

ROOT tips

How to check duplicate events in a TTree?

```
root[0] myTree->SetScanField(0);
root[1] myTree->Scan("eventNb:runNb:LS:zVtx:Reco_QQ_4mom.M()"); >tree.log
$ wc -l tree.log
$ sort -u tree.log | wc -l
```

If the tree has duplicated events, both `wc -l` will not give same number of lines.

git commands

- Download files from GitHub servers:

```
$ git cms-init
$ git remote add <repository-name> <HTTPS clone URL>
  $ git remote add CMS-HIN-dilepton git@github.com:CMS-HIN-dilepton/cmssw.git
$ git remote show origin : Try to checkout a remote branch after your local git repo is aware of
$ git remote update
$ git fetch
  $ git fetch <repository-name> : Check branches/tags of the repository
$ git branch : To check which branches are available. Chosen one is starred.
  $ git branch -d <local-branch> : Delete a local branch
$ git checkout <branch-name> : Change between branches
  $ git checkout HIN14015CodeSet/master
$ git checkout -b <localBranchName> <repo-name>/<branch-name>
  $ git checkout -b localBranchTest CMS-HIN-dilepton/onia_b20150311
$ git branch <localNewBranchName>
$ git push <remote-repo-name> <localNewBranchName> : create a new branch and commit to remote repo
  $ git push <REMOTENAME> <LOCALBRANCHNAME>:<REMOTEBRANCHNAME>

$ git cms-addpkg FWCore/Version : Download a directory from the chosen branch

$ git clone git@github.com:<your-user-name>/usercode src/UserCode/<your-name>
$ git clone username@host:/remote/repository/location
$ git clone -b <branch-name> --single-branch <remote-repo-address> : clone 1 branch from remote repo
  $ git clone -b CMSSW_7_5_X_stdMuVal --single-branch https://github.com/MiheeJo/cmssw.git
```

- To clone a repository into local, but to work with sub-directories
 - ◆ This only connects remote repository, not clone them all into local directories

```
$ git init <directory-name>
$ cd <directory-name>
$ git remote add -f <directory-name> <git clone URL (same as HTTPS address)>
  $ git remote add -f CMS-HIN-Dimuons git://github.com/CMS-HIN-dilepton/Dimuons.git
```

- Delete wrong changes/checkouts/commits/branches

```
$ git checkout -- <directory> : Overwrite local files with those from server, local changes will be lost
$ vi .git/info/sparse-checkout : Open this file and delete the directory you don't want to follow
$ git read-tree -mu HEAD : Run this command will actually remove the directory which is deleted from index

$ git reset HEAD~1 : delete the last commit
$ git branch -d [localbranchname] : delete local branch
$ git push [remotename] :[branch] : delete remote branch
```

- Add/modify files

```
$ git add <file-name>
```

```
$ git commit -a
$ git commit -m "leave message"
$ git commit --amend : Merge commits and commit again
```

- If git doesn't recognize changes on files for commit

```
$ git rm --cached <file-name>: This means to only remove the file from the index
$ git reset <file-name>: This tells git to reload the git index from the last commit
```

- Upload changes to git server

```
$ git push <remote-name> <branch-name>
$ git push origin master
$ git push Macros HEAD:master : When there's no branch in remote(=Macros) and push to the m
$ git push <remote-name> <local-branch-name>:<remote-branch-name>
$ git remote add origin <remote-server-address> : When this is the first time to tell the remote
```

- Copy and sync what is in the server to local

```
$ git cms-merge-topic <pull request id>
$ git cms-merge-topic <github-user>:<branch-name>
$ git cms-merge-topic -u CMS-HIN-dilepton:oniaHI_b20150311
```

- Sync forked repo to its original repo link [↗](#)

```
$ git checkout <local-branch-name>
$ git pull https://github.com/ORIGINAL_OWNER/ORIGINAL_REPOSITORY.git <original-repo-branch-name>
$ git commit -m"updating forked repo"
$ git push <remote-repo-name> <local-branch-name>
```

