High Power Computing

Accounts

- All PSI accounts use the same unified logon information
  - Third Party Authentication (Balrog)
  - Virtual Private Network (VPN)
  - Linux Login Cluster (LLC)
  - High Performance Computing (Merlin4)
  - File Archives (archivftp.psi.ch)

Accounts

- All PSI accounts use the same unified logon information
  - Third Party Authentication (Balrog)
  - Virtual Private Network (VPN)
  - Linux Login Cluster (LLC)
  - High Performance Computing (Merlin4)
  - File Archives (archivftp.psi.ch)

Documentation

- Merlin4 uses the Sun Grid Engine, with documentation widely available around the internet
- PSI has documentation available on the intranet
  - https://intranet.psi.ch/PSI_HPC/LocalHpcUserGuide
  - https://intranet.psi.ch/PSI_HPC/SGEUserGue

Accessing Intranet

- Request VPN account
  - To: helpdesk@psi.ch
  - CC: Stefan Ritt (stefan.ritt@psi.ch)
- Download client and config file for your system
  - http://www.psi.ch/computing/vpn-infos
  - Gateway, Group ID, Group Secret Phrase

Merlin4

- Two login clusters
  - merlinl01.psi.ch
  - merlinl02.psi.ch
- ~30 other non-login clusters
  - merlinc01.psi.ch
  - merlinc02.psi.ch
  - ...
  - merlinc30.psi.ch

Initial Setup

- Connect to Merlin4 login node
- Enter the following commands according to PSI IT documentation:

```
chmod 0700 ~/.ssh
ssh-keygen -t rsa
# press Enter to accept the default settings
cp ~/.ssh/id_rsa.pub ~/.ssh/authorized_keys
chmod 0600 ~/.ssh/authorized_keys
```

To check everything is okay:

```
for i in `seq -w 1 30` ; do
  h=merlinc$i
  ssh -o ConnectTimeout=10 $h hostname
done
```

Aside: Things I've learned about SSH

SSH Config File
(~/.ssh/config)

```
Host balrog
     HostName balrog.psi.ch
     User quirk_j

Host llc
     HostName llc.psi.ch
     User quirk_j

Host merlin
     HostName merlinl01.psi.ch
     User quirk_j

Host merlinl01
     HostName merlinl01.psi.ch
     User quirk_j

Host merlinl02
     HostName merlinl02.psi.ch
     User quirk_j

Host stallion
     HostName stallion.psi.ch
     User jrquirk

Host abner
     HostName abner.psi.ch
     User jrquirk

Host nedm
     HostName nedm-cluster.bu.edu
     User jrquirk
```

+

Private/Public key pair (ssh-keygen, ~/.ssh/id_rsa and ~/.ssh/id_rsa.pub)

Public key copied into remote ~/.ssh/authorized_keys file

=

Immediate login with:

- SSH
- SCP
- SFTP
- emacs

Just typed merlin!

```
jrquirk@planewave:~/alcap/data/hist$ ssh merlin
Last login: Wed Mar 19 17:50:29 2014 from wireless1x-155-41-88-129.bu.edu
Merlin4 Login Node running SL  is open for authorized users.
[quirk_j@merlinl01 ~]$
```

No password prompt!

Relevant Commands

- qstat: list jobs you've submitted
- qsub: submit jobs
- qdel: delete running job
- /usr/lpp/mmfs/bin/mmlsquota merliny [--block-size auto]: Check status of 300 GB quota
- module: Controls environment variables

qstat

```
[quirk_j@merlinl01 ~]$ qstat
job-ID  prior   name       user         state submit/start at     queue                          slots ja-task-ID
-----------------------------------------------------------------------------------------------------------------
 912011 0.60000 batch_alca quirk_j      qw    03/19/2014 18:56:15                                    1
 912012 0.55000 batch_alca quirk_j      qw    03/19/2014 18:57:34                                    1
 912013 0.53333 batch_alca quirk_j      qw    03/19/2014 18:58:24                                    1
 912015 0.52500 batch_alca quirk_j      qw    03/19/2014 18:59:54                                    1
 912016 0.52000 batch_alca quirk_j      qw    03/19/2014 19:00:43                                    1
 912017 0.51667 batch_alca quirk_j      qw    03/19/2014 19:02:03                                    1
 912018 0.51429 batch_alca quirk_j      qw    03/19/2014 19:02:58                                    1
 912019 0.51250 batch_alca quirk_j      qw    03/19/2014 19:03:09                                    1
 912020 0.51111 batch_alca quirk_j      qw    03/19/2014 19:04:46                                    1
[quirk_j@merlinl01 ~]$
```

Job names are just the name of the submitted SGE script

q: not important for us at this point?
w: waiting to be run
r: running

qsub [options] *cmd cmd_args*

- -e/-o *filename*: Redirect standard error/output to these files. By default is *~/job_name*.e###### and *~/job_name*.o######
- -v *var1*,*var2*,...: Variables to be sent to environment job will be run in (like MIDAS_HOME or MIDAS_SYS)
- -N *job_name*: Name of job, default is executable's name

