becoming more routine



# User Analysis: Ganga Robot



- For user analysis jobs there is a similar 'standard candle' analysis job sent every day through the ganga framework

GangaRobot

Monthly Summary: 03/2009

This report was generated: 19 March 2009 - 15:09

Site [01](#) [01](#) [01](#) [01](#) [02](#) [02](#) [02](#) [02](#) [02](#) [03](#) [03](#) [03](#) [03](#) [04](#) [04](#) [04](#) [04](#) [05](#) [05](#) [05](#) [05](#) [06](#) [06](#) [06](#) [06](#) [07](#) [07](#) [07](#) [07](#) [08](#) [08](#) [08](#) [08](#)

DE	01	01	01	01	02	02	02	02	02	03	03	03	03	04	04	04	04	05	05	05	05	06	06	06	06	07	07	07	07	08	08	08	08
CSCS-LCG2_DATADISK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
CSCS-LCG2_MCDISK	A	C	0	0	A	C	0	0	A	C	0	0	A	C	0	0	A	C	0	0	A	C	0	0	A	C	0	0	A	C	0	0	
CYFRONET-LCG2_DATADISK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
CYFRONET-LCG2_MCDISK	C	C	C	C	C	C	C	C	C	C	C	C	C	F	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C		
CYFRONET-LCG2_SCRATCHDISK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
DESY-HH_ALFA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
DESY-HH_DATADISK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
DESY-HH_MCDISK	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C		

Tests of different sw releases

Tests of different storage areas

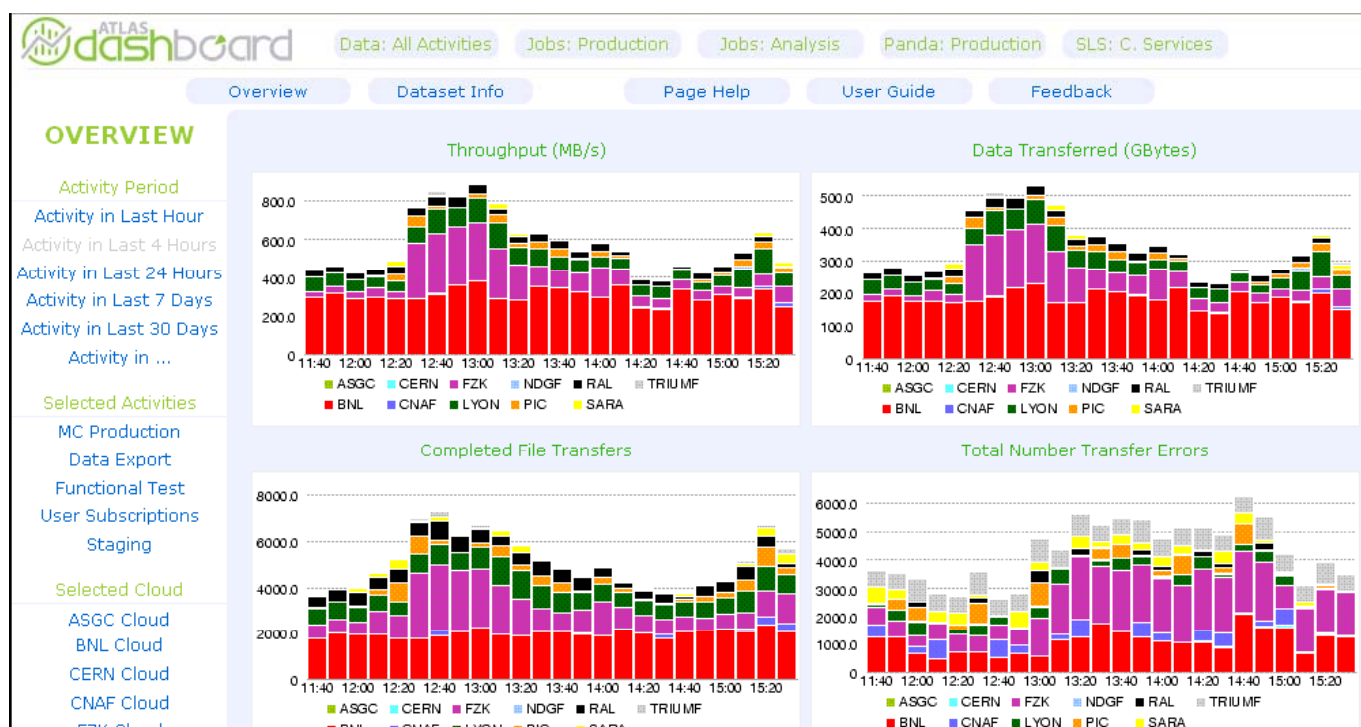
- If your site fails either of the last 2 tests it's automatically blacklisted for ATLAS user analysis in EGEE.
  - Distributed Analysis Shift Team will also raise GGUS tickets if a clear site problem is identified



