

Introduction

DPDAnalysis is a simple analysis framework based on ROOT. It can be used to analyze ATLAS D3PD data (ROOT TTree data format).

Package Organization

Package is lightweight. Consists of a core module which forms the basis on which users can develop and maintain their package(s). core provides a way of interfacing multiple TTree objects (found in the D3PD's) and accessing the data. core also provides analysis skeleton. All user modules are derived from DPDAnalysis base class. The user module typically consists of a Begin, ProcessEvent, and End methods. Typically the user may want to book histograms, configure the analysis in Begin. Begin is called once per user-module before event-loop execution.

Event Loop

Event loop is implemented in Process method and called for every event. Since arbitrary number of TTrees are supported, the event loop usually runs over a TTree (or TChain) that has the most number of entries. For all the other TTrees, the event access is synchronized until they run out of entries. Once they do, the last entry is returned for the remainder of the event-loop until processing TTree with most number of entries completes.

Data Preparation

- TOP_NTUP (D3PD's) produced using TopInputsD3PDMaker are slimmed using the slimmer.
- slimmer.py.txt: script to slim D3PD's and write out a smaller ROOT-tuple
- More info about slimming can be found here.
- A list of branches required for top analysis is in PhysicsBranches.txt
- Here's an example script to run the slimmer on grid. Modify according to requirements.

```
prun --athenaTag=16.0.1 --exec "./slimmer.py %IN -f PhysicsBranches.txt -o myTopD4PD.root" \
--inDS user.wbell.mc09_7TeV.108343.st_schan_enu_McAtNlo_Jimmy.merge.NTUP_TOP.e534_s765_s767_r1302
--outDS user.venkat.mc09_7TeV.108343.st_schan_enu_McAtNlo_Jimmy.D4PD.2010.11.16.v01 --nGBPerJob=6
--outputs "my*.root"
```



- Note: test locally before submitting the jobs. I have a sample file to test (on pcazNN @ CERN)

```
/raid01/venkat/dataset/user.wbell.mc09_7TeV.108346.st_Wt_McAtNlo_Jimmy.merge.NTUP_TOP.e624_s765_s
```

- To run on this file, use the following commands `ln -s /raid01/venkat/dataset/user.wbell.mc09_7TeV.108346.st_Wt_McAtNlo_Jimmy.merge.NTUP_TOP.e624 ./slimmer.py user.wbell.003248.EXT0._00003.NTUP_TOP.root -f PhysicsBranches.txt -o output.root`

Getting Started

Checkout packages from SVN repository

  You may use the checkout.sh script to checkout and compile the necessary packages from SVN OR

- You can setup a suitable ATLAS software release/cache (on a SLC5, x64 machine)
- setup ATLAS release in the usual way


```
cd ~/cmthome/
source setup -tag=15.6.12.4,AtlasProduction,slc5,gcc43,mydpd
```


Sample requirements file for setting up and compiling DPDAnalysis

- Check out packages from SVN repository


```
export SVNROOT= $SVNGRP
cmt co Institutes/Arizona/DPDAnalysis/core
cmt co Institutes/Arizona/DPDAnalysis/util
cmt co Institutes/Arizona/DPDAnalysis/topuser
export SVNROOT= $SVNOFF
cmt co DataQuality/GoodRunsLists
cmt co Reconstruction/Jet/JetUncertainties
```

Sample log-file of SVN checkout using CMT.

 Make sure you have SVN access to checkout the packages. You must have an AFS account for this step.

 If your afs username doesn't match your login name then you have to follow these additional steps

```
export SVNGRP=svn+ssh://yourafsusername@svn.cern.ch/repos/atlasgrp
export SVNOFF=svn+ssh://yourafsusername@svn.cern.ch/repos/atlasoff
```

 If you find it tedious to enter your password when prompted (multiple times) afsuser@svn.cern.ch's password you may want to obtain and cache your credentials using kinit \$USER@CERN.CH and then checkout the above packages.

Compile all packages

- if you are using ATLAS setup, its easy to compile. Just cd into cmt directory in user package and compile

```
cd $TestArea/Institutes/Arizona/DPDAnalysis/topuser/cmt/ ; cmt bro make; cd -
```

Sample log-file of compiling using CMT

Before running the program

- Create a link to JESUncertainty.root  Not required if you are using topuser-00-02-02 or later. Refer to the list tags below

```
cd $TestArea; ln -s Reconstruction/Jet/JetUncertainties/share/JESUncertainty.root
See JESUncertaintyProvider for details
```

- If running on data edit
=\$TestArea/Institutes/Arizona/DPDAnalysis/topuser/config/grl.config=
and modify the following line and specify a full path to the XML good runs list file. This step is not necessary if running on MC.

