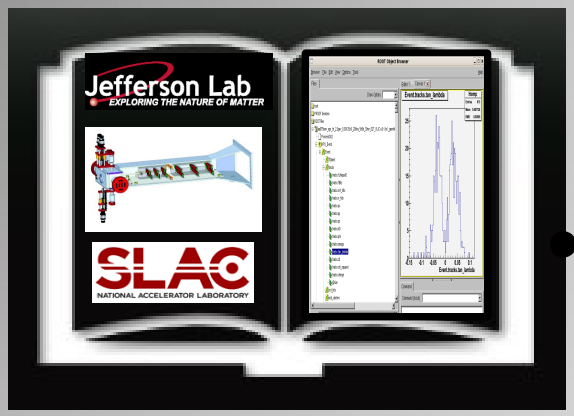
HPS

Data Handling and Storage

Presented by Homer NEAL for the



BANK of DOUBLE64s len (ints): 21 tag: 10 num: 20 datalen (bytes): 80 <#children: 0>

DOUBLE VALS
65.1108955719181
65.11049116003949
65.11049116003949
65.11049116003949
65.11049116003949
65.11049116003949
65.11049116003949
65.11049116003949
65.11049116003949
65.11049116003949

SOFTWARE REVIEW

@ SLAC NATIONAL ACCELERATOR LABORATORY

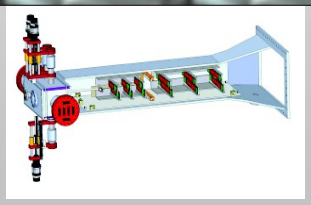
January 27th, 2014

Data Distribution

0.6 million CPU hours ⇒ ~280 slots needed constantly for three months

Jefferson Lab
EXPLORING THE NATURE OF MATTER

- Simulation +
- Reconstruction +
- Tape storage



clashps:: prod. account on ifarm

Recon (SLCIO)

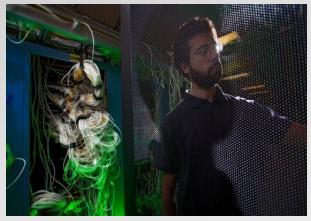
DST (ROOT)

~13 TB of output Expected for the Mock Data Challenge

DSTs are produced as part of the reconstruction process.

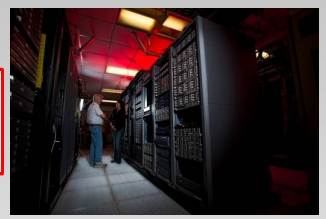
Transfers accomplished using → passwordless transfers from JLAB clashps account to hpdatsrv account on SLAC transfer machine using ssh keys.

Transfer machine: bbr-xfer14



SLAC
NATIONAL ACCELERATOR LABORATORY

30 TB disk: /nfs/slac/g/hps3



JLAB HPS Production Spaces

- /u/home/clashps has 2GB, backed up
- /u/group/hps has 20GB (backed up) and is visible from **ifarm** and **jlabl** machines

For production releases and scripts.

- /work/clas/clashps: 1TB of permanent (but not backed up)
  - export buffer
- /volatile/clas/clashps: 1TB for data staging
  - Note: This is 1TB reserved and 5 TB quota

*These two above are only visible from **ifarm** machines.*

Management of this space is typically done through FX.

scp -p clas12run@clonusr3:/misc/work/hps/hps_*dat /volatile/clas/clashps/data/

HPS raw data (EVIO) first arrives here

Access to JLAB production account (clashps)

```
ifarm1102> getent group clashps
clashps:*:936:weygand,clashps,desnault,fxgirod,ngraf,mgraham,homer,sarahp,omoreno,tvm,stepanya,boiarino,holtrop,hovanes,ungaro,uemura,mskolnik,gcharles
```

Our contact, FX, at JLAB is in charge of managing the clashps group.

Passwordless scp transfer of data to SLAC transfer server:

scp -v -q -p
/volatile/clas/clashps/data/<slcio file of reconstruction output>
hpdatsrv@bbr-xfer14.slac.stanford.edu:/hps/data/MDC/recon/

scp -v -q -p
/volatile/clas/clashps/data/<root file of DST output>
hpdatsrv@bbr-xfer14.slac.stanford.edu:/hps/data/MDC/DST/

An example script exists:

~clashps/exporthpsreclcio.sh

Simple adjustments needed for use for MDC and executed as a cron job.

Any new production output would be transferred to SLAC automatically.

- /hps/data/MDC/ is a staging area for data imports.
- By default, new files written there are staged out to HPSS (tape silo). Accessible through xrootd using:

  xrdcp root://sccs02//hps/data/....
- Files there have to be copied to the HPS data disk storage area: **/nfs/slac/g/hps3**

***Note**: I believe that I am the only ~HPS person that currently has access to the hpdatsrv account. Please contact Grace Tsai to adjust this. The SLAC expert assisting with this is Wilko Kroeger.*

Data Access

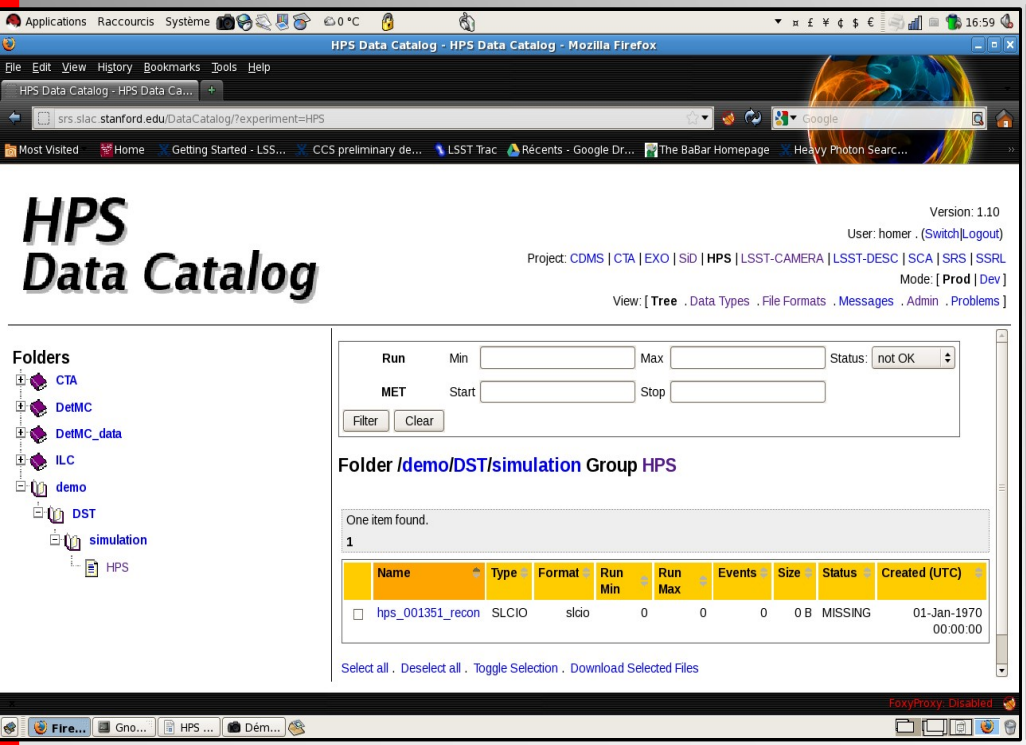
- In the SLAC SRS Data Catalog system:

