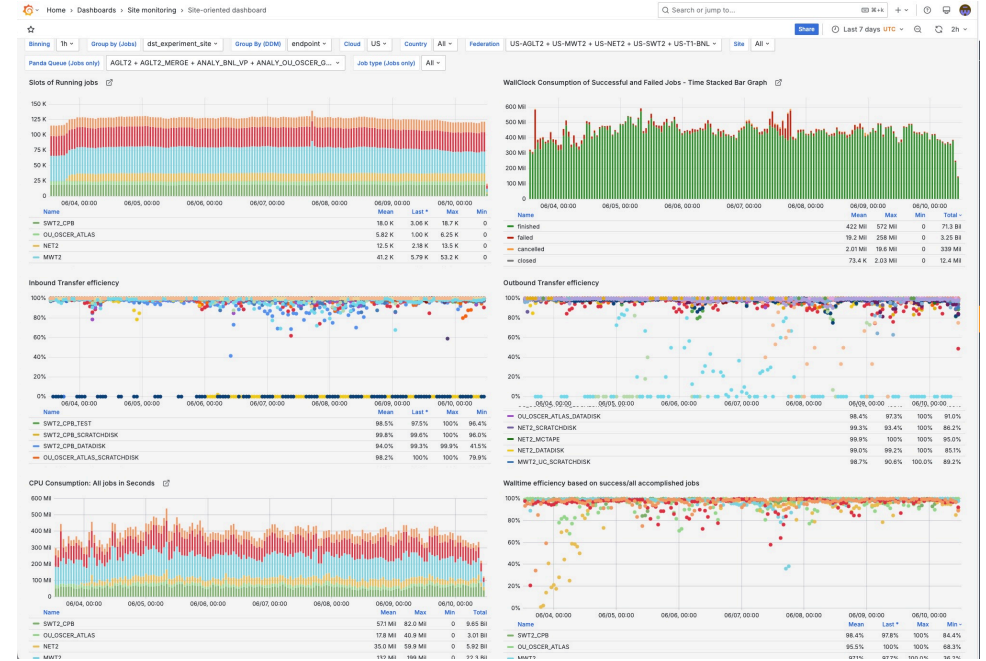
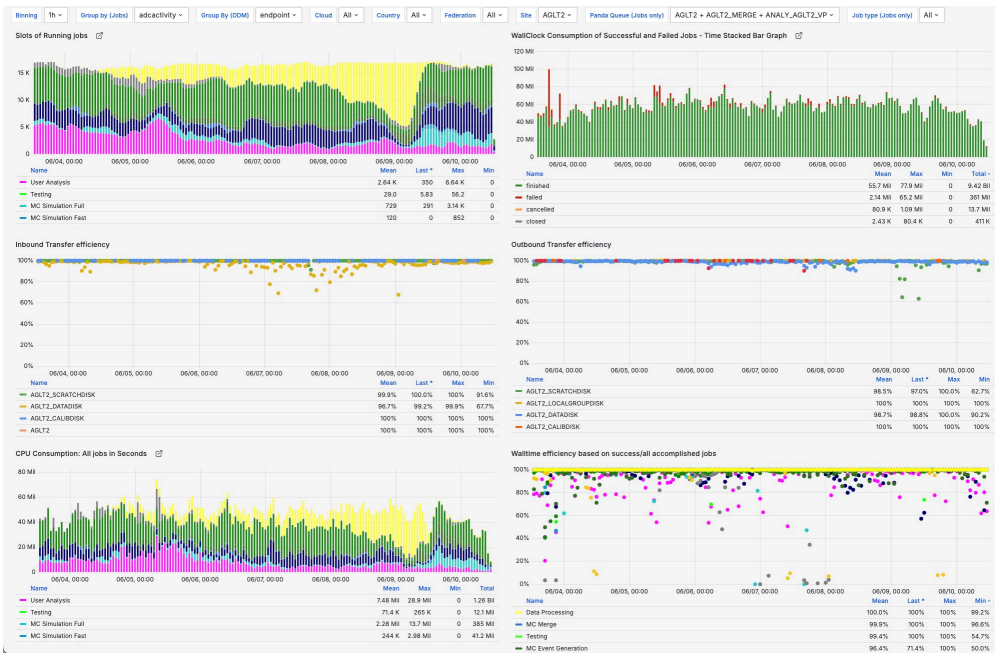