  Example:

  ```
  $ qsub -v DAQdir,MIDAS_HOME,MIDAS_SYS -N alcapana.run02800 -e
    alcapana.run02800.err -o alcapana.run02800.out
    $DAQdir/analyzer/batch/scripts/batch_alcapana.sge 2800
  ```

qsub [options] *cmd cmd_args*

- -e/-o *filename*: Redirect standard error/output to these files. By default is *~/job_name*.e###### and *~/job_name*.o######
- -v *var1*,*var2*,...: Variables to be sent to environment job will be run in (like MIDAS_HOME or MIDAS_SYS)
- -N *job_name*: Name of job, default is executable's name

  Example:

  ```
  $ qsub -v DAQdir,MIDAS_HOME,MIDAS_SYS -N alcapana.run02800 -e
    alcapana.run02800.err -o alcapana.run02800.out
    $DAQdir/analyzer/batch/scripts/batch_alcapana.sge 2800
  ```

Extension sge I think is just bookkeeping for user

qdel *job_id*

- Deletes a running job

```
[quirk_j@merlinl01 ~]$ qstat
job-ID  prior   name       user         state submit/start at     queue                          slots ja-task-ID
-----------------------------------------------------------------------------------------------------------------
 912057 0.00000 batch_alca quirk_j      qw    03/19/2014 23:02:55                                    1
[quirk_j@merlinl01 ~]$ qdel 912057
qstquirk_j has registered the job 912057 for deletion
[quirk_j@merlinl01 ~]$ qstat
[quirk_j@merlinl01 ~]$
```

qdel *job_id*

- Deletes a running job

```
[quirk_j@merlinl01 ~]$ qstat
job-ID  prior   name       user         state submit/start at     queue                          slots ja-task-ID
-----------------------------------------------------------------------------------------------------------------
 912057 0.00000 batch_alca quirk_j      qw    03/19/2014 23:02:55                                    1
[quirk_j@merlinl01 ~]$ qdel 912057
qstquirk_j has registered the job 912057 for deletion
[quirk_j@merlinl01 ~]$ qstat
[quirk_j@merlinl01 ~]$
```

module *cmd [cmd_args]*

- add *module_name* Setup environment variables so that you can use the program/package *module_name*
- rm *module_name* Reverse the above
- avail: list available modules
- list: list modules you've loaded

```
[quirk_j@merlinl01 ~]$ echo $ROOTSYS; echo $PATH

/gpfs/home/gridengine/sge6.2u5p2/bin/lx26-amd64:/usr/lib64/qt-3.3/bin:/usr/local/bin:/bin:/usr/bin:/usr/local/sbin:/usr/sbin:/sbin:/gpfs/home/quirk_j/bin
[quirk_j@merlinl01 ~]$ module list
No Modulefiles Currently Loaded.
[quirk_j@merlinl01 ~]$ module add root
[quirk_j@merlinl01 ~]$ module list
Currently Loaded Modulefiles:
  1) root/root-5.34
[quirk_j@merlinl01 ~]$ echo $ROOTSYS; echo $PATH
/opt/root/root-5.34
/opt/root/root-5.34/bin:/gpfs/home/gridengine/sge6.2u5p2/bin/lx26-amd64:/usr/lib64/qt-3.3/bin:/usr/local/bin:/bin:/usr/bin:/usr/local/sbin:/usr/sbin:/sbin:/gpfs/
[quirk_j@merlinl01 ~]$ module rm root
[quirk_j@merlinl01 ~]$ echo $ROOTSYS; echo $PATH

/gpfs/home/gridengine/sge6.2u5p2/bin/lx26-amd64:/usr/lib64/qt-3.3/bin:/usr/local/bin:/bin:/usr/bin:/usr/local/sbin:/usr/sbin:/sbin:/gpfs/home/quirk_j/bin
[quirk_j@merlinl01 ~]$ module list
No Modulefiles Currently Loaded.
[quirk_j@merlinl01 ~]$
```

Variables were set

module *cmd [cmd_args]*

- add *module_name* Setup environment variables so that you can use the program/package *module_name*
- rm *module_name* Reverse the above
- avail: list available modules
- list: list modules you've loaded

Only modules that seem useful
- root
- python versions
- boost
- gcc version

```
[quirk_j@merlinl01 ~]$ echo $ROOTSYS; echo $PATH

/gpfs/home/gridengine/sge6.2u5p2/bin/lx26-amd64:/usr/lib64/qt-3.3/bin:/usr/local/bin:/bin:/usr/bin:/usr/local/sbin:/usr/sbin:/sbin:/gpfs/home/quirk_j/bin
[quirk_j@merlinl01 ~]$ module list
No Modulefiles Currently Loaded.
[quirk_j@merlinl01 ~]$ module add root
[quirk_j@merlinl01 ~]$ module list
Currently Loaded Modulefiles:
  1) root/root-5.34
[quirk_j@merlinl01 ~]$ echo $ROOTSYS; echo $PATH
/opt/root/root-5.34
/opt/root/root-5.34/bin:/gpfs/home/gridengine/sge6.2u5p2/bin/lx26-amd64:/usr/lib64/qt-3.3/bin:/usr/local/bin:/bin:/usr/bin:/usr/local/sbin:/usr/sbin:/sbin:/gpfs/
[quirk_j@merlinl01 ~]$ module rm root
[quirk_j@merlinl01 ~]$ echo $ROOTSYS; echo $PATH

/gpfs/home/gridengine/sge6.2u5p2/bin/lx26-amd64:/usr/lib64/qt-3.3/bin:/usr/local/bin:/bin:/usr/bin:/usr/local/sbin:/usr/sbin:/sbin:/gpfs/home/quirk_j/bin
[quirk_j@merlinl01 ~]$ module list
No Modulefiles Currently Loaded.
[quirk_j@merlinl01 ~]$
```