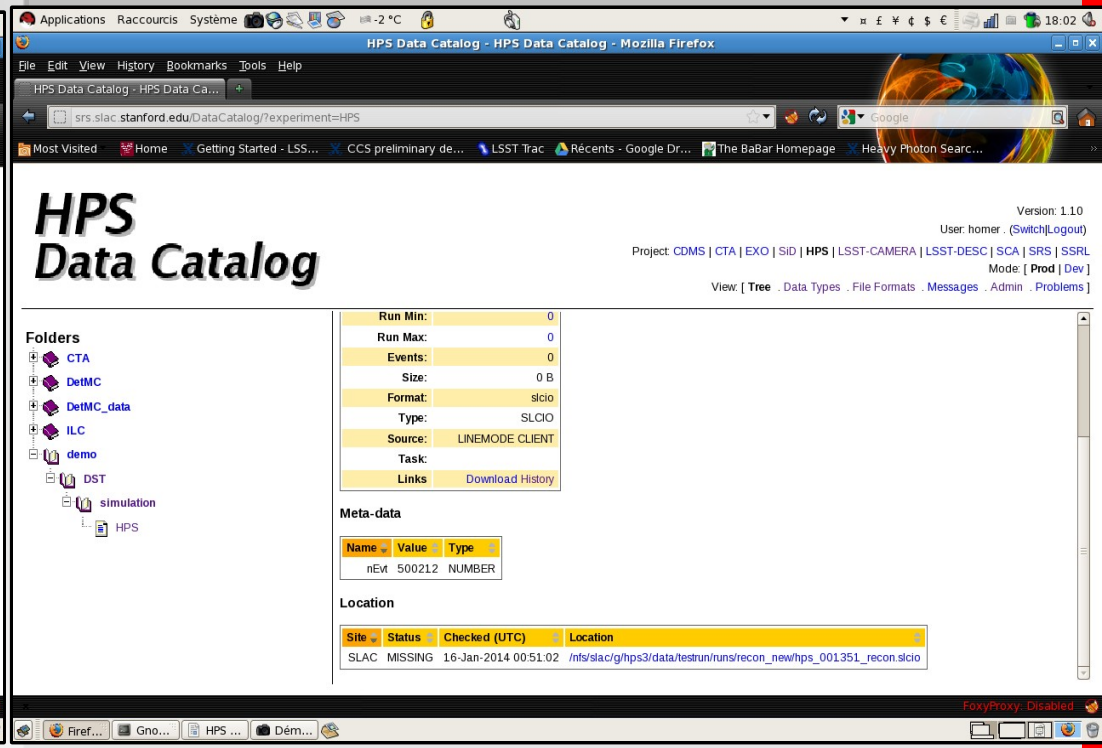
- A system for seeing what data is available with facilities for:
  - adding&viewing metadata
  - indicating quality
  - downloading data directly or making lists of paths to the data

