

HPS Data Management

Presented by Homer NEAL for the




Collaboration Meeting





January 23rd, 2014

Data Distribution

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



- Simulation +
- Reconstruction +
- Tape storage



clashps:: prod. account on ifarm

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

Re
con
(SLCIO)

D
S
T
(ROOT)

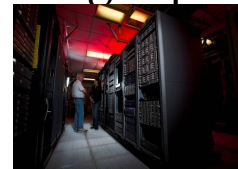
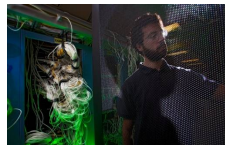
13 TB for the
Mock
Data Challenge

DSTs are produced as part of the
reconstruction process.

Transfer machine
bbr-xfer14

30 TB

/nfs/slac/g/hps3



JLAB HPS Production Spaces

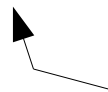
- /u/**home**/clashps has 2GB, backed up
- /**work**/clas/clashps: 1TB of permanent (but not backed up)
- /**volatile**/clas/clashps: 1TB for staging
 - Note: This is 1TB reserved and 5 TB quota

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

- /u/group/hps has 20GB (backed up) and is visible from **ifarm** and **jlabl** 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,ngra  
f,mgraham,homer,sarahp,omoreno,tvm,stepanya,boi  
arino,holtrop,hovanes,ungaro,uemura,mskolnik,gcha  
rles
```

If you will be running HPS production at JLAB then you will need to contact FX to be added to 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).
- 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:

***HPS
Data Catalog***

- 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

Version: 1.10
User: homer . (Switch|Logout)
Project: CDM5 | CTA | EXO | SID | HPS | LSST-CAMERA | LSST-DESC | SCA | SRS | SSRL
Mode: [Prod | Dev]
View: [Tree . Data Types . File Formats . Messages . Admin . Problems]

One item found.

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