Variables were set

/usr/lpp/mmfs/bin/mmlsquota merliny

- Simple, easy to remember command
- --block-size auto displays in user-readable format
- Recommend aliasing it in ~/.bash_profile:

  $ alias quota="/usr/lpp/mmfs/bin/mmlsquota merliny"

  $ alias quotah="/usr/lpp/mmfs/bin/mmlsquota --block-si

```
[quirk_j@merlinl01 ~]$ quota
                     Block Limits                                    |     File Limits
Filesystem type             KB      quota      limit   in_doubt    grace |    files   quota    limit in_doubt    grace  Remarks
merliny    USR      354780640 1048576000 1059061760        192     none |     8114       0        0        0     none hpcddn.psi.ch
[quirk_j@merlinl01 ~]$ quotah
                     Block Limits                                    |     File Limits
Filesystem type         blocks      quota      limit   in_doubt    grace |    files   quota    limit in_doubt    grace  Remarks
merliny    USR         338.3G      1000G      1010G       192K     none |     8114       0        0        0     none hpcddn.psi.ch
```

SGE Scripts

- Comments are like in regular shell scripts , lines beginning with #
- Additionally, qsub command line arguments can be passed inside the script if the line begins with #$

```
#!/bin/bash
# This is a comment
#$ -N job_name
#$ -e log.err
#$ -o log.out
# These options now don't need to be passed at the command line
# The below is apparently required, though some people might
# call the next two line an examples of
# cargo cult programming
source /usr/share/Modules/init/sh
export -n -f module

### THE REST OF THE SCRIPT ###
##############################
```

More examples on merlin at /gpfs/home/sge/examples/merlin4/sge/

Batch Scripts

- analyzer/batch/scripts/batch_alcapana.sh
  - #####: Run number to process
  - -r ##### #####: Range of runs to process
  - -p: FTP password for archivftp.psi.ch; if omitted you will be prompted for it
  - -n: Maximum number of alcapana jobs to process at a time; at this point in time I have no sense of what is a considerate maximum number to submit; default is 1
  - -t: Frequency with which to check if it's okay to submit a new run; default is every 30 seconds (there is a loop where, if the maximum number of jobs is submitted, it pauses for this many seconds)
- analyzer/batch/scripts/batch_alcapana.sge
  - #####: Run number to process
  - The above script submits this to the grid

Batch Scripts

Before starting, set DAQdir

```
[quirk_j@merlinl01 AlcapDAQ]$ echo $DAQdir

[quirk_j@merlinl01 AlcapDAQ]$ . thisdaq.sh
[quirk_j@merlinl01 AlcapDAQ]$ echo $DAQdir
/gpfs/home/quirk_j/AlcapDAQ
[quirk_j@merlinl01 AlcapDAQ]$
```

Batch Scripts

First: Downloads file

```
[quirk_j@merlinl01 scripts]$ ./batch_alcapana.sh -n 10 -t 10 -p [redacted] -r 2800 2820
Downloading run 2800...
--2014-03-19 18:55:06--  ftp://archivftp.psi.ch/mu2e/run2013/run02800.mid
           => “run02800.mid”
Resolving archivftp.psi.ch... 192.33.120.36
Connecting to archivftp.psi.ch|192.33.120.36|:21... connected.
Logging in as mucap ... Logged in!
==> SYST ... done.    ==> PWD ... done.
==> TYPE I ... done.  ==> CWD (1) /archiv/project/mucap/mu2e/run2013 ... done.
==> SIZE run02800.mid ... 294300147
==> PASV ... done.    ==> RETR run02800.mid ... done.
Length: 294300147 (281M) (unauthoritative)

53% [=============================================================================================================>
```