- Set git environment on a new local (reference: <https://help.github.com/articles/generating-ssh-keys/> [↗](#))

```
$ ssh-keygen -t rsa -C "mihee.jo@cern.ch" (first skip, second enter password)
$ eval `ssh-agent -s` : Try "Could not open a connection to your authentication agent." is printed
$ ssh-add ~/.ssh/id_rsa
$ cat ~/.ssh/id_rsa.pub : Copy and paste all contents to github ssh key section
$ ssh -vT git@github.com : Then add ssh key to github account
Move to CMSSW area
$ git cms-init
```

- Compare 2 files in different branches (2 possible ways)

```
$ git diff branch1:file branch2:file
$ git diff branch1 branch2 myfile.cpp
```

Linux shell programs

xargs

- `ls | xargs -n 1 tar zxvf`
- `cmsLs /store/user/miheejo/MC2011/NonPromptJpsi_template/ppReco/Tree/ | awk -v p="" '{if ($5!="") p=p" root://eoscms//eos/cms"$5}; END{print p}' | xargs -n 1 hadd /tmp/miheejo/NonPromptJpsi_template_ppReco_cmssw445p1.root`

rsync

- `rsync -avr -e "ssh -l user" --exclude 'logs_*' ./bit1_weightedEff_* remote:/my/dir`

diff

- `diff -ENwbur <dir1> <dir2>`

: Diff all files in 2 directories

- ◆ `-q` : Print only file names that have differences, `-Ewb` ignore the bulk of whitespace changes, `-N` detect new files, `-u` unified, `-r` recursive

Find

- `find fracfree_fin/___mb_pt/ -type f -name "2D_GG.txt" | wc -l`
- `find . -exec grep -l "searching text" {} \;` : this will print the file name that contains "searching text"
- `find . -name "*.bak" -type f -delete` : Find a file recursively and delete. `-delete` should be at the end of the command otherwise everything will be deleted.
- `find . -type d -print -exec chmod 644 {} \;` : Find a file recursively and change permission.

grep

- `grep -f file1 file2` : Compare files and print common contents. For exact matches (`-w`) without regex (`-F`) options can be used.

AWK

- Extract .py file name from LSFJOB fileFrom this string: `cmsRun /afs/cern.ch/user/d/dmoon/scratch0/Hi_TnP/cms442patch5/src/HiAnalysis/HiOnia/test/Batch/181910_182065/get/HisOnia1_11.py` with below command. Above string is written in a "LSFJOB" file.

```
awk -v p='cmsRun' '{if(p == $1) print $2}' <LSFJOB | awk 'BEGIN{FS="/"}; {print $NF}' | awk
```

- Get file names with regular expression.

```
list=$(rfdirdir $indir | awk '{if (/pT[0-9.]+-[0-9.]+/) print $9}')
```

- Rename files with simple regular expression. This changes old prefix of all files ("rp22_5000") to new one ("rp22").

```
ls | awk -v p="" '{if (index($0,"rp22") != 0) {p=$0; gsub(/rp22_5000/,"rp22"); print "mv " p " rename.sh" }' |
```

- Read multiple files and combine each records into 1 record

```
awk '{a=$0; if (getline <"f1") {print(a $0);}}' <f2
```

- ◆ other example:

```
awk '{ a=$1; b=$2; if (getline <"file2.txt") { printf("%.2f %.2f\n", a/$1, b/$2); } }
```

- Read 2 files and compare contents, print missing lines in 2nd file

- ◆ `awk 'FNR==NR{a[$1]="$2"; next} { if($1 in a){next}else{print $0} }'` dsf
- ◆ Array `a[.]` has same index and content: `$1` `$2`.
- ◆ 'next' is same as continue in c++: Stop processing and go to next record.
- ◆ `FNR` is number of records for each file, `NR` is total records sum of all files
 - ◇ So `FNR==NR{...}` allows `{...}` to be executed only for 1st file

VI

- Reuse matched pattern

\0 is the entire match, \1 is the first part of the matched pattern

- Searching a pattern in a non-greedy way with regular expression

```
^# to match the # character anchored at the start of the line (this answers question 1)
\s\+ to match any whitespace one or more times
\( to start a group (this answers question 2)
.\{-}\ to match any character 0 or more times in a non-greedy way; this is different from .*
\) to end the subgroup.
: matches a literal :
```

