Table of Contents

Glexec

Motivation

Worker nodes on the grid exhibit great diversity, making it difficult to offer uniform processing resources. A pilot job architecture, which probes the environment on the remote worker node before pulling down a payload job, can help. Pilot jobs become smart wrappers, preparing an appropriate environment for job execution and providing logging and monitoring capabilities. PanDA (Production and Distributed Analysis), an ATLAS and OSG workload management system, follows this design. However, in the simplest (and most efficient) pilot submission approach of identical pilots carrying the same identifying grid proxy, end-user accounting by the site can only be done with application-level information (PanDA maintains its own end-user accounting), and end-user jobs run with the identity and privileges of the proxy carried by the pilots, which may be seen as a security risk. To address these issues, we want to unable PanDA to use gLExec, a tool provided by EGEE which runs payload jobs under an end-user's identity. End-user proxies are pre-staged in a credential caching service, MyProxy, and the information needed by the pilots to access them is stored in the PanDA DB. gLExec then extracts from the user's proxy the proper identity under which to run.

Strategy

- First the end-user credentials are downloaded from a MyProxy server.
- A wrapper bash script is created where the entire environment is re-setup.Reason for that is after identity switch by gLExec the previously existing environment vanishes.
- Also a line to move the new current process from the new identity HOME directory to the previous pilot working directory is included in this wrapper script.
- And finally the actual payload command (buildJob or runAthena) is included in the wrapper.
- After creating this wrapper and modifying permissions to directories and files previously created in order to allow the new identity to read/write/execute files, gLExec is finally invoked to switch identity and run the wrapper under the new user.

Problems and issues.

There are some problems when using gLExec, mostly associated with the fact that the new user has no privilege anymore to read/write/execute files anywhere. So the old files and directories created by the pilot are almost always forbidden territory. Allowing the new user to write in the pilot working directory, in order to be able to generate the output root files and logs, is granted just by changing permissions (as commented in section Strategy). But there are a few more issues:

- The python_egg_cache. With recent Athena releases, at some sites it tries to unpack some python libraries. By default, setuptools tries to unpack the python libs in ~/.python_eggs. Instead of that, we are using the environment variable PYTHON_EGG_CACHE, making it to point to a random directory underneath /tmp/. Comment, the libraries can be unpacked at the installation time by using flag -Z.

- If a random directory underneath /tmp/ needs to be created, is not a good idea to use mktemp. Reason is that this type of commands make use of the value of the environment variable $TMPDIR, and there are sites where that variable is redefined by the local batch system to different directories where the new user may not have write permissions.

- With CREAM CE, by default, new directories and files are created with umask 0077, instead of the usual 0022. We are adding the command umask u=rwx,g=rwx,o=rwx to the glexec wrapper to facilitate the files created by the new user to have good read permissions so the pilot can stage them out.

- At some sites, the environment variable LD_PRELOAD points to a non-existing file. This makes gLExec to throw some content to the stderr, and this causes python interpreter to crash, unless the shell command gLExec is invoked using commands.get*output() method. But these methods are being deprecated. In order to be able to user subprocess, we are explicitly preventing the environment variable LD_PRELOAD from being inserted in the gLExec wrapper.

- Previously, the rule to setup a panda queue as gLExec enabled was to include in SchedConfig glexec=True. That rule has changed at the pathena level, now it requires to be glexec=uid in order to trigger the proxy delegation. The pilot code has been changed to follow the new policy.

- In order to allow matching between the pilot DN and the list of DNs used as a retrieval policy for the credentials stored in a MyProxy server, the regexp must include pipes | at the end and the beginning. For example:

```
$ myproxy-info -d -s myproxy.cern.ch
username: /O=GermanGrid/OU=LMU/CN=Johannes Elmsheuser
owner: /O=GermanGrid/OU=LMU/CN=Johannes Elmsheuser
   name: 125967faf3014f5e9fea7c68239d1ed2
   trusted retrieval policy: |Nurcan Ozturk|Jose Caballero|John R. Hover|
   timeleft: 166:04:47  (6.9 days)
```

Instead of

```
$ myproxy-info -d -s myproxy.cern.ch
username: /O=GermanGrid/OU=LMU/CN=Johannes Elmsheuser
owner: /O=GermanGrid/OU=LMU/CN=Johannes Elmsheuser
   name: 125967faf3014f5e9fea7c68239d1ed2
   trusted retrieval policy: Nurcan Ozturk|Jose Caballero|John R. Hover
   timeleft: 166:04:47  (6.9 days)
```

pandameta has been changed to add the pipes.