HPS Data Catalog - HPS Data Catalog - Mozilla Firefox

Version: 1.10
User: homer . (Switch|Logout)
Project: CDM5 | CTA | EXO | SID | HPS | LSST-CAMERA | LSST-DESC | SCA | SRS | SSRL
Mode: [Prod | Dev]
View: [Tree . Data Types . File Formats . Messages . Admin . Problems]

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

Applications Raccourcis Système 0 °C 16:59

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 Search...

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)
<input type="checkbox"/>	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

Fire... Gno... HPS ... Dém...
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 Search...

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

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 screenshot shows the HPS Data Catalog web interface in Mozilla Firefox. The page title is "HPS Data Catalog". The user is logged in as "homer". The project is "CDMS | CTA | EXO | SID | HPS | LSST-CAMERA | LSST-DESC | SCA | SRS | SSRL". The mode is "Prod | Dev". The view is "Tree". The left sidebar shows a folder tree with "demo" selected. The main content area shows a table of 6 items found, displaying all items. The table has columns: Name, Type, Format, Run Min, Run Max, Events, Size, Status, and Created (UTC). The items are:

Name	Type	Format	Run Min	Run Max	Events	Size	Status	Created (UTC)
hps_001362_evio_2_recon	SLCIO	slcio	0	0	0	0 B	UNSCANNED	
hps_001362_evio_1_recon	SLCIO	slcio	0	0	0	0 B	UNSCANNED	
hps_001362_evio_0_recon	SLCIO	slcio	0	0	0	8.6 GB	OK	18-Apr-2013 21:42:27
hps_001352_recon	SLCIO	slcio	0	0	0	0 B	MISSING	01-Jan-1970 00:00:00
hps_001351_recon	SLCIO	slcio	0	0	0	0 B	MISSING	01-Jan-1970 00:00:00
hps_001351_evio_0_recon	SLCIO	slcio	0	0	0	5.5 GB	OK	18-Apr-2013 22:05:11