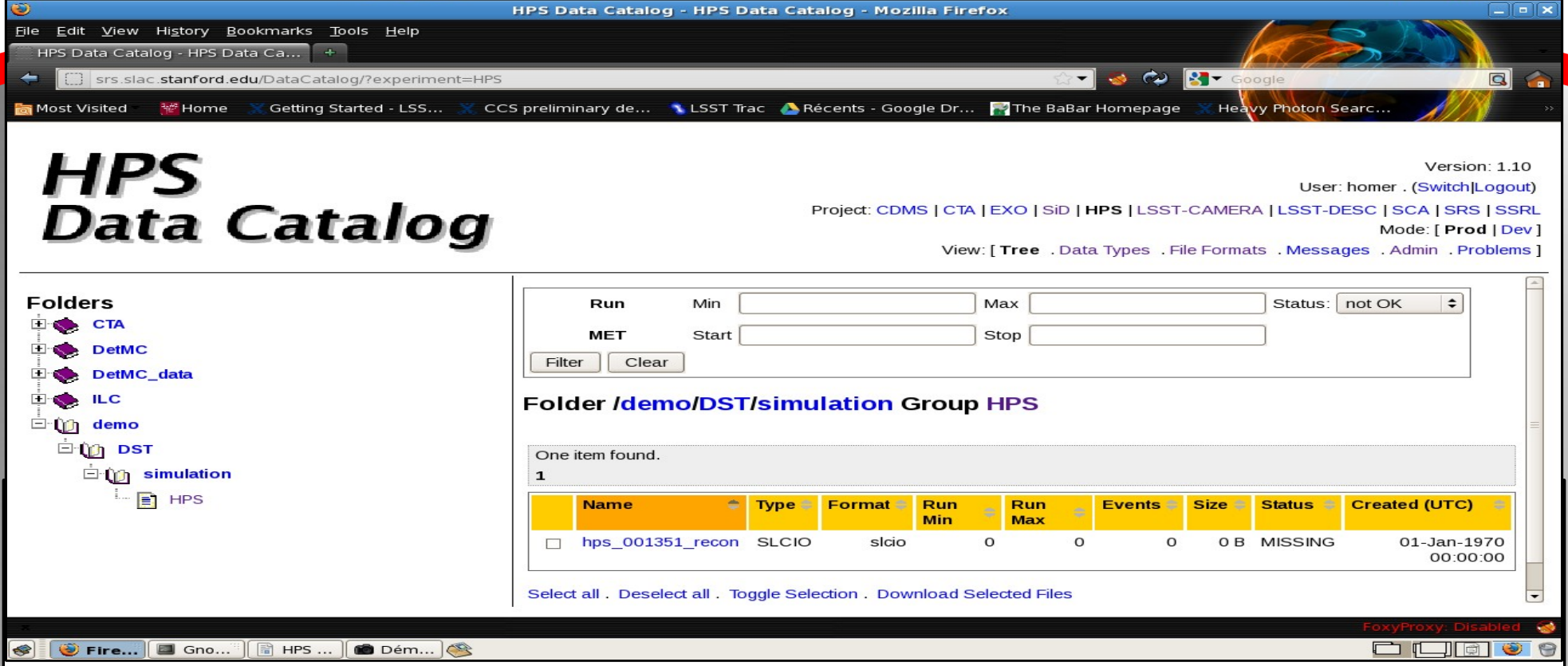
HPS Data Catalog Entry Insertion

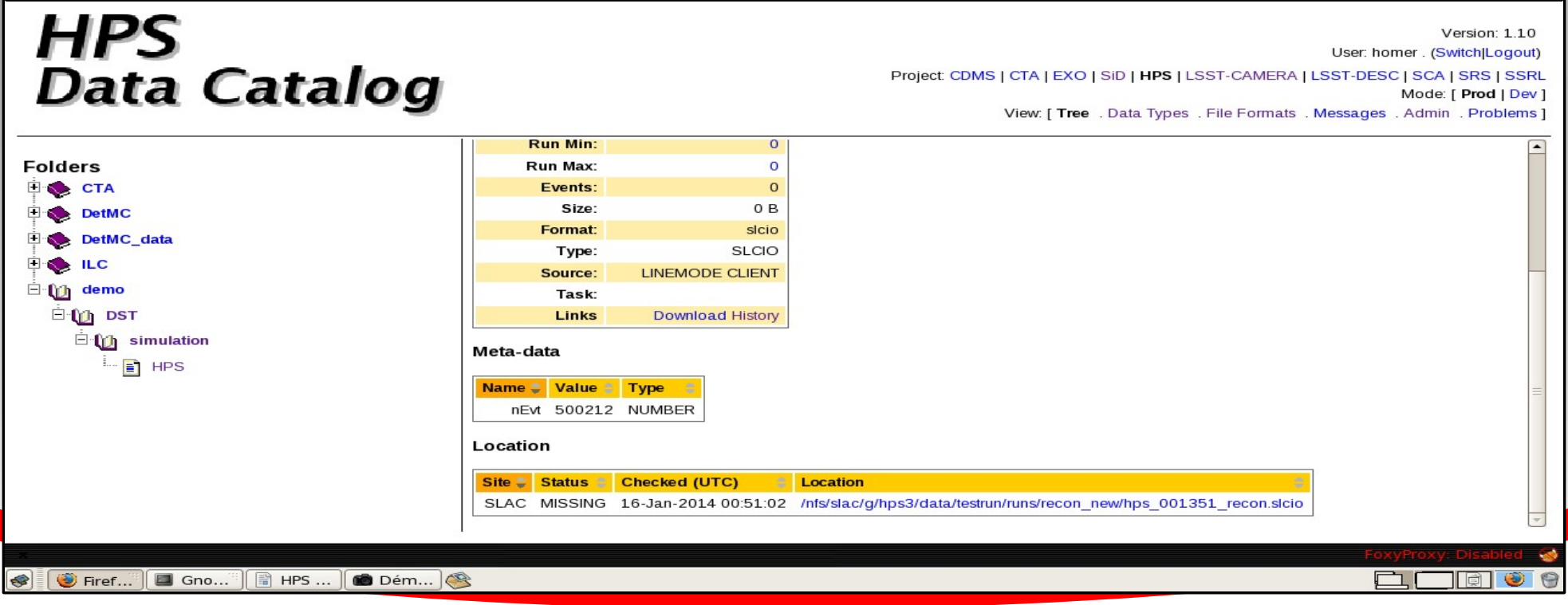
ifarm1102> **ssh homer@noric34.slac.stanford.edu "~srs/datacat/prod/datacat registerDataset --group HPS --define nEvt=500212 --site SLAC slcio demo/DST/simulation /nfs/slac/g/hps3/data/testrun/runs/recon_new/hps_001351_recon.slcio"**
homer@noric34.slac.stanford.edu's password:
Adding files to Data Catalog:
Site Name Type Location
SLAC hps_001351_recon slcio /nfs/slac/g/hps3/data/testrun/runs/recon_new/hps_001351_recon.slcio
INFO: 1 dataset registerd.

Note: An account with passwordless access will be needed for this. The hpdatsrv account can not be used for this.






HPS Data Catalog - HPS Data Catalog - Mozilla Firefox
File Edit View History Bookmarks Tools Help
HPS Data Catalog - HPS Data Ca...
srs.slac.stanford.edu/DataCatalog/?experiment=HPS
Most Visited Home Getting Started - LSS... CCS preliminary de... LSST Trac Récents - Google Dr... The BaBar Homepage Heavy Photon Searc...
HPS
Data Catalog
Version: 1.10
User: homer . (Switch|Logout)
Project: CDMS | CTA | EXO | SiD | HPS | LSST-CAMERA | LSST-DESC | SCA | SRS | SSRL
Mode: [Prod | Dev]
View: [Tree . Data Types . File Formats . Messages . Admin . Problems]
Folders
CTA
DetMC
DetMC_data
ILC
demo
DST
simulation
HPS
Run Min Max Status: not OK
MET Start Stop
Filter Clear
Folder /demo/DST/simulation Group HPS
One item found.
1
Name Type Format Run Min Run Max Events Size Status Created (UTC)
hps_001351_recon SLCIO slcio 0 0 0 0 B MISSING 01-Jan-1970 00:00:00
Select all . Deselect all . Toggle Selection . Download Selected Files
FoxyProxy: Disabled
Fire... Gno... HPS ... Dém...


HPS
Data Catalog
Version: 1.10
User: homer . (Switch|Logout)
Project: CDMS | CTA | EXO | SiD | HPS | LSST-CAMERA | LSST-DESC | SCA | SRS | SSRL
Mode: [Prod | Dev]
View: [Tree . Data Types . File Formats . Messages . Admin . Problems]
Folders
CTA
DetMC
DetMC_data
ILC
demo
DST
simulation
HPS
Run Min: 0
Run Max: 0
Events: 0
Size: 0 B
Format: slcio
Type: SLCIO
Source: LINEMODE CLIENT
Task:
Links Download History
Meta-data
Name Value Type
nEvt 500212 NUMBER
Location
Site Status Checked (UTC) Location
SLAC MISSING 16-Jan-2014 00:51:02 /nfs/slac/g/hps3/data/testrun/runs/recon_new/hps_001351_recon.slcio
FoxyProxy: Disabled
Firef... Gno... HPS ... Dém...

