Using the GRID at SLAC

This page documents some SLAC specific examples of using the GRID for dataset production and retrieval.

GRID Certificates

The centrally maintained ATLAS page for starting on the GRID is the place to start

- ATLAS GRID Workbook

DQ2 Setup at SLAC

DQ2 is the ATLAS data management system. There is significant documentation on it's general prinicples, usage, and troubleshooting here

- "How To" for DQ2
- DQ2 Troubleshooting Guide
- A DQ2 Tutorial

In order to being using the DQ2 tools at SLAC one simply needs to source the following script

```
source /afs/slac.stanford.edu/g/atlas/etc/hepix/GridSetup.(c)sh
```

**This script is automatically run for you if you are using the standard ATLAS setup and run "bash",
as described here: SLAC ATLAS Computing Environment**

You then need to get a GRID ticket, which will last 12 hours or so:

```
voms-proxy-init -voms atlas
```

Some simple dq2 commands

To add your files to a personal dataset (which can also be used by others):

```
dq2-put --long-surls -p lcg -L SLACXRD_LOCALGROUPDISK -s mydirectory user09.AndrewHaas477621.
SG_pythia_real_1000GeV.ESD.v1
```

where "mydirectory" is a directory containing the files you want to add to the dataset.
If you don't like answering "yes" to all the questions, include option "-a".
If you get "LFC exception [Could not create path with error Permission denied ... ", it's possible that the group membership of /grid/atlas/dq2/user09 is wrong. Complain to Non-Grid Jobs at SLAC!

If your directory contains ".pool.root" files, you need to setup a release first, so dq2-put can calculate the GUID for each file using
"pool_extractFileIdentifier". **Note: the GUID is created at the time the file is created, based upon the name of the file, machine, etc. To be safe,** *create*
the file with an original filename, by inserting some random string in it!

The dataset name you use should conform to the format "user09.DN.name.datatype.version", as above,
where DN is your identifier extracted from your certificate, and can be computed from:

```
python /afs/cern.ch/atlas/offline/external/GRID/ddm/Operations/utils/dn.py
```

To list the files in a dataset (note, you can use wildcards...):

```
dq2-ls -f user09.AndrewHaas477621.SG_pythia_real*.ESD.v1
```

To freeze the dataset:

```
dq2-freeze-dataset user09.AndrewHaas477621.SG_pythia_real_1000GeV.ESD.v1
```

To get the dataset:

```
cd /tmp
dq2-get user09.AndrewHaas477621.SG_pythia_real_1000GeV.ESD.v1
ls -lh user09.AndrewHaas477621.SG_pythia_real_1000GeV.ESD.v1/
```

Transfering large datasets

This is the old way:

To request an import of a large dataset to SLAC (it must be avilable first at BNL!):

```
dq2-register-subscription --archive <dataSet> SLACXRD_LOCALGROUPDISK
```

(the --archive flag makes sure it doesn't automatically get deleted after a week)

There's similar code that works with DQ2 containers:

```
dq2-register-subscription-container --archive  data09_cos.00121416.physics_L1Calo.merge.DPD_CALOCOMM.r733_p37/
SLACXRD_LOCALGROUPDISK
```

The new way:

Go to:

http://panda.cern.ch:25980/server/pandamon/query?mode=ddm_req

You'll have to first [register once with Panda](#) and [have your GRID certificate approved by the system](#) (make sure you have it imported into your Firefox browser)

It will take some time for the data to appear. You can check with:

```
dq2-ls -f <dataSet>
```

to see how many files are available locally.

And you can make a PoolFileCatalog.xml file directly:

```
dq2-ls -P <dataSet>
sed s%srm://osgserv04.slac.stanford.edu:8443/srm/v2/server?SFN=/xrootd/atlas%root://atl-xrdr//atlas/xrootd%g
PoolFileCatalog.xml >! PoolFileCatalog.xml
```

PATHENA Setup at SLAC

A commonly used set of tools for distributed analysis is [PANDA](#)

- [The PanDA Production and Distributed Analysis System](#)
- [Performing distributed analsysis with PANDA: PATHENA](#)
- [Client tools for Panda analysis jobs](#)

In order to being using these tools at SLAC, one simply needs to source 1 script and set 1 environment variable

```
source /afs/slac/g/atlas/packages/panda-client/etc/panda/panda_setup.sh
export PATHENA_GRID_SETUP_SH="/afs/slac/package/vdt/wlcg-client/setup.sh"
```