ArizonaPhysicsAnalysis < Main < TWiki

You can find predefined grl's here [GRL Lists](#)

```
grl.GoodRunsXML: Institutes/Arizona/DPDAnalysis/topuser/config/mygoodruns.xml
```

- If running on data edit
=`$TestArea/Institutes/Arizona/DPDAnalysis/topuser/config/topanalysis.conf`
and modify the following `Set false` for data, and `true` for MC

```
top.IsSimulation: true
```

- Make a file list

```
ls /afs/cern.ch/user/v/venkat/public/TOP/D3PD/TopInputs.D3PD.root >  
$TestArea/mytopd3pd.mc.list
```

- Now you are ready to run locally. See `topanalysis.config` for available options.



Running the program

Locally

```
cd $TestArea  
runDPD.py -f mytopd3pd.mc.list -c  
Institutes/Arizona/DPDAnalysis/topuser/config/topuser.config  
For help on command line options runDPD.py --help
```

On the grid using PANDA run

for prun (Version: 0.2.96 or higher)

-   You may want to use the `scripts/runTop.sh` if you need to submit jobs in a batch (several dataset containers).
- Just copy this script to your test area and type in `./runTop.sh` for help.
- setup your testarea. same area you are using to test topuser locally will suffice for grid submission
- setup panda. see [here](#) for for setup
- check to make sure you are using a recent prun version using: `prun --version`
- compile all packages using `cmt` (must do before job submission)
`cd Institutes/Arizona/DPDAnalysis/topuser/cmt; cmt bro make; cd -;`
- submit job using `prun` (modify your user name and `--inDS` and `--outDS` accordingly)

```
prun --useAthenaPackages \  
--match "*D3PD*root*" --exec "echo %IN | runDPD.py -g -c ./Institutes/Arizona/DPDAnalysis  
--inDS your.input.dataset.container.here/ \  
--outDS user.username.output.dataset.container.here \  
--outputs "my*.root"
```

- for a list of all available options for `prun` you can do `prun --help` and use them with the above shown above.
- see [this link](#) for a finished job.

instructions for prun (< 0.2.96)

You need to setup PANDA. You also need to clean the \$TESTAREA

```
cd $TestArea
find -name $CMTCONFIG | awk '{print "rm -rf " $1}'
this removes the i686 directories
```

```
rm -rf InstallArea
this removes the InstallArea
```

to submit to the grid

```
cp /raid/johns/testarea/DPDAnalysis/15.6.10.4/runGrid.sh .
cp /raid/johns/testarea/DPDAnalysis/15.6.10.4/gridsetup.sh .
at some point, these files should be placed in the core/share area in SVN
```

edit runGrid.sh as you wish

```
./runGrid.sh 21 ANALY_BNL_ATLAS_1
the 21 gets appended to your job name and the ANALY_BNL_ATLAS_1 is the
site
```

Output

A ROOT file containing histograms and ntuple of topological variables.
See output file contents.

Merging output

Merge functionality is provided. Merging histograms and ntuples is accomplished using the following commands. Assuming one wants to merge myoutput.1.root and myoutput.2.root, the following commands are equivalent and yield a merged_output.root. Merging works for arbitrary number of input files. By default, only histograms are merged. See [merge.config](#) for options.

Locally

```
echo myoutput.1.root myoutput.2.root | runDPD.py -m
ls myoutput*.root | runDPD.py -m
ls -cl myoutput*.root > files.list; runDPD.py -m -f files.list
```

On the grid using PANDA run

for prun (Version: 0.2.96 or higher)

```
prun --useAthenaPackages \  
--match "*D3PD*root*" --exec "echo %IN | runDPD.py -m -c ./Institutes/Arizona/DPDAnalysis/core \  
--inDS your.input.dataset.container.here/ \  
--outDS user.username.output.dataset.container.here \  
--outputs "merged*.root"
```

Previous Datasets

- Dataset using pre-V5 TopInputsD3PDMaker are listed in TopObsData (Aug - Sept 2010)
- Dataset for 2010/11 TopMriond (Oct 2010-Feb 2011)

List of Tags

Newest tag always contains latest updates. Please use those unless otherwise specified. The code is in SVN repository. Links below:

- [core](#)
- [util](#)
- [topuser](#)

A list of tags for specific Athena releases is tabulated:

List of tags for 2011 data analysis

Athena	core	util	topuser	Comment
16.6.3.5.1	00-00-11	00-00-04	00-05-04	2011 data analysis for PLHC