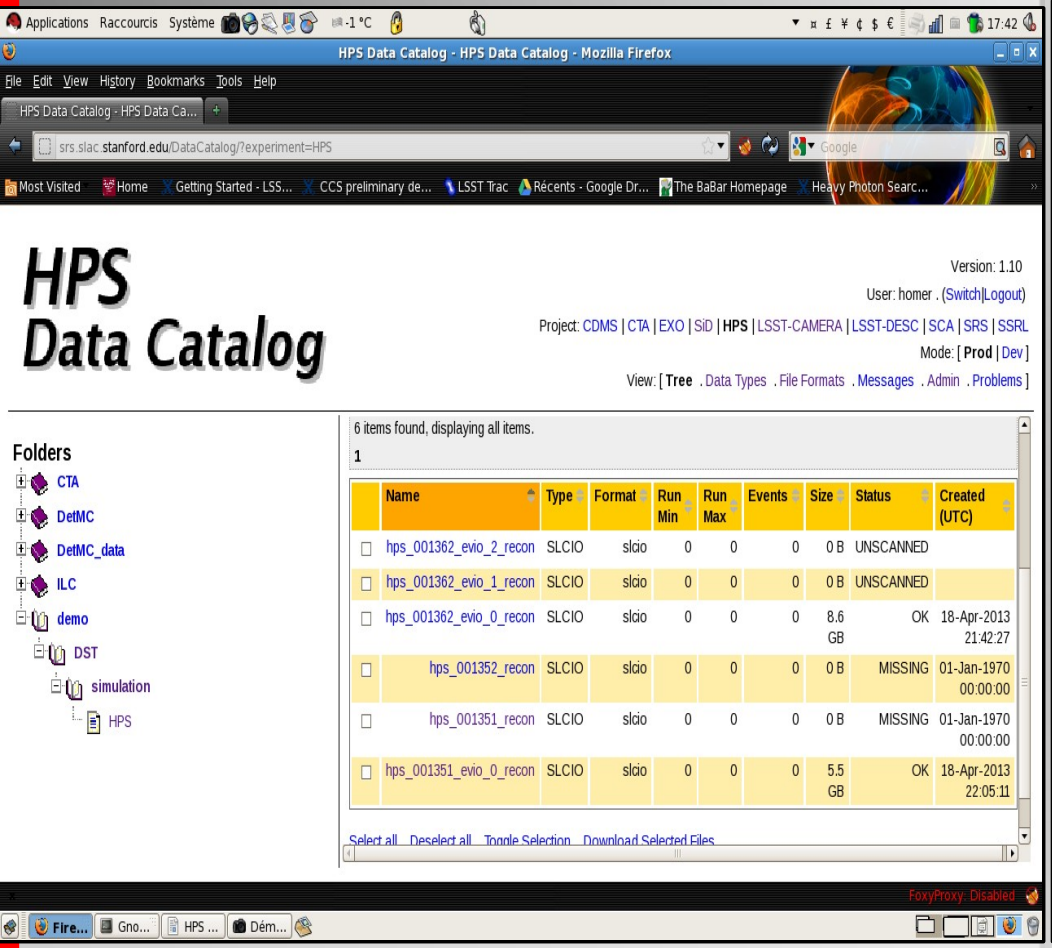
HPS Data Catalog & the Crawler

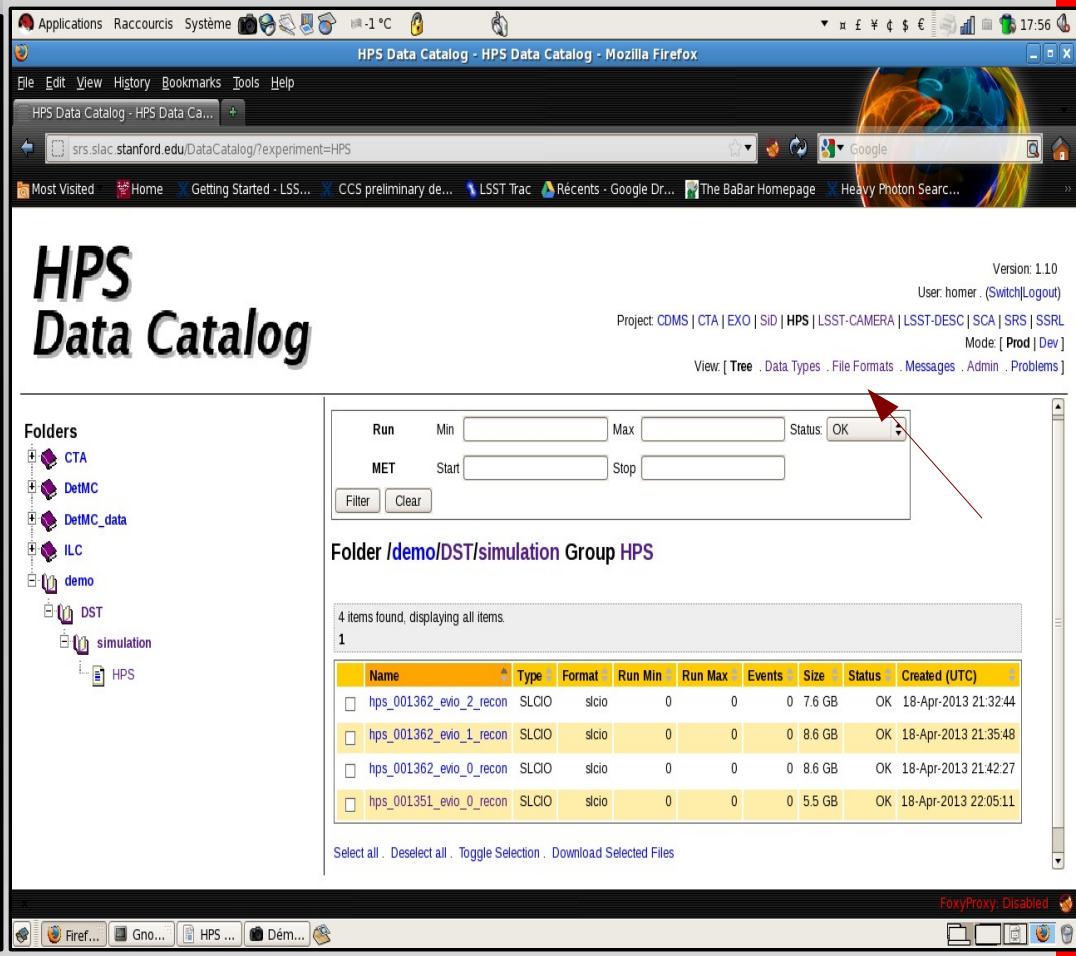
ifarm1102> whoami
clashps
ifarm1102> **ssh homer@noric34.slac.stanford.edu "~srs/datacat/prod/datacat registerDataset --group HPS --site SLAC slcio demo/DST/simulation /nfs/slac/g/hps3/data/testrun/runs/dcat/hps_001362_evio_0_recon.slcio"**
homer@noric34.slac.stanford.edu's password:
Adding files to Data Catalog:
Site Name Type Location
SLAC hps_001362_evio_0_recon slcio /nfs/slac/g/hps3/data/testrun/runs/dcat/hps_001362_evio_0_recon.slcio
INFO: 1 dataset registerd.

ifarm1102> **ssh homer@noric34.slac.stanford.edu "~srs/datacat/prod/datacat registerDataset --group HPS --site SLAC slcio demo/DST/simulation /nfs/slac/g/hps3/data/testrun/runs/dcat/hps_001362_evio_1_recon.slcio"**
homer@noric34.slac.stanford.edu's password:
Adding files to Data Catalog:
Site Name Type Location
SLAC hps_001362_evio_1_recon slcio /nfs/slac/g/hps3/data/testrun/runs/dcat/hps_001362_evio_1_recon.slcio
INFO: 1 dataset registerd
.

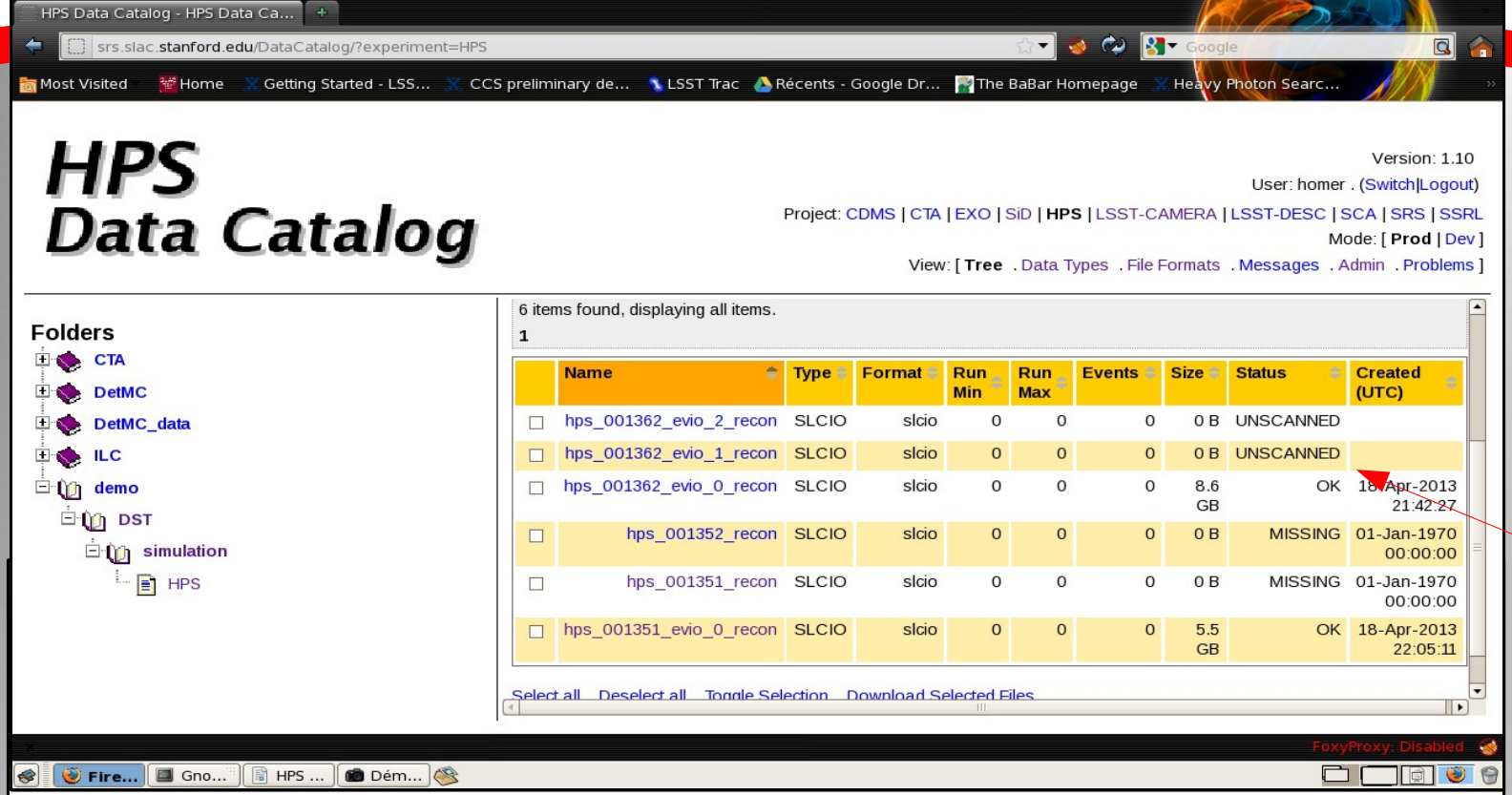
ifarm1102> **ssh homer@noric34.slac.stanford.edu "~srs/datacat/prod/datacat registerDataset --group HPS --site SLAC slcio demo/DST/simulation /nfs/slac/g/hps3/data/testrun/runs/dcat/hps_001362_evio_2_recon.slcio"**
homer@noric34.slac.stanford.edu's password:
Adding files to Data Catalog:
Site Name Type Location
SLAC hps_001362_evio_2_recon slcio /nfs/slac/g/hps3/data/testrun/runs/dcat/hps_001362_evio_2_recon.slcio
INFO: 1 dataset registerd.