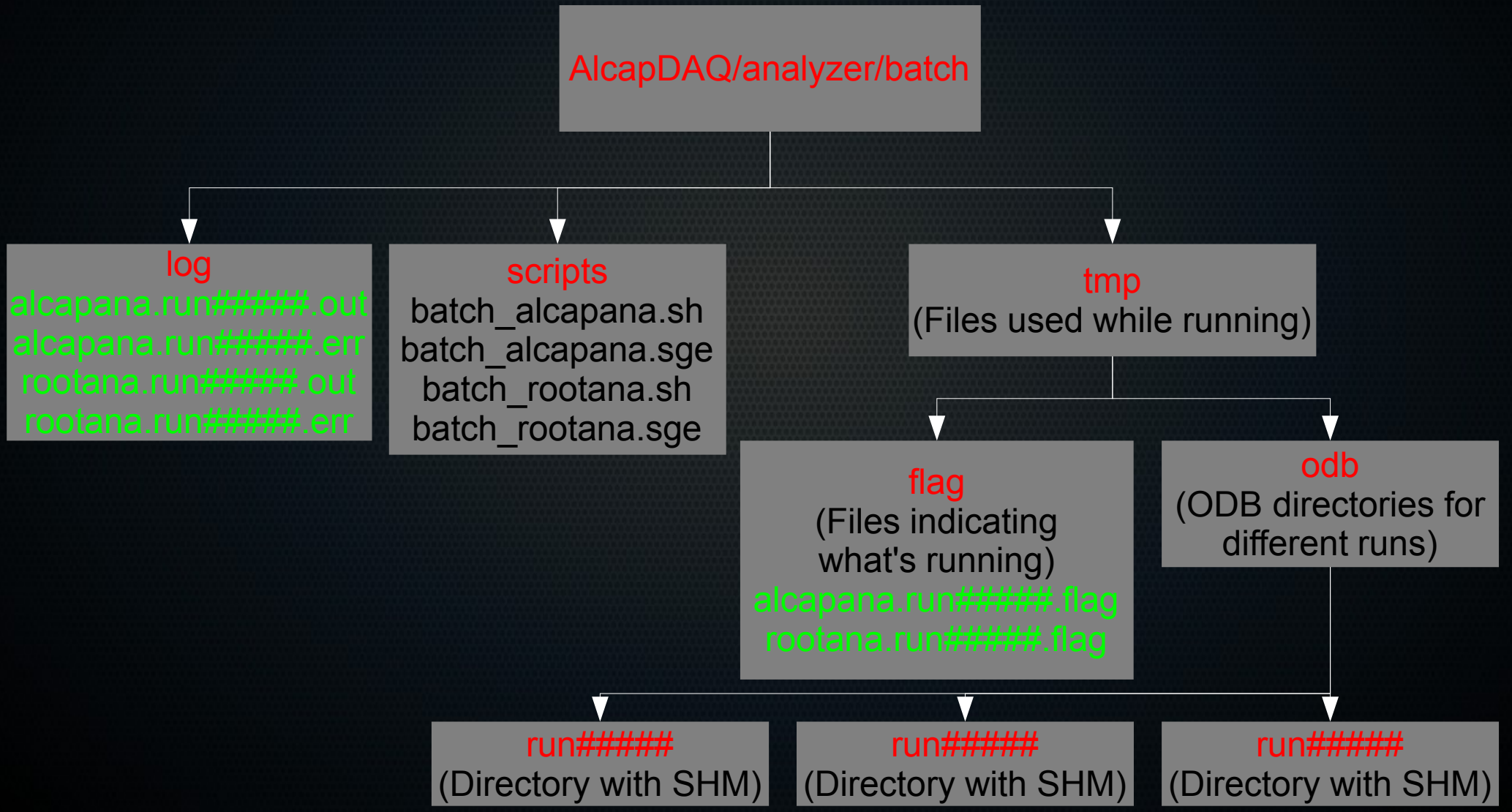
Batch Scripts

Second: Submits job

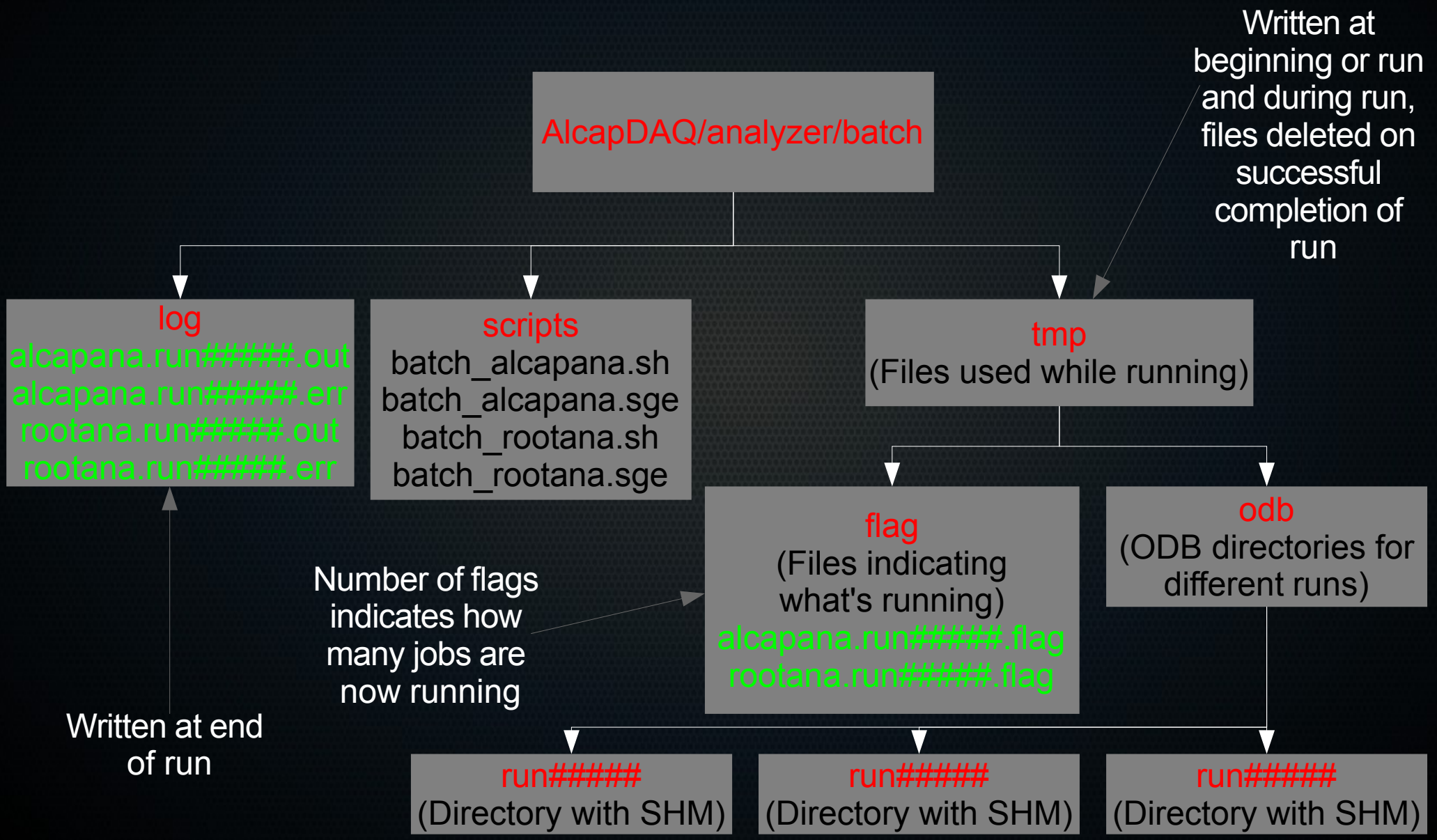
```
2014-03-19 18:56:14 (4.10 MB/s) - “run02800.mid” saved [294300147]

Your job 912011 ("batch_alcapana.sge") has been submitted
Downloading run 2801...
```

