

Table of Contents

Penn Data Samples.....	1
Submitting the Trimming Job.....	1
At CERN.....	1
Data.....	1
MC.....	2
At Penn.....	2
Data.....	2
MC.....	2

Penn Data Samples

Submitting the Trimming Job.

The Trimming script is located `~olivito/public/gridD3PDtest/egammaD3PDTrimmerForData.py`, In it you can find the branches that we are currently saving. There is a separate script: `egammaD3PDTrimmerForMC.py` to be used on MC.

To run the trimming job locally do:

```
python egammaD3PDTrimmer.py `less fileList.txt`
```

where `fileList.txt` is a comma separated list of D3PDs to be trimmed.

To run the trimming job on the grid do:

1. Setup the grid tools and pathena.
2.

```
prun --exec "python egammaD3PDTrimmer.py %IN" --athenaTag=15.6.8 --inDS data10_7TeV.00154817.physics_L1Calo.merge.NTUP_EGAMMA.f255_p150/ --outDS user10.JohnAlison.00154817.L1Calo.NTUP_EGAMMA.f255.p150.trimmed.v3 --outputs "trimmed*root"
```

💡 Skip's Tips: Putting `"trimmed*root"` in the outputs is very important if your subjobs create more than one file (which root automatically creates if the file size exceeds 1.9G). If you just list a specific filename, that's the only output you'll get. So if you're sure your output will be less than 1.9G, you can just use `trimmed.root`.

💡 Skip's Tips: There's a limit of ~7G on the work directory size for each subjob. If it exceeds this, the job fails. This often happens when trying to trim MC datasets. One solution is to reduce the number of files per subjob with the `--nFilesPerJob <N>` argument.

When you get the output from the grid, if you used `"trimmed*root"` the ntuples will be zipped up in files like `*.trimmedXYZroot.tgz`. There's a script to untar all of these and name the ntuples uniquely:

```
~olivito/public/gridD3PDtest/untar_trimmed_ntuples.py <dir>
```

A list of current egamma D3PD datasets can be found on this page. (not current anymore as of 24/05/10!)

At CERN

Because of a mistake in the command used for the grid jobs, we only got back a fraction of the statistics. We're now rerunning the trimming (and including HLT decision info). The status flag in the table tells whether the run has been reprocessed yet or not.

Data

Dataset
data10_7TeV.00152409.physics_L1Calo.merge.NTUP_EGAMMA.r1239_p150_tid134458_00
data10_7TeV.00152441.physics_L1Calo.merge.NTUP_EGAMMA.r1239_p150_tid134923_00
data10_7TeV.00152777.physics_L1Calo.merge.NTUP_EGAMMA.r1239_p150_tid134455_00
data10_7TeV.00153565.physics_L1Calo.merge.NTUP_EGAMMA.f251_p150_tid134638_00
user10.JohnAlison.00154817.L1Calo.NTUP_EGAMMA.f255.p150.trimmed.v3
user10.JohnAlison.data10_7TeV.00154822.physics_L1Calo.merge.NTUP_EGAMMA.f255_p150.trimmed
user10.JohnAlison.data10_7TeV.00155073.physics_L1Calo.merge.NTUP_EGAMMA.f255_p150.trimmed

user10.JohnAlison.data10_7TeV.00155112.physics_L1Calo.merge.NTUP_EGAMMA.f255_p150.trimmed.withFix
user10.JohnAlison.data10_7TeV.00155116.physics_L1Calo.merge.NTUP_EGAMMA.f255_p150.trimmed.withFix/
user10.JohnAlison.data10_7TeV.00155118.physics_L1Calo.merge.NTUP_EGAMMA.f255_p150.trimmed.withFix.v2
user10.JohnAlison.data10_7TeV.00155160.physics_L1Calo.merge.NTUP_EGAMMA.f255_p150.trimmed.withFix

MC

Dataset
user.olivito.perf-egamma.mc09_7TeV.106046.PythiaZee_no_filter.recon.ESD.e468_s765_s767_r1253.NTUP_EGAMMA
user10.JohnAlison.mc09_7TeV.106043.PythiaWenu_no_filter.merge.NTUP_EGAMMA.e468_s765_s767_r1253_p150
user10.JoshuaMiloKunkle.mc09_7TeV.105001.pythia_minbias.merge.NTUP_EGAMMA.e517_s787_s767_r1250_p150
user.olivito.mc09_7TeV.105001.pythia_minbias.merge.AOD.e517_s764_s767_r1250_r1260.Egamma.110510.db10.4
user.olivito.mc09_7TeV.105724.PythiaB_bbe3X.merge.AOD.e511_s765_s767_r1253_r1260.Egamma.110510.db10.4
user.olivito.mc09_7TeV.105723.PythiaB_cce3X.merge.AOD.e511_s765_s767_r1253_r1260.Egamma.110510.db10.4
user.olivito.mc09_7TeV.105802.JF17_pythia_jet_filter.merge.AOD.e505_s765_s767_r1253_r126.Egamma.110510.db10.4

NOTE: for data on pc-penn-d-01, when this is mounted on another machine, the path is:

/project/data_d01_0/egammaD3PDs

At Penn

Data

Dataset	Machine	Path
---------	---------	------

MC

Dataset	Machine	Path
---------	---------	------

-- JohnAlisonUpenn - 18-May-2010 -- DominickOlivito - 17-May-2010

This topic: Sandbox > PennData

Topic revision: r27 - 2010-05-31 - DominickOlivito



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