HPS Data Catalog & the Crawler

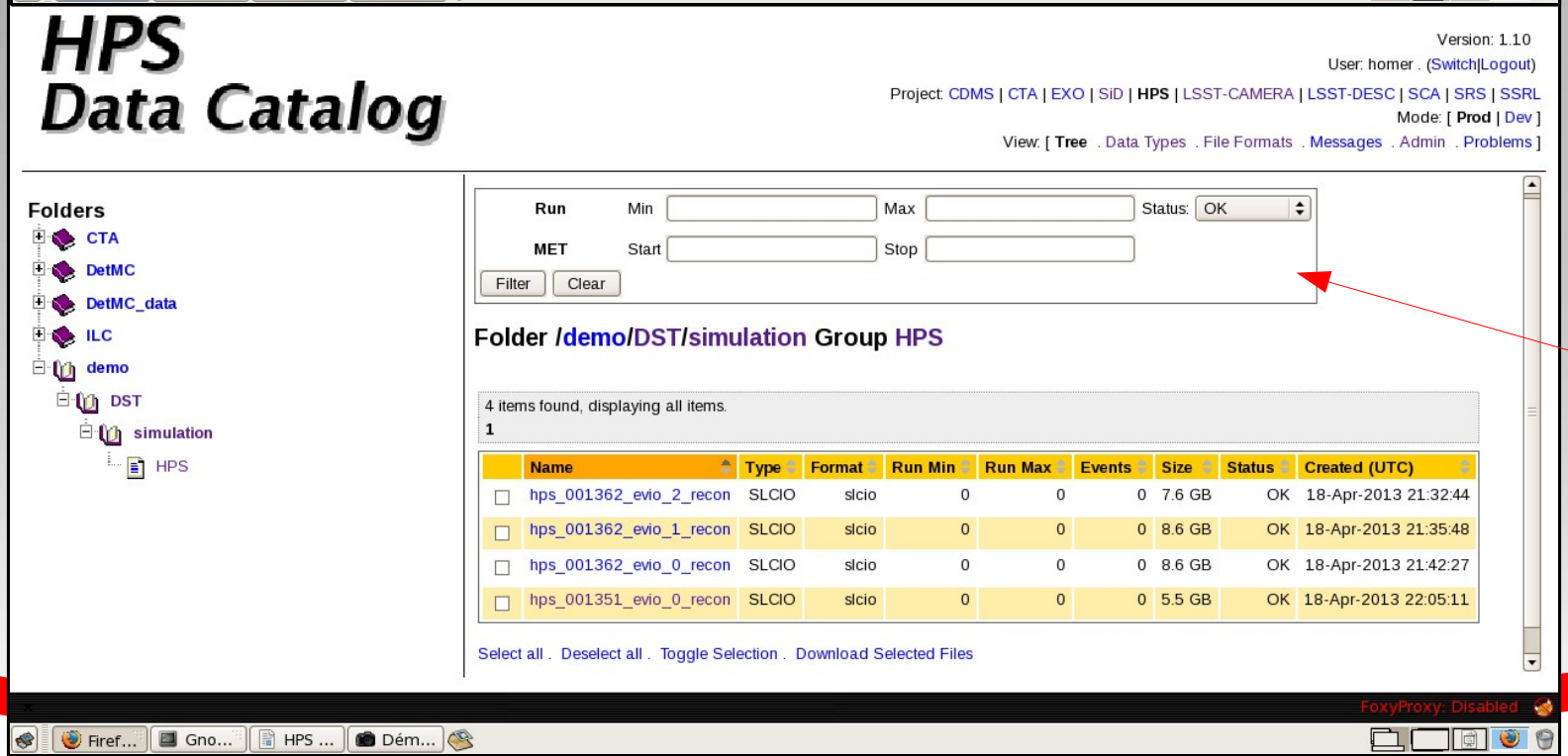




The Crawler will check if the file exists and report to the catalog if the file was found and its size and creation date.