The screenshot shows the HPS Data Catalog web interface in Mozilla Firefox, displaying a detailed view of a folder. The folder is "demo/DST/simulation Group HPS". The user is logged in as "homer". The project is "CDMS | CTA | EXO | SID | HPS | LSST-CAMERA | LSST-DESC | SCA | SRS | SSRL". The mode is "Prod | Dev". The view is "Tree". The left sidebar shows a folder tree with "demo" selected. The main content area shows a table of 4 items found, displaying all items. The table has columns: Name, Type, Format, Run Min, Run Max, Events, Size, Status, and Created (UTC). The items are:

Name	Type	Format	Run Min	Run Max	Events	Size	Status	Created (UTC)
hps_001362_evio_2_recon	SLCIO	slcio	0	0	0	7.6 GB	OK	18-Apr-2013 21:32:44
hps_001362_evio_1_recon	SLCIO	slcio	0	0	0	8.6 GB	OK	18-Apr-2013 21:35:48
hps_001362_evio_0_recon	SLCIO	slcio	0	0	0	8.6 GB	OK	18-Apr-2013 21:42:27
hps_001351_evio_0_recon	SLCIO	slcio	0	0	0	5.5 GB	OK	18-Apr-2013 22:05:11

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

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 Search...

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](#)]

6 items found, displaying all items.

1

	Name	Type	Format	Run Min	Run Max	Events	Size	Status	Created (UTC)
<input type="checkbox"/>	hps_001362_evio_2_recon	SLCIO	slcio	0	0	0	0 B	UNSCANNED	
<input type="checkbox"/>	hps_001362_evio_1_recon	SLCIO	slcio	0	0	0	0 B	UNSCANNED	
<input type="checkbox"/>	hps_001362_evio_0_recon	SLCIO	slcio	0	0	0	8.6 GB	OK	18-Apr-2013 21:42:27
<input type="checkbox"/>	hps_001352_recon	SLCIO	slcio	0	0	0	0 B	MISSING	01-Jan-1970 00:00:00
<input type="checkbox"/>	hps_001351_recon	SLCIO	slcio	0	0	0	0 B	MISSING	01-Jan-1970 00:00:00
<input type="checkbox"/>	hps_001351_evio_0_recon	SLCIO	slcio	0	0	0	5.5 GB	OK	18-Apr-2013 22:05:11