# DDM Dashboard



- Shifters also monitor the ATLAS DDM Dashboard
  - This monitors file transfer success rates across the grid





# Drill Down to Sites

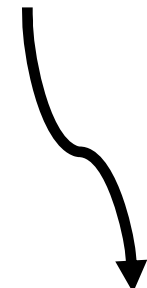


Activity Summary ('2009-03-19 11:40' to '2009-03-19 15:40')

Click on the cloud name to view list of sites

Cloud	Transfers			Registrations		Errors			Services
	Efficiency	Throughput	Successes	Datasets	Files	Transfer	Registration	Services	Grid
ASGC	0%	0 MB/s	0	0	0	0	0	0	
BNL	64%	317 MB/s	49194	166	49408	27177	0	0	
CERN	0%	0 MB/s	0	0	0	451	0	0	
CNAF	20%	3 MB/s	1561	22	1561	6192	0	0	
FZK	47%	120 MB/s	31849	822	31879	35594	0	0	
LYON	76%	83 MB/s	20199	280	20240	6269	0	0	
NDGF	7%	1 MB/s	30	0	30	425	0	0	
PIC	54%	20 MB/s	6645	247	6699	5737	0	0	
RAL	83%	36 MB/s	12463	192	12468	2528	0	0	
SARA	35%	7 MB/s	3144	199	3174	5812	0	0	
TRIUMF	7%	3 MB/s	1376	92	1377	17015	0	0	

