



Squid Cache for Generator Job Options

Graeme Stewart
2011-7-4



Reminder



CERN/Production

- Friday:
 - A set of production tasks (*e825) were DOS'ing <http://atlas-computing.web.cern.ch> to download a file MCopts.tar.gz
 - GDP paused/aborted the tasks
 - Server is now load balanced across 2 machines.
 - Will try using squid to cache the file.



From Dan's
AMOD report
last week

How to do this

- Detect site squid cache(s) in transform from environment
 - Use `http_proxy`, `ATLAS_PROXIES`, `FRONTIER_SERVER`
 - `ATLAS_PROXIES` we just invented - is it useful?
 - Comma separated proxy list, e.g.,
 - `ATLAS_PROXIES=nat005.gla.scotgrid.ac.uk:3128,atlas-squid.gridpp.rl.ac.uk:3128`
- Change the job options download to:
 - `wget --waitretry=5 --tries=3 --connect-timeout=20 --read-timeout=120`
 - Should work faster and better
 - Completely broken squid will waste ~70s wallclock
 - Squid which breaks on data download will waste ~360s
- This is currently in validation

Site/Cloud to check

- That there is a working squid cache at each site
 - Should be - needed by frontier
- Use http_proxy setting for complete site control - we respect it
- Otherwise ensure that the cache used by frontier allows downloads from
 - cern.ch and atlas-computing.web.cern.ch for authorised clients
 - AFAIU, there was older advice to secure squid caches based on destination, rather than source
 - You can check this with
 - \$ export http_proxy=my.cache:3128
 - \$ wget http://atlas-computing.web.cern.ch/atlas-computing/links/kitsDirectory/EvgenJobOpts/MC11JobOpts-00-00-19.tar.gz