Convert

- Merge multiple PDF files into one. But resolution is poorer.

```
convert input1.pdf input2.pdf ... output.pdf
```

- Convert pdf to png :

```
ls | grep pdf | awk 'BEGIN{FS=".pdf"}; {print "convert -gravity center -define pdf:use-crop=0 " $1 ".png"}'
```

pdfnup

- Need the pdftjam package. Merge several pdf files into 1 pdf file, with matrix format.

```
pdfnup --nup [column]x[row] --outfile [output file name] [input file 1] [input file 2] [input file 3]
```

Kerberos and OpenAFS on (k)Ubuntu, Debian

These instructions work as-is on Ubuntu Karmic Koala (version 9.10) and Lucid Lynx (version 10.04). They have been tested on Ubuntu 12.04.2 LTS (Precise Pangolin) and work if you use `apt-get` instead of `aptitude`. This web page [will give up-to-dated information regarding CERN kerberos](#). Original reference of this instruction can be found [here](#). Or check [OpenAFS wiki](#)

- Install the necessary packages.

```
sudo apt-get install openafs-krb5 openafs-client krb5-user module-assistant
openafs-modules-dkms
```

NOTE: The `openafs-modules-dkms` package automatically does the compiling and installation of the `openafs` kernel module. More importantly, it keeps the kernel module up-to-date as software updates upgrade the kernel.

- When asked, supply these answers to the following questions:

AFS cell this workstation belongs to:	cern.ch
Size of AFS cache:	512000 (this means 0.5GB, can be higher if you have the space)
DB server host names for your home cell:	(blank)
Run openafs client now and at boot?	(user preference)

- Unfortunately, at its default resolution, `debconf` won't ask all the necessary questions on the first pass. You'll need to reconfigure. Run:

```
sudo dpkg-reconfigure krb5-config openafs-client
```

- When asked, supply these answers to the following questions:

Default realm:	CERN.CH (all caps)
Does DNS contain pointers to your realm's Kerberos Servers?	Yes
AFS cell this workstation belongs to:	cern.ch
Size of AFS cache:	512000 (this means 0.5GB, can be higher if you have the space)
Run openafs client now and at boot?	(user preference)
Look up AFS cells in DNS?	Yes
Encrypt authenticated traffic with AFS fileserver?	Yes
Dynamically generate the contents of /afs	Yes
Use fakestat?	Yes
DB server host names for your home cell:	(blank)
Run openafs client now and at boot?	(user preference)

- Compile the openafs kernel module, then restart the service. (This step is not necessary on recent versions with the `openafs-modules-dkms` package.)

```
sudo m-a prepare
sudo m-a auto-install openafs
sudo modprobe openafs
```

- Restart service.

```
sudo service openafs-client restart
```

- Things should be working. Test the functionality with:

```
kinit (cern username) && klist
aklog && tokens
```

NOTE: You will have to repeat the kernel module installation whenever the kernel gets upgraded. If, after a reboot, AFS stops working, type the following:

```
ls /lib/modules/`uname -r`/fs/openafs.ko
```

If you get `No such file or directory`, you probably need to rebuild the `openafs` module for the currently-running kernel.

Programming languages

python

- `if not [check for check in ("0.0-1.2", "0.0-2.4", "1.2-1.6", "1.6-2.4") if str(rap) in check] continue`
 - ◆ `in` can not be changed to `is`. It doesn't work.
- `pieces = [p for p in re.split("(|\t|\\\\".)*?\\\\"|'.*?')", line) if p.strip()]`
- for `thiscent` in `centrality`:
 - for `thisrap` in `rapidity`:
 - `open("fit_"+thiscent+"_"+thisrap, 'w').close()`

On lxplus (CAF)

- cmsLs : list file name for in the CAF directories

```
cmsLs /store/caf/user/miheejo/HIExpressPhysics/hiexp-hirun2011-r181530-reco-v1_2/
```

- cmsStage, cmsRm : copy/remove CAF directory files to local directory. Use directory location same as cmsLs.