Scheduled Downtime

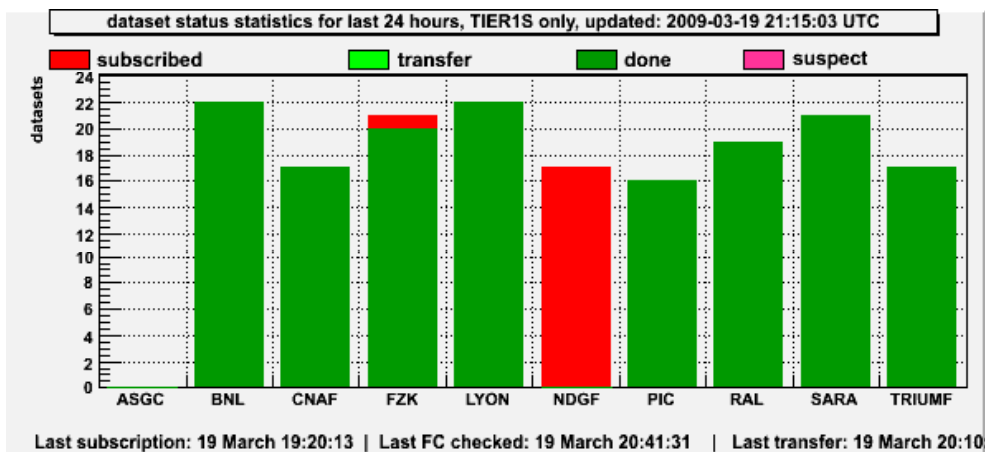


	Efficiency	Throughput	Successes	Datasets	Files	Transfer	Registration	Services	Grid	
+ CERN-PROD_MCDISK	0%	0 MB/s	0	0	0	0	445	0	0	sched
[FTS] FTS State [Failed] FTS Retries [1] Reason [DESTINATION error during TRANSFER_PREPARATION phase: [CONNECTION_ERROR] failed to contact on remote SRM [http://srm-atlas.cern.ch:8443/srm/managerv2]. Givin' up after 3 tries] Source Host [dcsrm.usatlas.bnl.gov]										253
[FTS] FTS State [Failed] FTS Retries [1] Reason [DESTINATION error during TRANSFER_PREPARATION phase: [CONNECTION_ERROR] failed to contact on remote SRM [http://srm-atlas.cern.ch:8443/srm/managerv2]. Givin' up after 3 tries] Source Host [ccsrm.in2p3.fr]										192

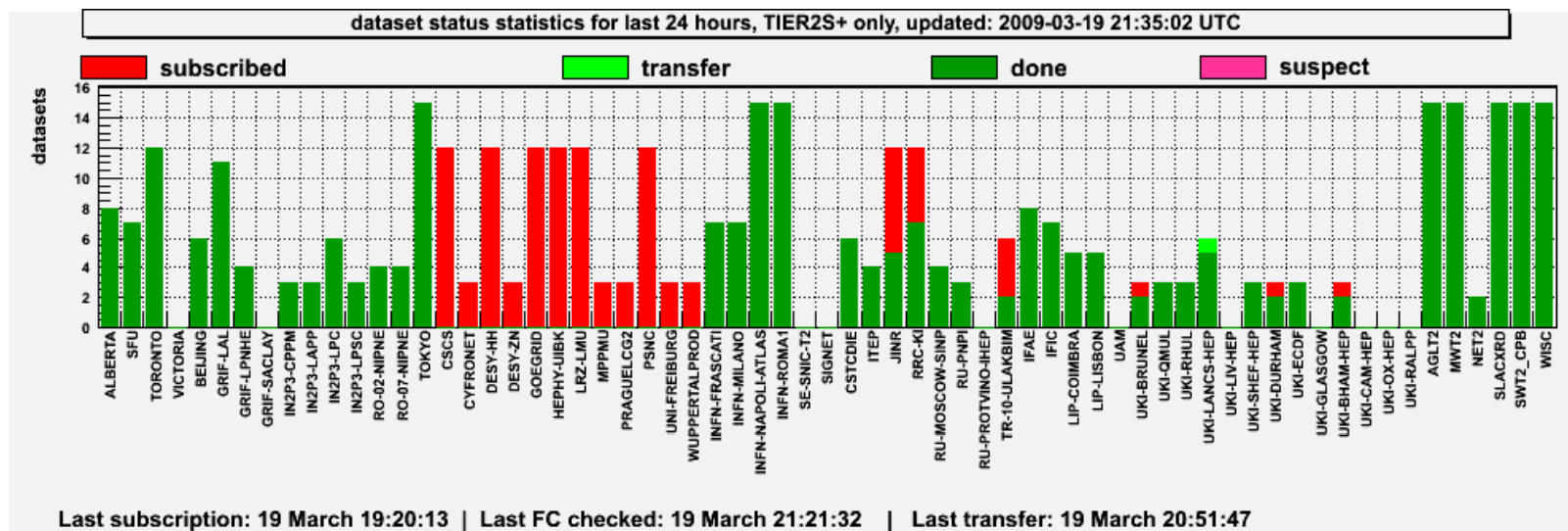




# Data Distribution Tests



- ATLAS also runs weekly Distributed Data Management Functional Tests
  - These tests distribute a small amount of /dev/random data to each ATLAS site according to the ATLAS computing model
  - As these run all the time then they test the system's functionality even when there are no other activities