Select all Deselect all Toggle Selection Download Selected Files

I had just inserted these entries.

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](#)]

Folder /demo/DST/simulation Group HPS

4 items found, displaying all items.

1

	Name	Type	Format	Run Min	Run Max	Events	Size	Status	Created (UTC)
<input type="checkbox"/>	hps_001362_evio_2_recon	SLCIO	slcio	0	0	0	7.6 GB	OK	18-Apr-2013 21:32:44
<input type="checkbox"/>	hps_001362_evio_1_recon	SLCIO	slcio	0	0	0	8.6 GB	OK	18-Apr-2013 21:35:48
<input type="checkbox"/>	hps_001362_evio_0_recon	SLCIO	slcio	0	0	0	8.6 GB	OK	18-Apr-2013 21:42:27
<input type="checkbox"/>	hps_001351_evio_0_recon	SLCIO	slcio	0	0	0	5.5 GB	OK	18-Apr-2013 22:05:11

Select all Deselect all Toggle Selection Download Selected Files

List only OK files.

← files found

HPS Data Catalog File Listing

The screenshot shows a Mozilla Firefox browser window displaying the HPS Data Catalog. The address bar shows the URL `srs.slac.stanford.edu/DataCatalog/?experiment=HPS`. The page title is "HPS Data Catalog - HPS Data Catalog - Mozilla Firefox". The main content area features the "HPS Data Catalog" logo on the left and a sidebar on the right. The sidebar lists folders: CTA, DetMC, DetMC_data, ILC, demo, DST, simulation, and HPS. The main content area displays the "Folder /demo/DST/simulation Group HPS" and lists several files with their full paths, including `/nfs/slac/g/hps3/data/testrun/runs/dcat/hps_001351_evio_0_recon.slcio` and `/nfs/slac/g/hps3/data/testrun/runs/dcat/hps_001362_evio_2_recon.slcio`. The top right of the page shows the version (1.10), user (homer), and various project links (CDMS, CTA, EXO, SID, HPS, LSST-CAMERA, LSST-DESC, SCA, SRS, SSRL). The bottom of the browser window shows the taskbar with open applications like Firefox, Gnome, and HPS.

