

How to install/configure Magda

UTA HEP/Computing/0006

Tomasz Wlodek

Abstract: I explain how I did install MAGDA database on UTA Atlas farm.
(heppc16.uta.edu)

Introduction: Magda is a prototype of database for cataloguing Atlas files developed by Torre Weneaus et al at BNL. We would like it to be included in the Atlas Grid-test bed.

Prerequisites: The machine on which Magda will run must have AFS client installed and must be able to access AFS server at BNL. Read my note *AFS client installation under Linux* in order to learn how to do that.

Where to learn more: There is a fair amount of Magda related stuff on the www under: <http://atlassw1.phy.bnl.gov/magda/info>

How to configure it:

1. Install AFS client. Once you can see BNL under /afs/usatlas.bnl.gov/ go to next point.
2. Go to page <http://atlassw1.phy.bnl.gov/magda/dyAdd.pl> and fill the registration forms. For UTA I have defined the site to be UT Arlington, host uta-atlas, the UTA files collection is named atlas-files-uta. File replication does not work yet.
3. A bunch of commands need to be executed before you can actually use Magda. You can find them in /home/mcfarm/magda/setup_magda. They are:
export MAGDA_HOST=uta-atlas
echo MAGDA_HOST=\$MAGDA_HOST
export OLDPATH=\$PATH
export MAGDA_PERL_DIR=/afs/usatlas.bnl.gov/software/perl/anl/bin
export MAGDA_DIRECTORY=/afs/usatlas.bnl.gov/project/magda/current
export PERL5LIB=/afs/usatlas.bnl.gov/project/magda/current
export PATH=\$MAGDA_PERL_DIR:\$PATH:\$MAGDA_DIRECTORY
They define what the Magda host is, they add the Perl directory to standard search path and they add magda source directory to search path.

Ok, now you can use Magda; basic commands are:

- magda_findfile *filename*
- magda_getfile *filename*
- magda_releasefile *filename*

- magda_putfile *filename location*
- magda_usage

They find the requested file, create a soft link from current directory to an existing file, delete a local soft link to remote file, put a file into requested location and print magda usage statistics.

The spider

There is a Perl script which looks at remote locations and updates data catalogues: dyFileSpider.pl. I wrote a Python wrapper that executes it at requested time intervals and keeps it in background. To start it go to /home/mcfarm/magda and do

```
python start_spider
```

to stop it do

```
python stop_spider
```

The spider is executed every 30 minutes (this can be tuned by changing the variable TimeBetweenSpiderChecks in file run_spider). Its output is logged to /home/mcfarm/magda/spider.log. (The logfile can be changed by modifying LogFileName in run_spider).

That's all folks for now. When I'll learn how to do data replication in Magda I'll write another note.

Acknowledgements: Many thanks to Torre Weneaus for his help.