- There is an issue with the proxy delegation mechanism. This is the diagnosis:
    - I submit a job with pathena to ANALY_BNL_GLEXEC. Because that site has glexec='uid', a proxy was delegated.
    - This is what the MyProxy server says:

```
$ myproxy-info -s myproxy.cern.ch -l
/DC=org/DC=doegrids/OU=People/CN=Jose\ Caballero\ 511275
username: /DC=org/DC=doegrids/OU=People/CN=Jose Caballero 511275
owner: /DC=org/DC=doegrids/OU=People/CN=Jose Caballero 511275
name: 1edc84c65753441a8ef051f641bf92e0
trusted retrieval policy: Nurcan Ozturk|Jose Caballero|John R. Hover
timeleft: 166:39:34  (6.9 days)
```

-
    - After that I have tried to run a pilot on a job in

ANALY_GLASGOW_GLEXEC. It failed to retrieve the proxy from the MyProxy server, but Rodney is not an authorized retriever.

The conclusion is that when pathena delegates a proxy, it is retrievable only by the pilots who are supposed to run on that queue. But pilots trying to run jobs for the same users in different queues fail. There are two possible solutions:

-
    - to include the entire list of pilot DNs at the momment of the delegation. The problem is that that list is too long, and it could pass the maximum limit for the myproxy-init option -Z. Seems also that with the version of MyProxy server at CERN, that limit is still 255 characters.
    - to use a retrieval policy based on FQANs instead of DNs. That requires a version of MyProxy server higer than 3.6, as can be read here

http://grid.ncsa.illinois.edu/myproxy/voms/

This is the output of a test. This test seems to say that the MyProxy server at CERN is capable to handle retrieval policies based on FQAN. However, Maarten Litmaath ensures it shouldn't because it was not compiled to accept that mechanism.

```
$ myproxy-info -s myproxy.cern.ch -l /DC=org/DC=doegrids/OU=People/CN=Jose\ Caballero\ 511275
username: /DC=org/DC=doegrids/OU=People/CN=Jose Caballero 511275
owner: /DC=org/DC=doegrids/OU=People/CN=Jose Caballero 511275
 name: TEST1
 trusted retrieval policy: |Jose Caballero|John R. Hover|/atlas/Role=production|
 timeleft: 145:41:29  (6.1 days)
owner: /DC=org/DC=doegrids/OU=People/CN=Jose Caballero 511275
 name: 1edc84c65753441a8ef051f641bf92e0
 trusted retrieval policy: |Nurcan Ozturk|Jose Caballero|John R. Hover|Rodney Walker|
 timeleft: 129:12:06  (5.4 days)


$ voms-proxy-info -all
subject   : /DC=org/DC=doegrids/OU=People/CN=Nurcan Ozturk 18551/CN=proxy
issuer    : /DC=org/DC=doegrids/OU=People/CN=Nurcan Ozturk 18551
identity  : /DC=org/DC=doegrids/OU=People/CN=Nurcan Ozturk 18551
type      : proxy
strength  : 1024 bits
path      : x509_sm
timeleft  : 44:43:15
=== VO atlas extension information ===
VO        : atlas
subject   : /DC=org/DC=doegrids/OU=People/CN=Nurcan Ozturk 18551
issuer    : /DC=ch/DC=cern/OU=computers/CN=voms.cern.ch
attribute : /atlas/usatlas/Role=production/Capability=NULL
attribute : /atlas/Role=production/Capability=NULL
attribute : /atlas/lcg1/Role=NULL/Capability=NULL
attribute : /atlas/team/Role=NULL/Capability=NULL
attribute : /atlas/usatlas/Role=NULL/Capability=NULL
attribute : /atlas/Role=NULL/Capability=NULL
attribute : nickname = nurcan (atlas)
timeleft  : 44:43:14
uri       : voms.cern.ch:15001


$ myproxy-logon -s myproxy.cern.ch --no_passphrase -l '/DC=org/DC=doegrids/OU=People/CN=Jose Caba
A credential has been received for user /DC=org/DC=doegrids/OU=People/CN=Jose Caballero 511275 in
```

PanDA queues with gLExec enabled

queue	site	comments	HammerCloud tests
ANALY_BNL_GLEXEC	ANALY_BNL_GLEXEC	pilots submitted by hand	OK
ANALY_TEST-APF	ANALY_TEST-APF	pilots submitted with AutoPyFactory	OK
ANALY_GLASGOW_GLEXEC	ANALY_GLASGOW_GLEXEC	pilots submitted by hand	--
ANALY_OXFORD_GLEXEC	ANALY_OXFORD_GLEXEC	pilots submitted by hand	--
ANALY_GLEXEC_TRIUMF	ANALY_GLEXEC_TRIUMF	pilots submitted by hand	--
ANALY_CERN_GLEXEC	ANALY_CERN_GLEXEC	pilots submitted by hand	OK

Major updates:
-- JoseCaballero - 14-Mar-2011

Responsible: JoseCaballero

Never reviewed

This topic: PanDA > Glexec
Topic revision: r7 - 2011-04-23 - JoseCaballero