Watching for missing files at AGLT2

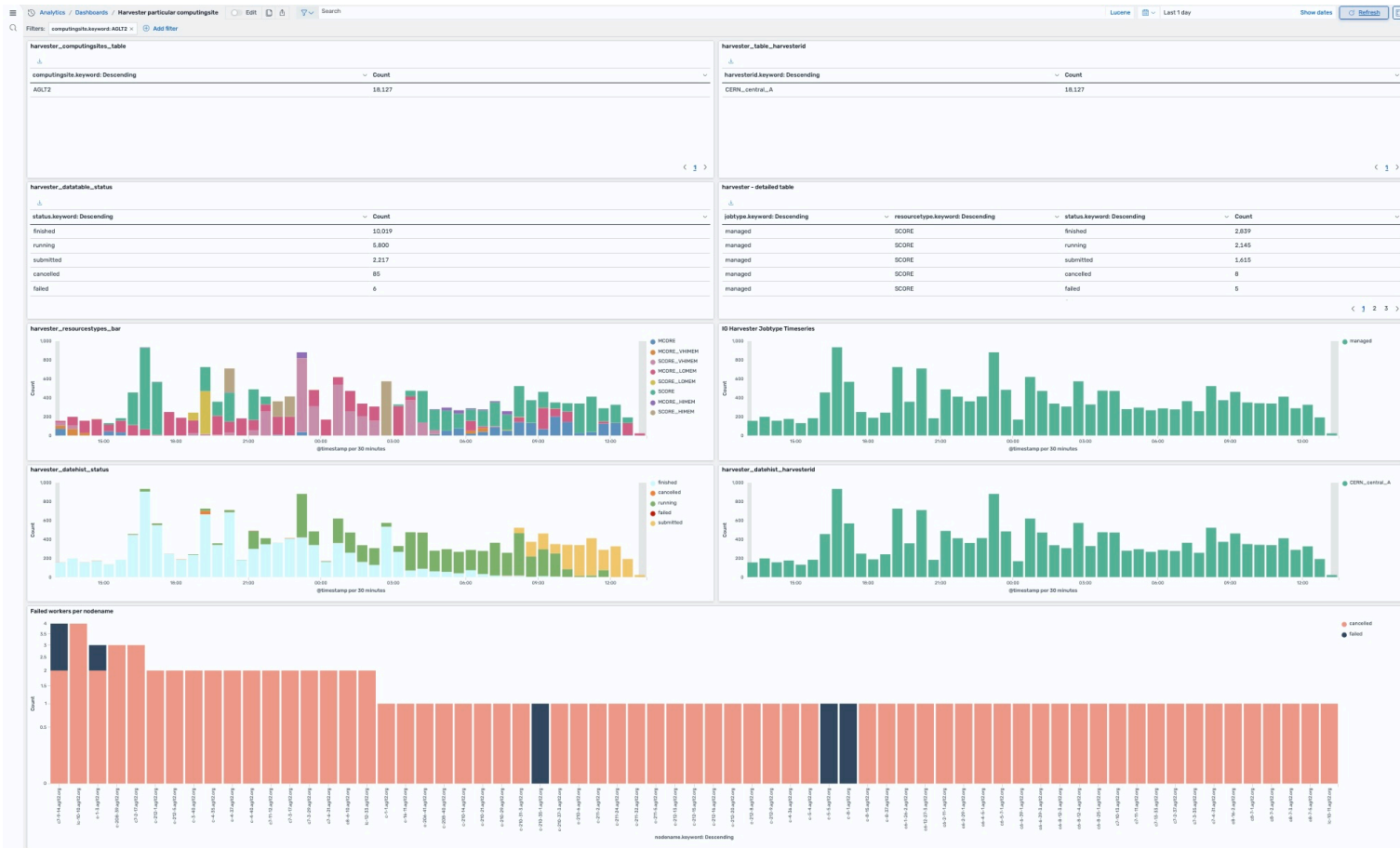
HTC26/F2F - Philippe Laurens - 10-Jun-2026

Quick daily check for operational issues

- The Panda, Harvester, WLCG site dashboards are very informative
- Over time we have built a better understanding/expertise for what to look for and how to interpret the wrapper, pilot and job errors.
- 2 minutes check through a few key pages to verify there is no issue.
- Note: The daily morning Ops meeting was a very useful addition.