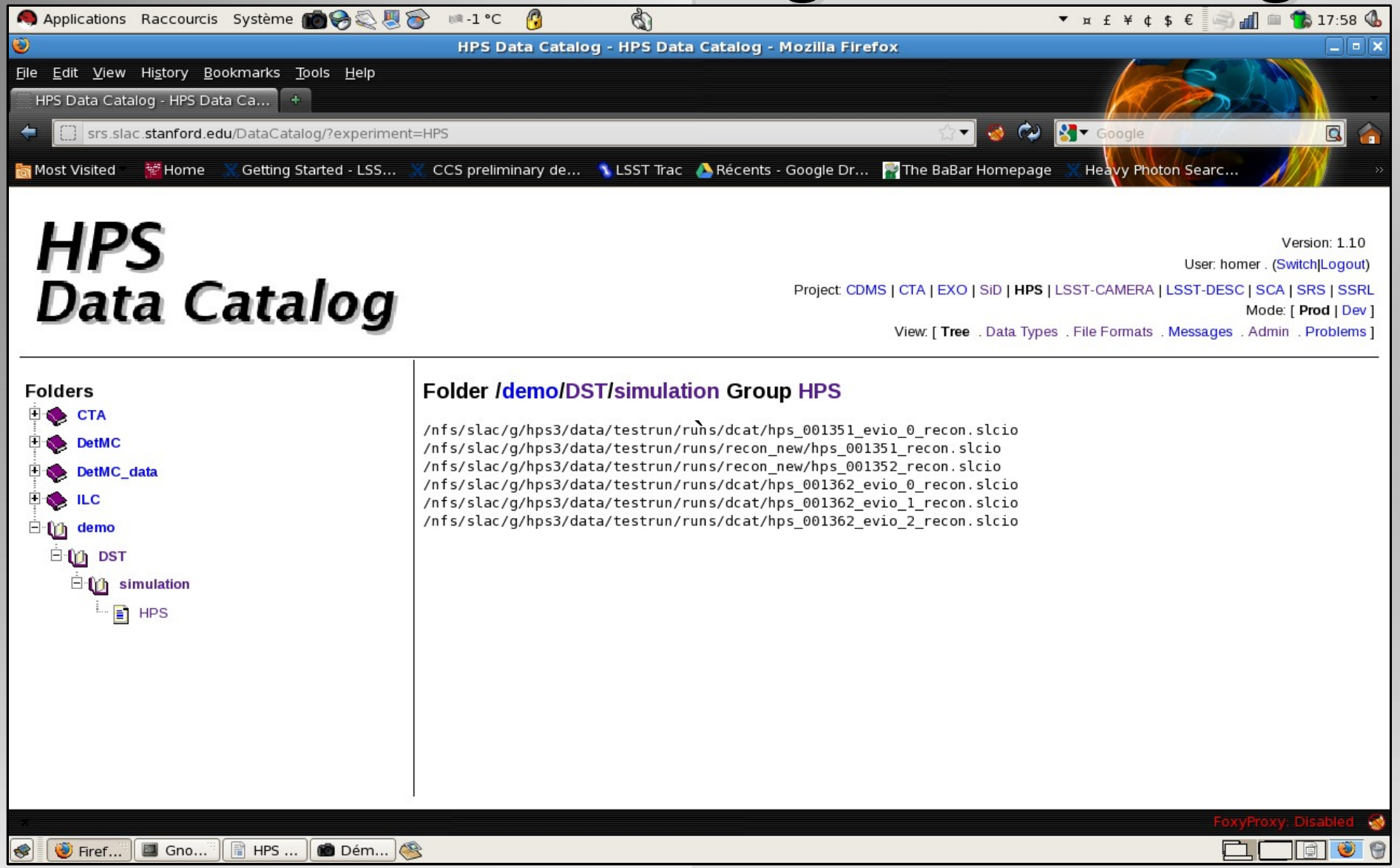
I had just inserted these entries.



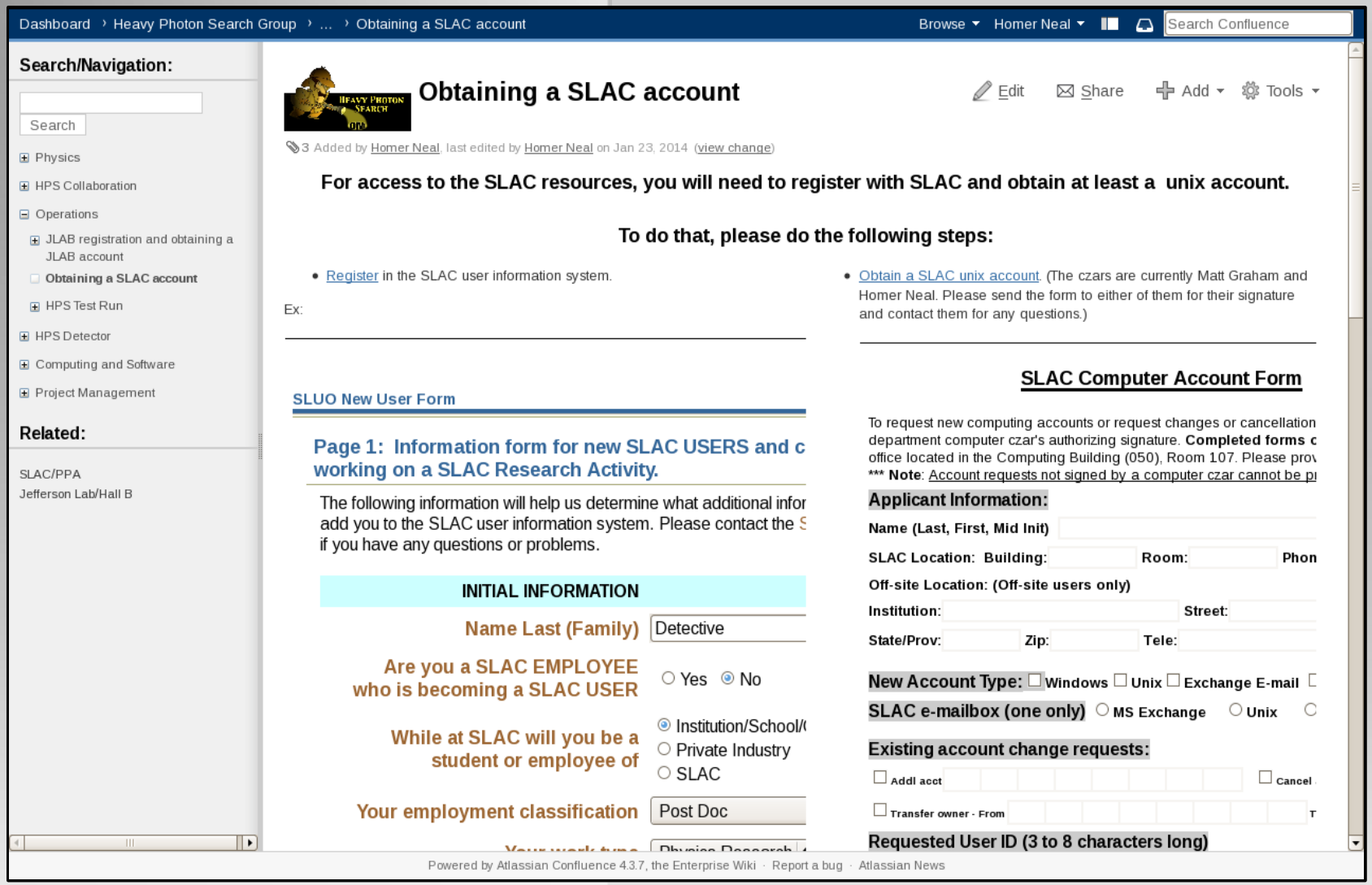
List only OK files.