Applications Raccourcis Système -1 °C 17:58

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

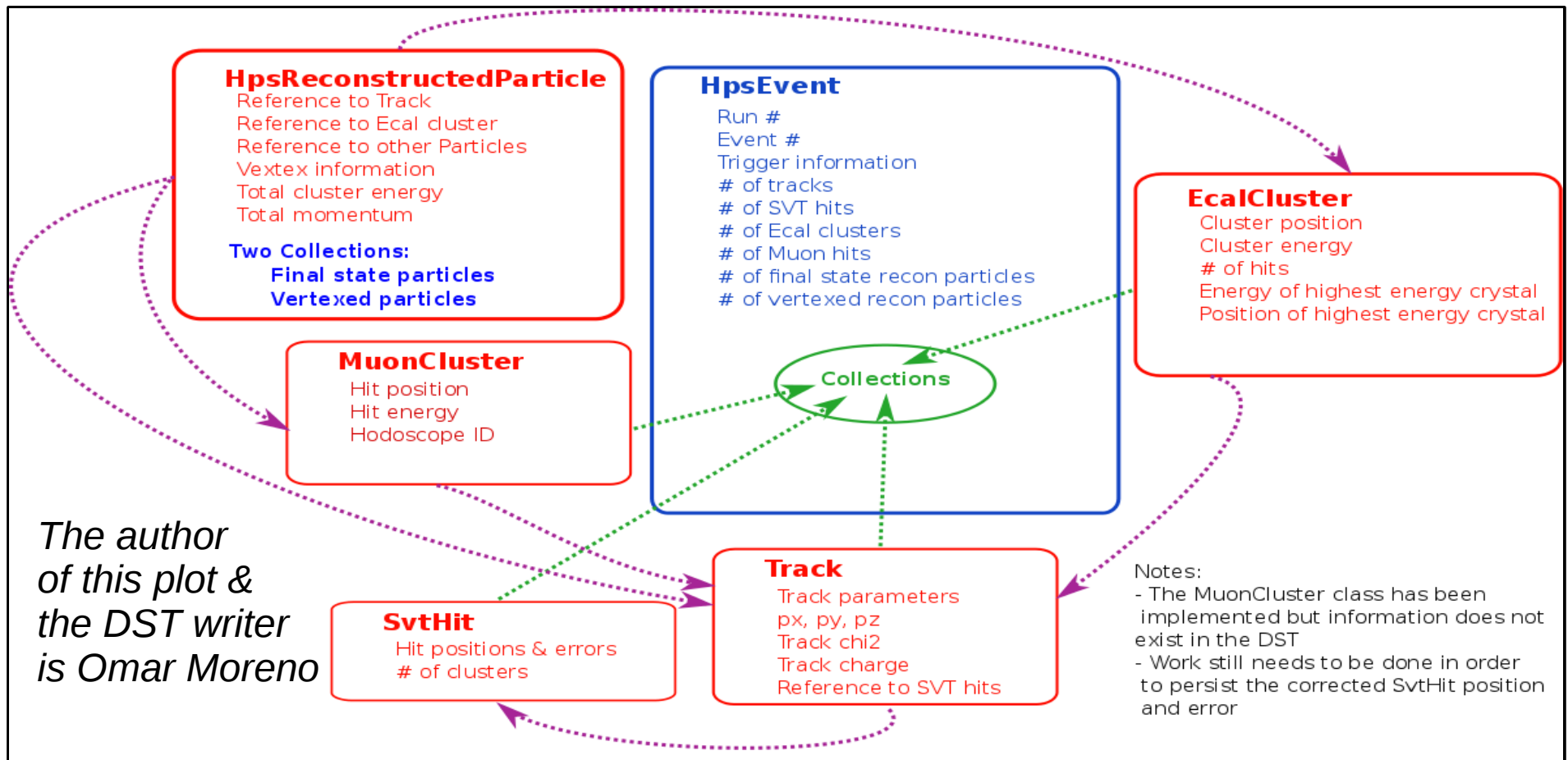
Folder /demo/DST/simulation Group HPS

- /nfs/slac/g/hps3/data/testrun/runs/dcat/hps_001351_evio_0_recon.slcio
- /nfs/slac/g/hps3/data/testrun/runs/recon_new/hps_001351_recon.slcio
- /nfs/slac/g/hps3/data/testrun/runs/recon_new/hps_001352_recon.slcio
- /nfs/slac/g/hps3/data/testrun/runs/dcat/hps_001362_evio_0_recon.slcio
- /nfs/slac/g/hps3/data/testrun/runs/dcat/hps_001362_evio_1_recon.slcio
- /nfs/slac/g/hps3/data/testrun/runs/dcat/hps_001362_evio_2_recon.slcio

FoxyProxy: Disabled

Firef... Gno... HPS ... Dém...

The HPS DST Format

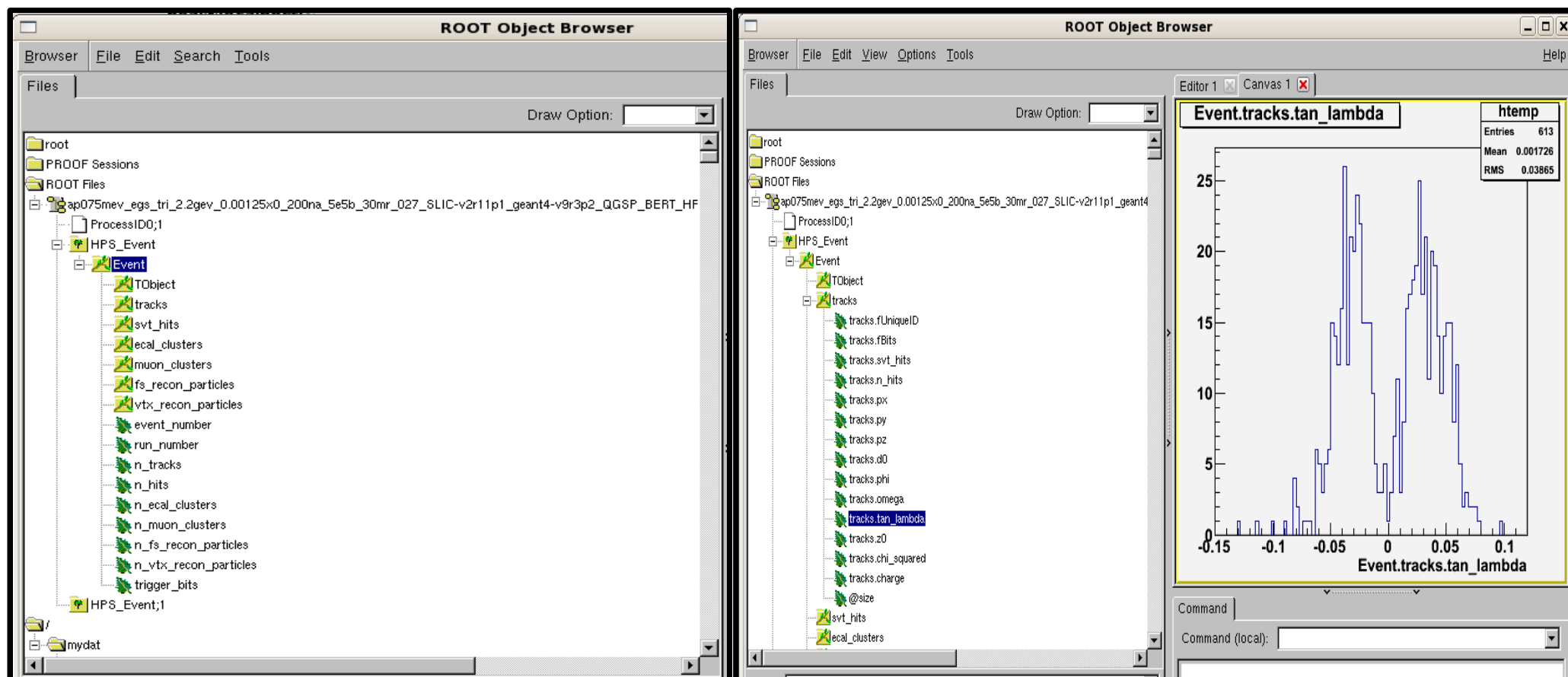


For a guide on how to use it see Omar's page at:
<https://confluence.slac.stanford.edu/display/hpsg/HPS+Data+Summary+Tapes>

The following are some interactive examples ...

Browsing the HPS DST

Using the TBrowser command from ROOT to see the structure:



Scanning values in the HPS DST

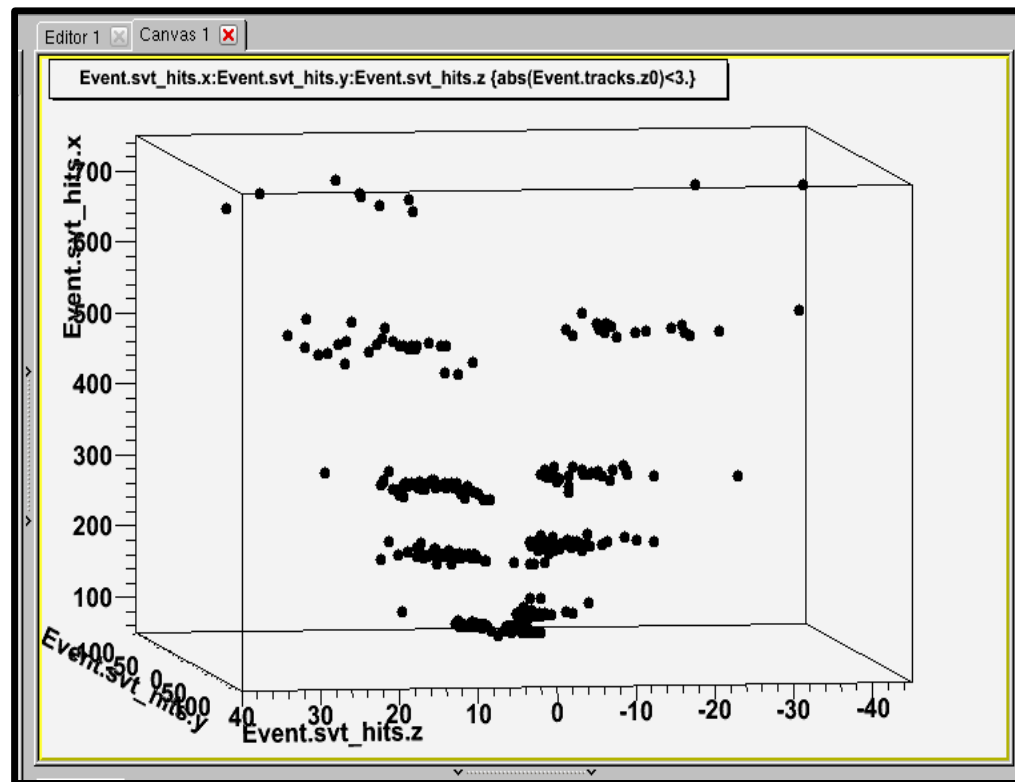
```

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 *

```

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

- See basic analysis example accessible from:
<https://confluence.slac.stanford.edu/display/hpsg/HPS+Data+Summary+Tapes>
- This is a good starting point for your analysis.

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.



Data Sitting in the Stage Area

```
-bash-3.2$ date;du /hps/data/testrun/runs/  
Thu Jan 23 08:55:19 PST 2014  
1555015833    /hps/data/testrun/runs/evio  
199153049     /hps/data/testrun/runs/slci0  
1568758151    /hps/data/testrun/runs/recon  
3322927036    /hps/data/testrun/runs
```