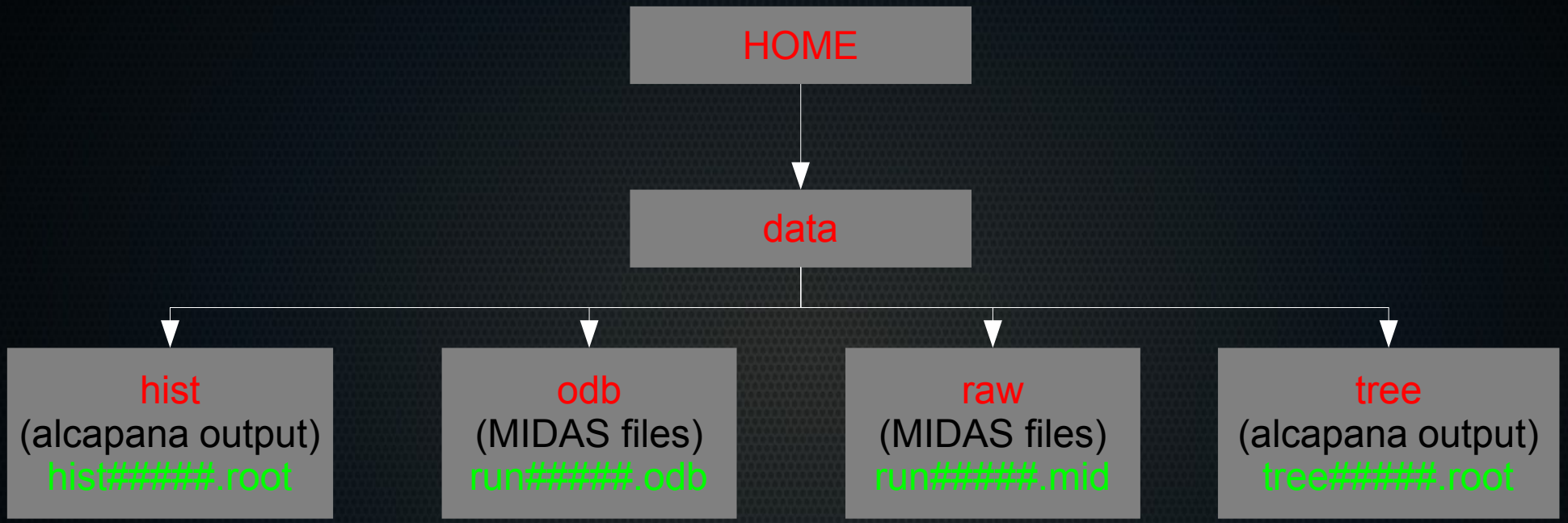
How the SH script works



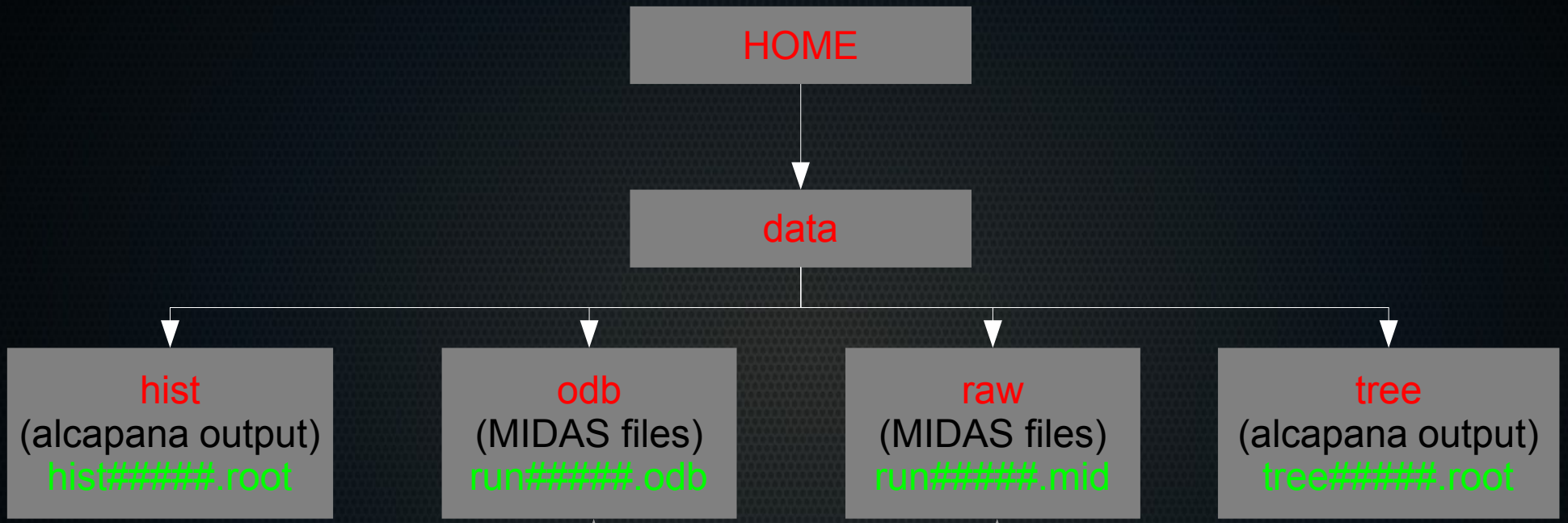
How the SH script works



How the SH script works



How the SH script works



Nam has a copy of all ODBs; not in archivftp; also at the beginning of each run file

Copied over from archivftp before processing with alcapana, deleted after successful processing

How the SH script works

- Requires DAQdir to be set
- Keeps track of how many jobs are running by creating empty flag files: analyzer/batch/tmp/flag/alcapana.run#####.flag
- If there are fewer than the maximum number requested, downloads raw data from PSI archiv to HOME/data/raw/run#####.mid
- Touches new flag file
- Deletes old output and error logs analyzer/batch/log/alcapana.run#####.{err,out} corresponding to this run
- Runs analyzer/work/HOME/alcapana (by invoking batch_alcapana.sge script) on MIDAS file and saves output in HOME/data/tree and HOME/data/hist
- Deletes raw file when done (when flag has been deleted)

How SGE script works

- Requires DAQdir to be set
- Loads ROOT environment (module add root)
- Looks to see if raw file exists in HOME/data/raw; quits if not found
- Looks to see if hist or tree files already exist in HOME/data/hist and HOME/data/tree; quits if found
- Looks to see if the executable analyzer/work/HOME/alcapana exists; quits if not found
- Sets up ODB directory as analyzer/batch/tmp/odb/run#####; this is to prevent multiple runs trying to access the same ODB when they shouldn't be
- Loads all ODB files in analyzer/odb in order ls lists them
- Looks for run specific ODB file in HOME/data/odb
  - If found, loads it
  - If not found, sets ODB key "Analyzer/Load ODB" to "y", which should load the ODB from the raw file itself
- Runs alcapana, removes flag and ODB when finished
- If it takes longer than 30 minutes to run, job is killed

Division of Labor

User	Runs
W. Chen	0000-1962
John Q.	1963-2704
Ed H.	2705-2898
Jim M.	2899-3098
Ben K.	3099-3234
Andy E.	3235-3444
Phil L.	3445-3581
Joe G.	3582-3689
T. H. Nam	3690-3838

Made this list based on who had accounts at the time and the disk space assigned to each user.