← files found

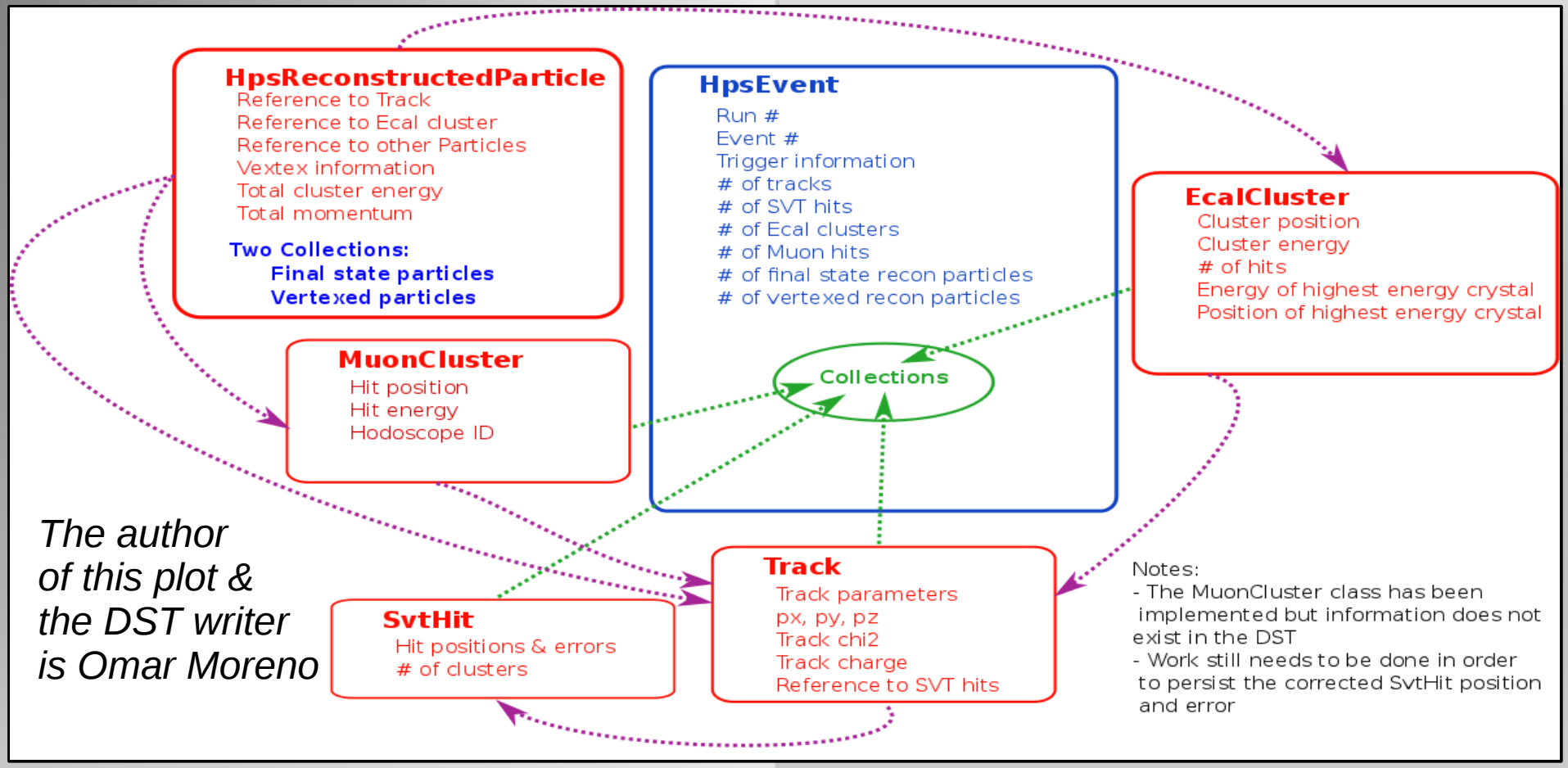
HPS Data Catalog File Listing



Access to SLAC Computing Resources



The HPS DST Format

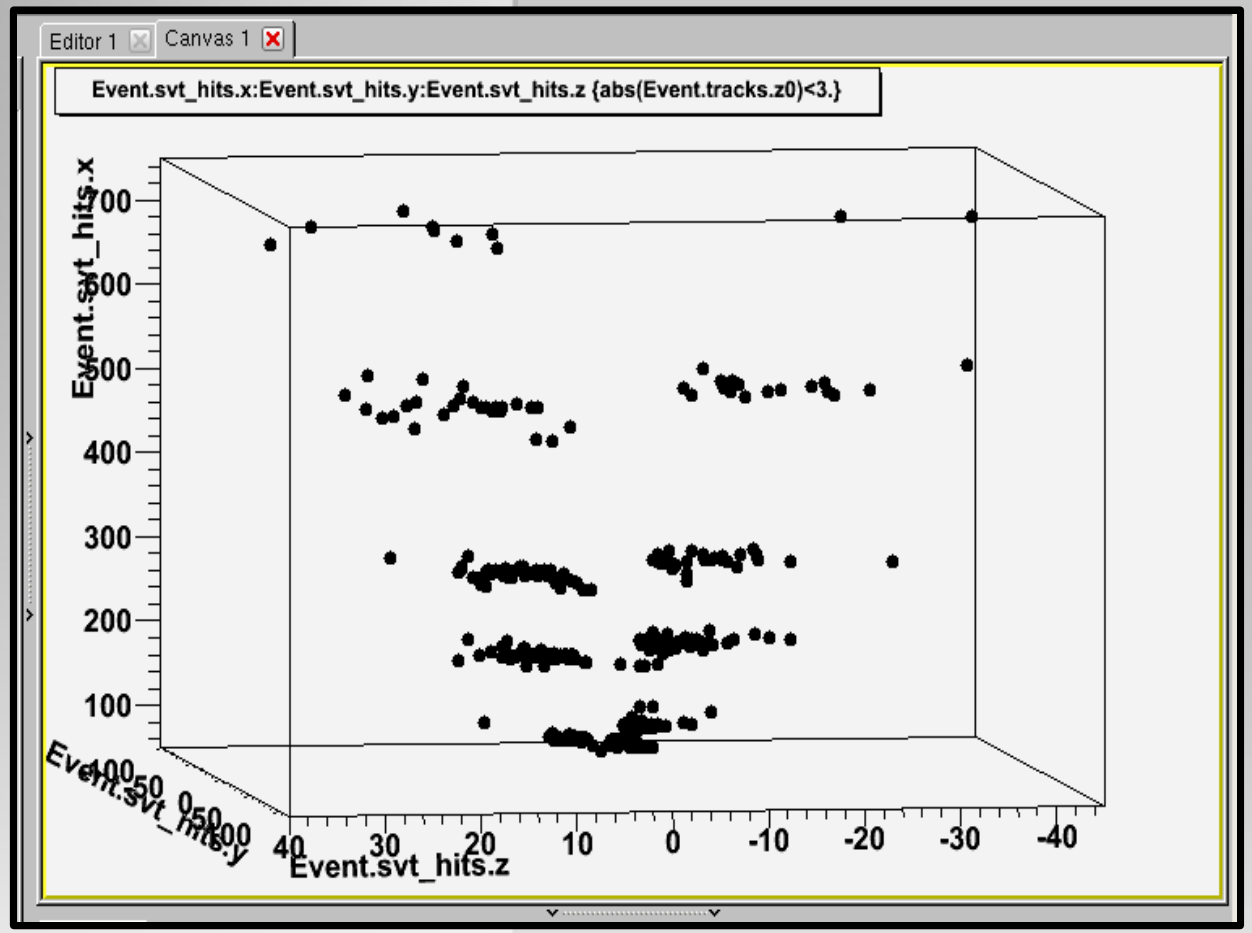


For a guide on how to use it see Omar's page at:
https://confluence.slac.stanford.edu/display/hpsg/HPS+Data+Summary+Tapes

The following are some interactive examples ...