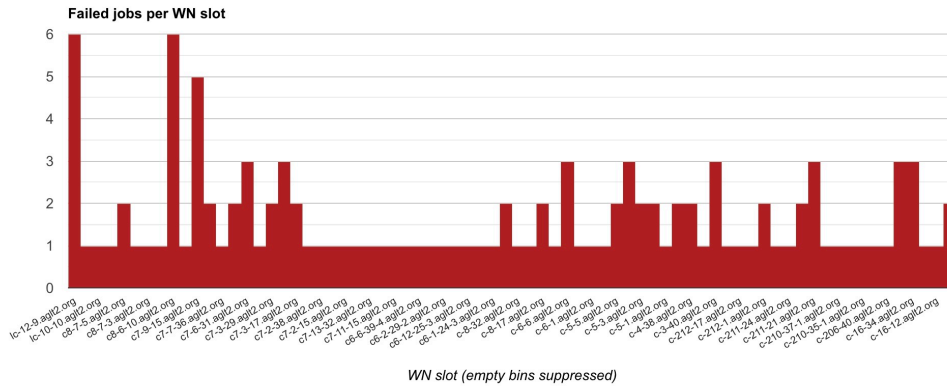


- For out site, glance at: number of jobs, proportion of failures, transfer efficiency (thanks!)
- And compare to the rest of USALTAS



- Glance at: histogram of jobs submitted/running and Nodes with high failures

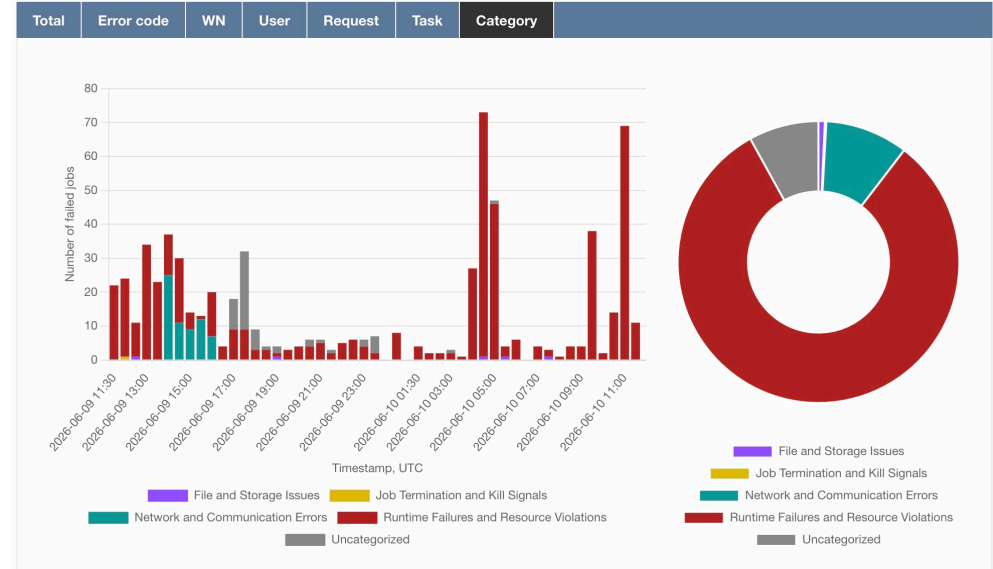
Worker node slot performance at AGLT2, sort by: count, copy time locked link



673 jobs found, go to job list: [failed](#), [all](#). Sort by [count](#), [alpha](#). Jump to [overall](#), [site](#), [user](#), [task](#) summary

Site: **AGLT2**

Job modification times in this listing range from 2026-06-09 11:48:31 to 2026-06-10 11:48:31, [copy time locked link](#)



- Glance at: Nodes with problems (again), magnitude/distribution of job errors, now color coded and with pie chart (thanks!)
- If there are concerns, look below for the list of errors, with sample error messages
- Node issues were somewhat frequent when cvmfs was unstable; not a problem at the moment.

Special Scrutiny for us: File Access Errors

- We had a dCache problem in Dec-2024 and had to restore databases from backup
- Over a few days there were files created by jobs on our site, or copied to our site, and written to our disks, but their pnfs information was lost.
- Rucio thought we had those files but jobs (and dCache) couldn't access them.
- For a period of months, we looked through job errors, identified missing files and also looked at the rest of each data set to preemptively find more missing files and declared them lost.
- (That issue also created some “dark data” which we finally cleared out quite recently)
- ... we still keep an eye on job file access errors ...