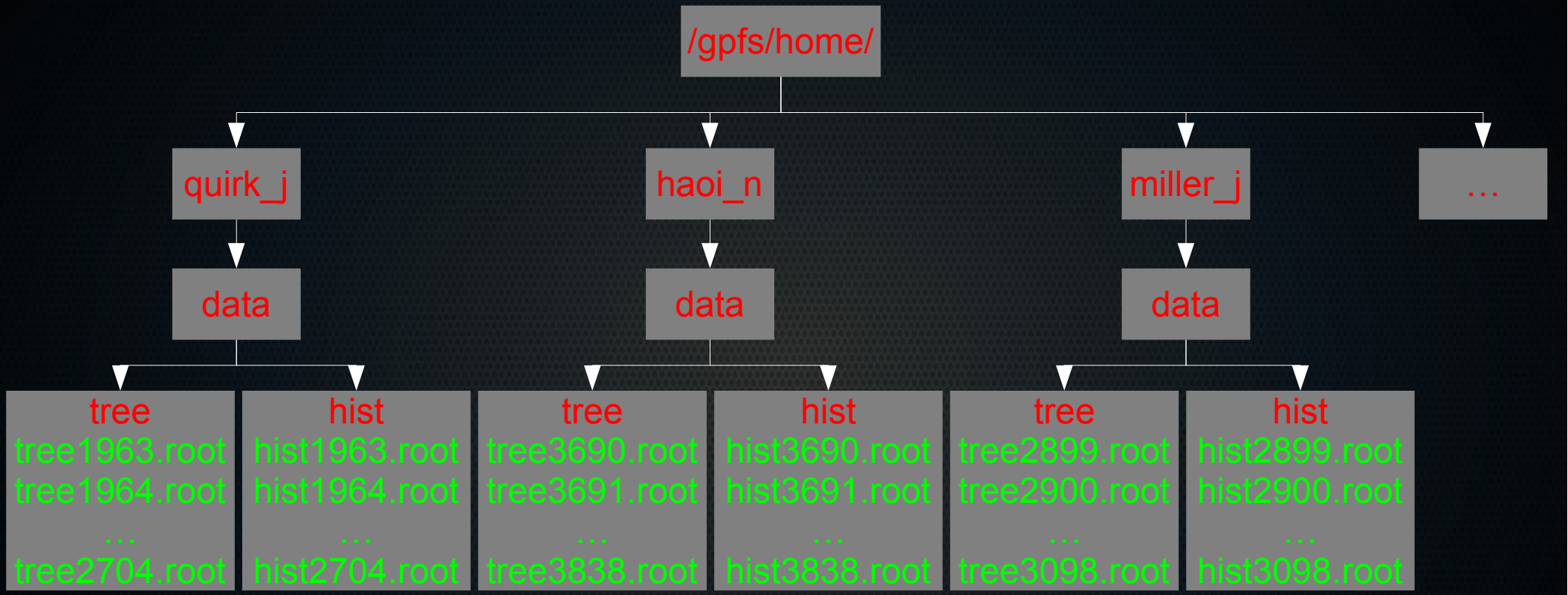
Check missing runs

- Sometimes there are problems
  - Can't download file
  - Error during run

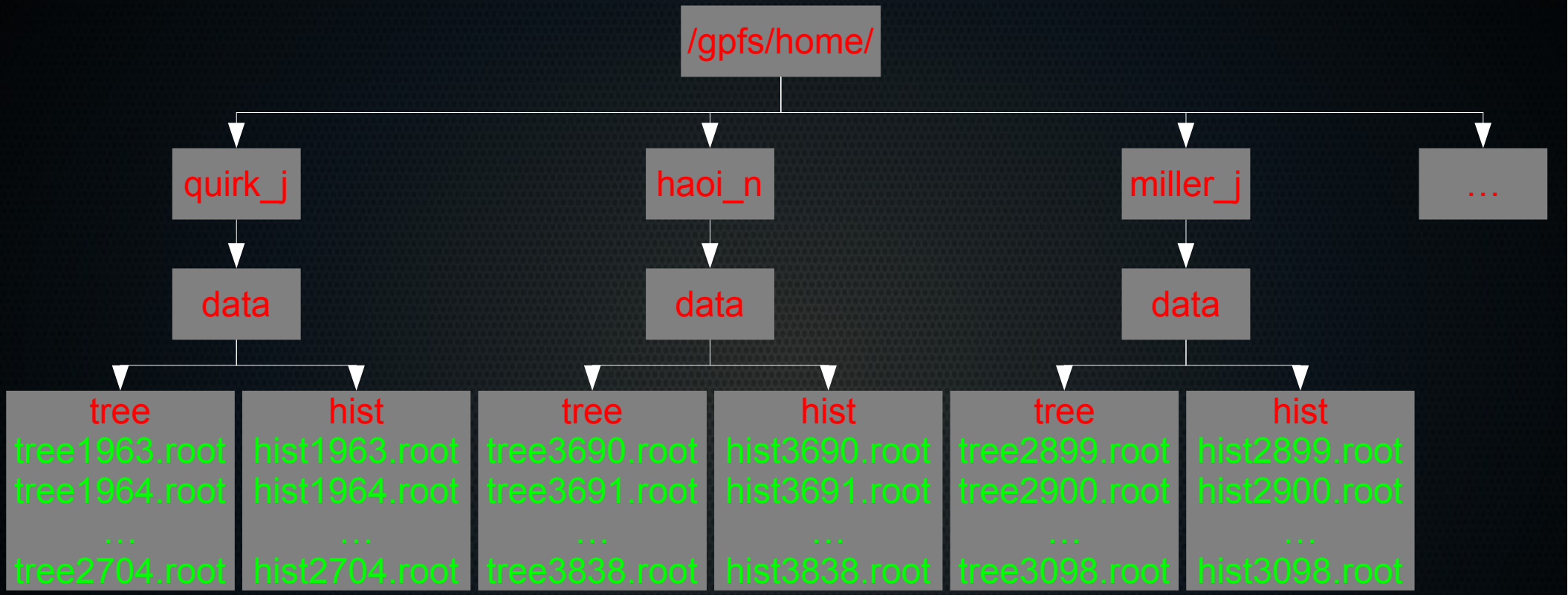
AlcapDAQ/analyzer/batch/scripts/check_missing_runs.sh

```
[quirk_j@merlinl01 scripts]$ ./check_missing_runs.sh 1963 2704
Checking for missing runs in ~/data/tree
/gpfs/home/quirk_j/data/tree/tree02177.root is missing
/gpfs/home/quirk_j/data/tree/tree02211.root is missing
/gpfs/home/quirk_j/data/tree/tree02214.root is missing
/gpfs/home/quirk_j/data/tree/tree02302.root is missing
/gpfs/home/quirk_j/data/tree/tree02303.root is missing
/gpfs/home/quirk_j/data/tree/tree02304.root is missing
/gpfs/home/quirk_j/data/tree/tree02305.root is missing
/gpfs/home/quirk_j/data/tree/tree02306.root is missing
/gpfs/home/quirk_j/data/tree/tree02307.root is missing
```