Making a quick plot using the HPS DST

root [47] **HPS_Event->SetMarkerStyle(20);**
root [48] **HPS_Event->Draw("Event.svt_hits.x:Event.svt_hits.y:Event.svt_hits.z","abs(Event.tracks.z0)<3.");**



Real Analysis with the DSTs

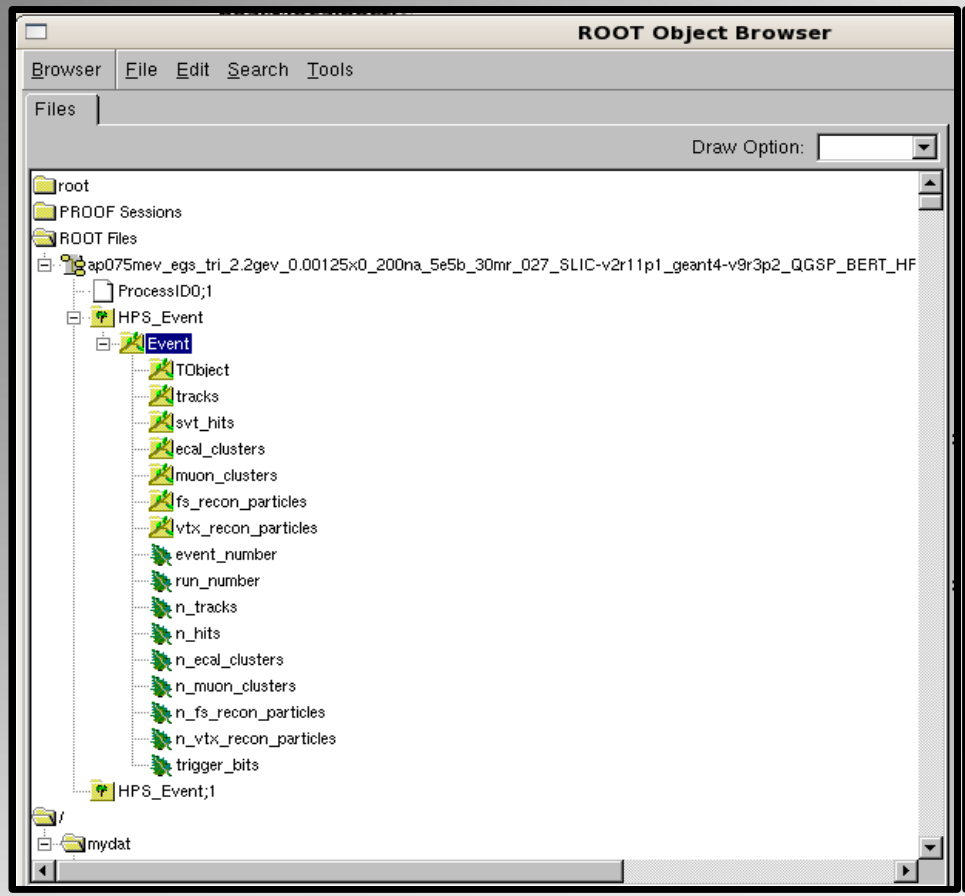
- Basic analysis examples are available on the following confluence page: https://confluence.slac.stanford.edu/display/hpsg/HPS+Data+Summary+Tapes
- This is a good starting point for analyses.

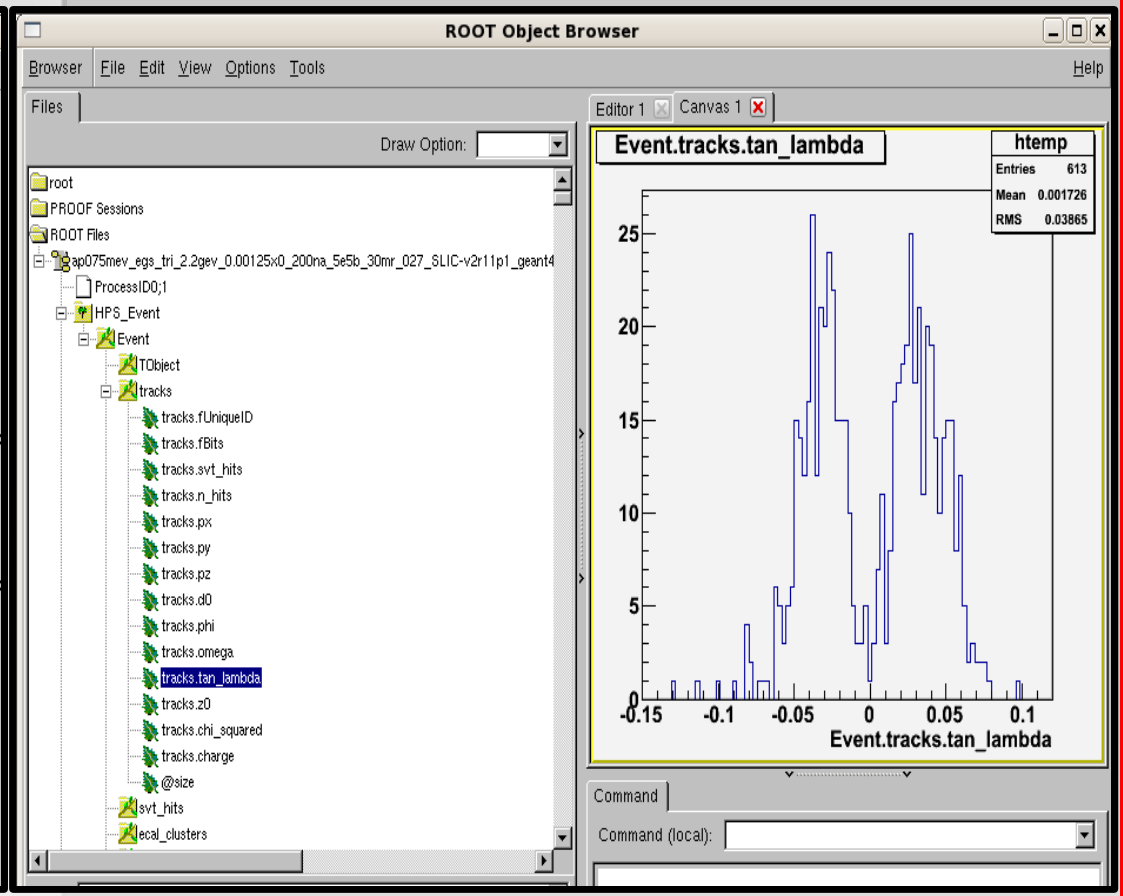
Summary

- Resources available at JLAB and SLAC for the MDC
- A minor amount of script work needed to auto transfer the slcio and DST files and update the HPS catalog
- DST will be available in ROOT format and are easy to use. See examples provided by Omar to get started.

Browsing the HPS DST

Using the TBrowser command from ROOT to see the structure:





Scanning values in the HPS DST

```
root [31] HPS_Event->Scan
("Event.tracks.z0:Event.tracks.tan_lambda","Event.tracks.z0>10.");
************************************************
*    Row   * Instance * Event.tra * Event.tra *
************************************************
*        8 *        0 * 10.089396 * -0.033867 *
*       29 *        0 * 23.519870 * -0.100035 *
*       50 *        0 * 11.912596 * 0.0091170 *
*       52 *        4 * 12.507448 * -0.027359 *
*       52 *        5 * 16.199352 * -0.045461 *
*       61 *        0 * 30.311225 * -0.129605 *
*       65 *       11 * 10.305232 * -0.028102 *
*       67 *        1 * 22.503559 * -0.074096 *
*       75 *        1 * 10.275684 * -0.033444 *
*       79 *        3 * 12.408655 * -0.052596 *
*       79 *        4 * 17.865940 * -0.047133 *
*       88 *        2 * 13.836064 * -0.053875 *
*       88 *        3 * 11.295079 * -0.081828 *
*       92 *        0 * 10.116567 * -0.059614 *
*      113 *        1 * 10.647219 * 0.0365686 *
*      118 *        0 * 10.838727 * -0.081098 *
*      134 *        0 * 31.822881 * -0.112879 *
*      140 *        4 * 17.993574 * -0.079575 *
*      140 *        6 * 16.415493 * -0.053416 *
*      158 *        4 * 10.280500 * -0.043641 *
*      168 *        5 * 14.319632 * -0.055593 *
*      170 *        1 * 23.455913 * -0.089834 *
*      185 *        5 * 25.864082 * -0.061572 *
*      188 *        8 * 11.057923 * -0.021803 *
*      192 *        0 * 24.503961 * -0.039328 *
```