c8-7-3.aglt2.org	311	Total errors in 105 jobs.
	105	pilot:1099 Failed to stage-in file: mc23_13p6TeV:AOD.50108963_003791.pool.root.1 from AGLT2_DATADISK, Error on XrdCl:CopyProcess:Run(): [ERROR] Server responded with an error: [3011 No such file (source)]:failed to transfer files using copytools=['rucio']
	103	executor:5318 payload execution failed with 1
	103	transform:18 Invalid configuration of a reduction job
c8-7-5.aglt2.org	231	Total errors in 77 jobs.
	77	executor:5318 payload execution failed with 1
	77	pilot:1099 Package.xAOD::Event ERROR /build/atnight/localbuilds/nightlies/AnalysisBase/main/athena/Control/xAODRootAccess/Root/EventIO.cxx:81 (static std::unique_ptr<xAOD::Event> xAOD::Event::createAndReadFrom(TFile&)) Could not recognize file: root://xrootd.aglt2.org//atlasdatadisk/rucio/data24_13p6TeV/e1/6a/DAOD_JETM12.44070233_000012.pool.root.1
	77	transform:18 Invalid configuration of a reduction job
c-208-39.aglt2.org	219	Total errors in 73 jobs.
	73	executor:5318 payload execution failed with 9
	73	pilot:1099 Py:TrigBtagEmulationToolCfg ERROR BTagBtagToJetAssociator not found in input file root://xrootd.aglt2.org:1094//atlasdatadisk/rucio/data18_13TeV/e0/f7/DAOD_PHYS.46693036_000034.pool.root.1. Emulation of b-jet chains will be incorrect.
	73	transform:18 Invalid configuration of a reduction job
c8-7-6.aglt2.org	205	Total errors in 69 jobs.
	69	pilot:1099 MetaDataSvc ERROR Could not process new metadata source root://xrootd.aglt2.org:1094//atlasdatadisk/rucio/mc23_13p6TeV/8b/42/DAOD_LLJ1.48509947_000003.pool.root.1

- Pilot:1099 (file access errors for stage-in) “No such file” errors are the ones to check

c-212-16.aglt2.org	7	Total errors in 7 jobs.
	7	pilot:1361 Remote file could not be opened:Remote file(s) could not be opened: ['root://xrootd.aglt2.org:1094//atlasdatadisk/rucio/data25_13p6TeV/e0/52/DAOD_PHYS.LITE.50623100._000310.pool.root.1']
c7-9-14.aglt2.org	6	Total errors in 6 jobs.
	6	pilot:1361 Remote file could not be opened:Remote file(s) could not be opened: ['root://xrootd.aglt2.org:1094//atlasdatadisk/rucio/data25_13p6TeV/6b/ca/DAOD_PHYS.50623321._001815.pool.root.1']
c-6-1.aglt2.org	5	Total errors in 5 jobs.
	5	pilot:1361 Remote file could not be opened:Remote file(s) could not be opened: ['root://xrootd.aglt2.org:1094//atlasdatadisk/rucio/data25_13p6TeV/5b/4e/DAOD_PHYS.50623321._001802.pool.root.1']
c7-9-15.aglt2.org	5	Total errors in 5 jobs.
	5	pilot:1361 Remote file could not be opened:Remote file(s) could not be opened: ['root://xrootd.aglt2.org:1094//atlasdatadisk/rucio/data25_13p6TeV/b5/53/DAOD_PHYS.50623321._001804.pool.root.1']
c-212-18.aglt2.org	4	Total errors in 4 jobs.
	4	pilot:1361 Remote file could not be opened:Remote file(s) could not be opened: ['root://xrootd.aglt2.org:1094//atlasdatadisk/rucio/data25_13p6TeV/4c/72/DAOD_PHYS.50623321._001796.pool.root.1']

- Pilot:1361 (direct-access) “could not be opened” errors are the ones to check, but...

Error occurrences by job attribute	
computingelement (2)	gate01.aglt2.org:9619 (37) gate03.aglt2.org:9619 (27)
computingsite (1)	AGLT2 (64)
eventservice (1)	ordinary (64)
jobstatus (1)	failed (64)
processingtype (1)	eventIndex (64)
prodsourcelabel (1)	managed (64)
produsername (1)	sgonzale (64)
reqid (3)	69122 (32) 69119 (17) 69121 (15)
specialhandling (3)	ddm:rucio (64) hc:US (64) pc (64)
transformation (1)	POOLtoEI_tf.py (64)
workinggroup (1)	AP_SOFT (64)

- Pilot:1361 (direct-access) continued:
 - learned that a few users get occasional access errors even though we have the files.
 - In this case... no further check
 - For other users, we do check

What would help

- The pilot:1099/1361 errors seem “overloaded”.
- Having some dedicated code(s) for just access would be cleaner.
- Automatically building a (daily) list of missing files even better
- Automation (of any kind) would need to watch for false positives or do its own double check trying to access those files.