Data Sharing



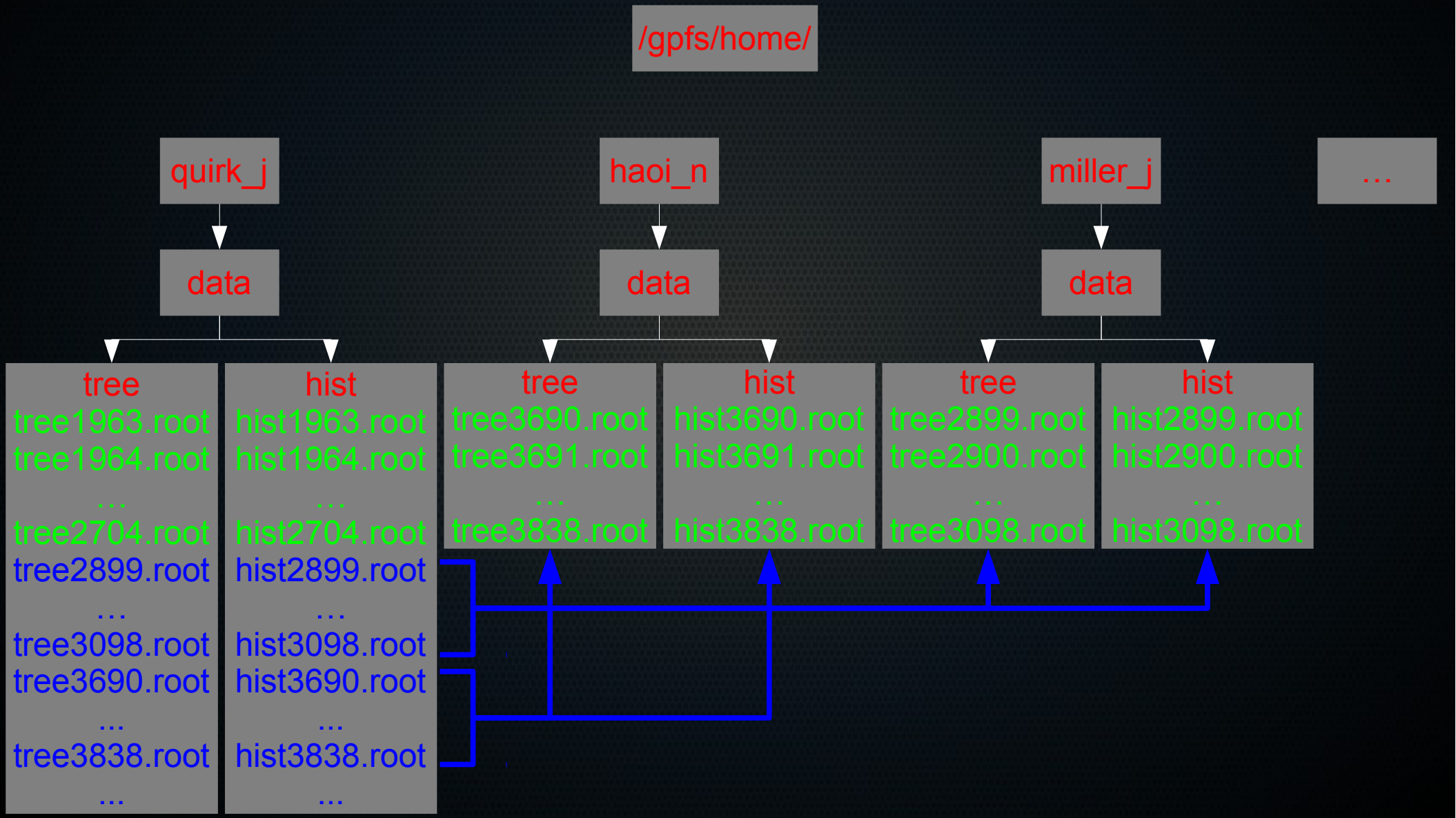
Data Sharing



quirk_j runs AlcapDAQ/analyzer/batch/scripts/check_missing_runs.sh

```
Making link to hoai_n's hist03690.root
Making link to hoai_n's hist03691.root
Making link to hoai_n's hist03692.root
Making link to hoai_n's hist03693.root
Making link to hoai_n's hist03694.root
```

Data Sharing



Data Sharing

- If you and a colleague each have a copy of the same file, a warning is printed

```
Warning: tree03838.root exists in hoai_n's data directory and yours.
Warning: hist03838.root exists in hoai_n's data directory and yours.
```

- If colleague deletes file, removes link

```
Removing broken link (hist03838.root)
```

- We have a group, *unx-alcap*, so we should all have access to each other's data

Data Storage and Sharing

- analyzer/scripts/link_data.sh
  - Checks all of your colleagues' HOME/data/tree and HOME/data/hist directories and creates soft links in your HOME/data/* directories to the files you don't have
  - If you both have a file with the same name, prints warning but doesn't overwrite yours
  - If you have a link to a file of theirs that they have deleted, removes your link
  - We have a group, unx-alcap, so that we can read eachother's data
- Home folder is
  - /gpfs/lastname_firstinitial for Merlin cluster
  - /afs/psi.ch/user/lastinitial/lastname_firstinitial for LLC cluster
  - LLC home folder is accessible from Merlin, Merlin home folder is not accessible from LLC
  - Up until now, HOME has always referred to your Merlin HOME