- cmsMkdir

```
/store/group/phys_heavyions/dileptons/MC2011/NonPromptJpsi_template/regit
```

- Load logical files :

```
root://eoscms//eos/cms/store/hidata/HIRun2011/HIMinBiasUPC/RAW/v1/000/182/296/F6A6C5C7-461
```

- eos chmod 777 [eos directory] : change permission of eos directory

- cmsPfn : will give the location of a root file in the eos area

- cmsLs /store/user/miheejo/MCsample_pythiagun_445patch1/HiOnia2MuMuPAT | awk '{if (\$5!="")

```
: Erase files in a directory totally
```

- To check the quota: eos quota | grep -A 4 "Quota Node: /eos/cms/store/caf/user/" | head -5

CMSSW related

- edmProvDump : check edm root file's contents. Not only the process names but also their contents, name of sub-branches
- edmPluginRefresh
- edmPluginDump : shows available plugins under the cmssw area
- Get step 1,2,3,4 files in certain release:

```
runTheMatrix.py -l 140 -n -e
```

LXPLUS Batch jobs

- Change the queue of submitted jobs

```
bjobs | grep PEND | awk '{print $1}' | xargs -n 1 bmod -q lnw
```

CRAB jobs

Installation

```
$ wget http://cmsdoc.cern.ch/cms/ccs/wm/www/Crab/Docs/CRAB_current.tgz
$ ./configure
```

CRAB2 configuration

- When "establishing gsissh controlpath" takes too much time: `crab -cleanCache`
- Check integrated luminosity on the crab job
 - ◆ After retriving all outputs, get report.

```
$ crab -c [crab job name] -report
```

- ◆ In the res/, json files are created. Run lumiCalc2.py.

```
lumiCalc2.py -i [json file location] overview
```

- To submit CRAB jobs on CAF with multicrab
 - ◆ Main scripts are [here](#)
 - ◆ Need =multicrab.cfg=(should be), =crabCaf.cfg=(loaded from multicrab.cfg and its contents will be overwritten for conflict items)

```
"multicrab.cfg"
[MULTICRAB]
cfg=crabCaf.cfg

[COMMON]
#CMSSW.dbs_url = http://cmsdbsprod.cern.ch/cms_dbs_caf_analysis_01/servlet/DBSServlet
CMSSW.datasetpath=/HIExpressPhysics/HIRun2011-Express-v1/FEVT
CMSSW.total_number_of_lumis      = -1
CMSSW.lumis_per_job              = 10
CMSSW.pset                       = HiTrigAna_data_fromReco_cfg.py
USER.user_remote_dir              = HIExpressPhysics

[hiexp-hirun2011-r181695-reco-v1-collisionEvents_lowerSC_autohlt]
CMSSW.runselection                = 181695

"crabCaf.cfg"
[CRAB]
jobtype      = cmssw
scheduler    = caf
#server_name = caf

[CAF]
queue = cmscaf1nd

[CMSSW]
#dbs_url = http://cmsdbsprod.cern.ch/cms_dbs_caf_analysis_01/servlet/DBSServlet
datasetpath=/HIExpressPhysics/CommissioningHI-Express-v1/FEVT
pset = HiTrigAna_data_fromRaw_cfg.py
total_number_of_lumis = 2
lumis_per_job = 1
output_file = openhlt_data.root

[USER]
#additional_input_files = BSEarlyCollision120909.db
copy_data = 1
storage_element=T2_CH_CAF
```

- Prompt Reco and RAW data transferred to EOS can be accessed by jobs running on LXBATCH queues directly. CRAB can be used to send jobs to the local scheduler similarly to the way it sends jobs to the CAF, please make the following changes to your crab.cfg file like below. Then stageout to T2_CH_CERN or your CASTOR directory as you would normally.

```
[CRAB]
jobtype = cmssw
scheduler = lsf
[LSF]
queue=<choose one of 1nd, 1nh, 1nw, 2nd, 2nw, 8nh, or 8nm depending on the type of job>
```

This topic: [Main > TipsForComputing](#)

Topic revision: r54 - 2017-02-03 - MiheeJO



Copyright &© 2008-2021 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

or Ideas, requests, problems regarding TWiki? use [Discourse](#) or [Send feedback](#)