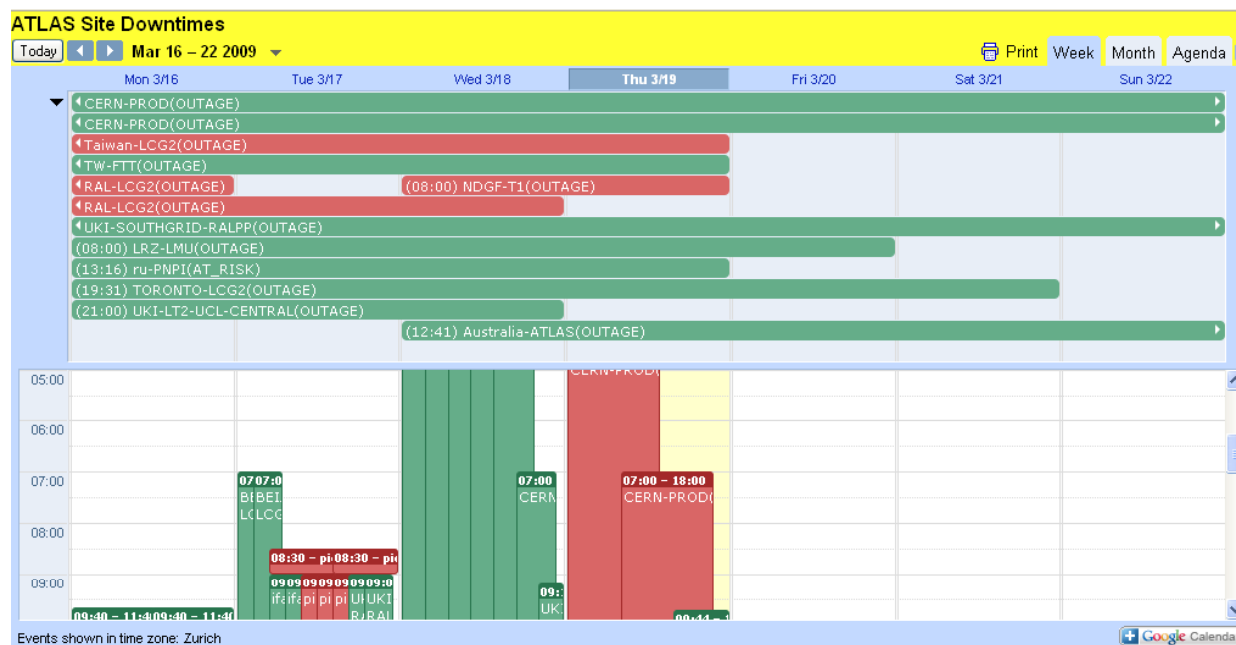
# Failing Site?

- The procedure is the usual one
  - E-log (for us)
  - GGUS (for site)
- If the problem would affect MC production the site will also be taken offline for production
  - But often a broken SE means you can't get input data anyway...
- If the problem is very grave then the site will be removed from the DDM site services machine and/or the subscription engine will be stopped to that cloud (e.g., T1 problems) because it swamps the dashboard
  - This prevents any transfer attempts to that site at all
    - Although DDM can try to get data from the site
  - It is a manual operation
    - So we don't like to do it, because it's easy to forget that a site/cloud was removed
  - This will improve soon with ATLAS Grid Information System (AGIS)
- After a period of suspension a cloud/site must succeed in DDM functional tests for 36 hours before being allowed to take ATLAS data



# How do we know about downtimes?

- Going to the GOC to see if a site is in downtime is far too slow for shifters triaging dozens of problems
- There is a feed from the GOC to an ATLAS Grid Downtime Calendar



- Problems:
  - Extensions not shown
  - Downtimes can be marked for secondary services



# Communications



- From us to you
  - We primarily use *GGUS* tickets for problems
    - Direct ticketing of sites is generally used and is much preferred by us
  - We also use our cloud contacts and *WLCG* operations meetings for requests
    - E.g., requests to change space token setup
  - And we have operational mailing lists which sites should sign up to
    - In particular: atlas-project-adc-operations@cern.ch
  - And weekly *ADC* operations meetings, several jamborees a year
    - To which sites are not just welcome, but encouraged to come
- From you to us
  - You can use *ggus* tickets
    - But responses may be slower as the ticket needs to be routed to the correct *ATLAS* responsables
  - Please do use your *ATLAS* cloud contacts
    - You should know who they are!
  - Or ask a question
    - On a mailing list
    - In a meeting - including the daily *WLCG* operations meeting