List of tags for 2010 data analysis

Athena	core	util	topuser	Comment
16.0.3.8.2	00-00-11	00-00-04	00-05-00	re-processed 2010 data analysis for PLHC
15.6.12.4	00-00-10	00-00-04	00-01-00	use these tags together
15.6.12.7	00-00-11	00-00-04	00-01-03	fix for grid issue on DE, IT sites. Pileup histograms/tables for note III, and fixed highest Pt Sum
15.6.12.7	00-00-11	00-00-04	00-02-00	QCD tool update with e/mu. Full sample validated for v5 D3PD tag only for D3PDs with new branches and branch names
15.6.13.1	00-00-11	00-00-04	00-02-01	Split histograms into pre/post b-tag
15.6.13.3	00-00-11	00-00-04	00-02-03	run analysis with nominal, jesplus, jesminus settings
16.0.1	00-00-11	00-00-04	00-02-05	macros to make data v. mc comparison and topological likelihood
16.0.1	00-00-11	00-00-04	00-02-08	build likelihood discriminant, extract signal fraction using fitter and

				estimate sys/stat uncertainties
16.0.1	00-00-11	00-00-04	00-02-13	added hitfit, included topxsec cuts
16.0.1	00-00-11	00-00-04	00-03-00	new ttres cuts
16.0.3.3	00-00-11	00-00-04	00-03-02	fixed and storing truth info, added additional variables to topotree
16.0.3.3.1	00-00-11	00-00-04	00-04-00	rel 16. mc challenge updates, updated reading of branches from new tree, JERProvider tool
16.0.3.3.3	00-00-11	00-00-04	00-04-01	rel 16. lepton sf, multijetJES, MMQCD, scripts & share dirs, new lhood fitter
16.0.3.3.3	00-00-11	00-00-04	00-04-02	rel 16. bug-fix in jes
16.0.3.3.3	00-00-11	00-00-04	00-04-03	added useful scripts
16.0.3.3.3	00-00-11	00-00-04	00-04-07	final update for Pre-Mbriond data analysis

Backward compatibility for older (< v5) D3PD

- With topuser tags \geq 00-02-00, it is possible to use older D3PDs by enabling a switch in the topanalysis.config. By default (top.D3PDVersion: 5). Set it equal to a version number (\geq 2) as shown below.

```
top.D3PDVersion: 2
```

HowTo

- Described in AZAnalysisHowTo

Extras

- D0 Color Scheme
- Binning for basic plots
- Alternate scheme used in ICHEP 2010

```
color_index=5; // diboson
color_index=92; // W+j
color_index=95; // Z+j
color_index=62; // single top
color_index=619; // other
ttbar is empty with a black line.
```

- JES calculation

ArizonaPhysicsAnalysis < Main < TWiki

```
Method to scale jets and correct MET
pxJetsBefore = Sum(jet_px_before)
Scale the jet 4vec (keep the sign of px,py,pz components same)
pxJetsAfter = Sum(jet_px_after)
DeltaPx = pxJetsAfter - pxJetsBefore
Old_MEx = -Sum(px_allobjs) = -Sum(jet_px_before) - Sum(px_all_other_objs)
New_MEx = -Sum(jet_px_after) - Sum(px_all_other_objs)
          = -pxJetsAfter - Sum(px_all_other_objs)
          = -(DeltaPx + pxJetsBefore) - Sum(px_all_other_objs)
          = -DeltaPx + Old_MEx

Similarly find New_MEy; New_MET = sqrt(New_MEx^2 + New_MEy^2)

Old_SumET = Sum(abs(pt_all_objects)) = Sum(abs(jets_pt)) + Sum(abs(pt_of_other_objs))
New_SumET = Sum(abs(jets_pt_new)) + Sum(abs(pt_of_other_objs))
          = Sum(scaleF * abs(jets_pt)) + Sum(abs(pt_of_other_objs))
          = [scaleF * Sum(abs(jets_pt))] + Sum(abs(pt_of_other_objs))
          = [(1 +- delta) * Sum(abs(jets_pt))] + Sum(abs(pt_of_other_objs))
          = Old_SumET +- [delta * Sum(abs(jets_pt))]
```

-- VenkateshKaushik - 24-Jun-2010

- APS_10-16-2010.pdf: APS 4 corners Presentation by Mario Aletti

This topic: Main > ArizonaPhysicsAnalysis

Topic revision: r75 - 2011-04-28 - VenkateshKaushik



Copyright &© 2008-2019 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback