



Skims



Overall idea

NanoAOD are taking ages to be produced

Skim on NanoAOD does not exist centrally, need to transfer the data locally to use the skim from Korea group + JSON application etc

→ Not easy when too many MC samples etc

→ Try to use CRAB as much as we can!

Change strategy! Produce our own NanoAOD skimmed files!

Instead of skimming after NanoAOD, skim before they are produced

→ Stored only Skimmed NanoAOD

→ Use Crab to apply JSON files etc

→ Tier3 transfer will be done via Crab etc

→ We will not depend from CMS production as soon as MiniAOD are available



NanoAOD production (CMSSW_10_6_27) IP2I

<https://twiki.cern.ch/twiki/bin/view/CMS/PdmVDataReprocessingReNanoAODULFall2020>

<https://gitlab.cern.ch/cms-nanoAOD/nanoaod-doc/-/wikis/Instructions/Private%20production>

[/wikis/Instructions/Private%20production](https://gitlab.cern.ch/cms-nanoAOD/nanoaod-doc/-/wikis/Instructions/Private%20production)

+ <https://gitlab.cern.ch/cms-nanoAOD/nanoaod-doc/-/wikis/Releases/NanoAODv9>

```
cmsDriver.py NANO -s NANO --mc --conditions ${GT} --era ${ERA} --eventcontent  
NANOAODSIM --datatier NANOAODSIM --  
customise_commands="process.add_(cms.Service('InitRootHandlers', EnableIMT =  
cms.untracked.bool(False)));process.MessageLogger.cerr.FwkReport.reportEvery=100" -n -1  
--no_exec
```

With \${GT} and \${ERA} coming from: (for NanoAODv9)

Input dataset	GT	Era
RunIISummer20UL16MiniAOD/RunIISummer19UL16MiniAODv2 MC	106X_mcRun2_asymptotic_v17 or 106X_mcRun2_asymptotic_preVFP_v11	Run2_2016,run2_nanoAOD_106Xv2
RunIISummer20UL17MiniAOD/RunIISummer19UL17MiniAODv2 MC	106X_mc2017_realistic_v9	Run2_2017,run2_nanoAOD_106Xv2
RunIISummer20UL18MiniAOD/RunIISummer19UL18MiniAODv2 MC	106X_upgrade2018_realistic_v16_L1v1	Run2_2018,run2_nanoAOD_106Xv2
2016 UL data (Summer19/20)	106X_dataRun2_v35	Run2_2016,run2_nanoAOD_106Xv2
2017 UL data (Summer19/20)	106X_dataRun2_v35	Run2_2017,run2_nanoAOD_106Xv2
2018 UL data (Summer19/20)	106X_dataRun2_v35	Run2_2018,run2_nanoAOD_106Xv2



Vtx Selector

https://cmsddt.cern.ch/lxr/source/PhysicsTools/UtilAlgos/test/testPrimaryVertexFilter_cfg.py?v=CMSSW_10_6_20

```
process.primaryVertexFilter =  
cms.EDFilter("PrimaryVertexFilter",  
pvSrc = cms.InputTag("offlineSlimmedPrimaryVertices"),  
minNdof = cms.double( 4 ),  
maxZ = cms.double( 24.0 ),  
maxRho = cms.double(2.0),  
NPV = cms.int32(1))
```



Basic Cuts

IP2I

<https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuidePhysicsCutParser>

```
process.goodMuons = cms.EDFilter("CandViewSelector",  
src = cms.InputTag("slimmedMuons"),  
cut = cms.string ('isGlobalMuon & pt > 100. & abs(eta) < 2.5 &  
(trackIso+caloIso)/pt < 0.2')  
)
```

```
process.goodElecs = cms.EDFilter("CandViewSelector",  
src = cms.InputTag("slimmedElectrons"),  
cut = cms.string ('et > 50. & abs(eta) < 2.5 &  
(dr03TkSumPt+dr03EcalRecHitSumEt+dr03HcalTowerSumEt)/et < 0.2' )  
)
```

```
process.goodPatJets = cms.EDFilter("CandViewSelector",  
    cut = cms.string ('pt > 300. & abs(eta) < 2.5 '),  
    src = cms.InputTag("slimmedJets"),  
)
```

→ Cuts can take many form, check the twiki page to find out



Basic Cuts

IP2I

```
process.countJets = cms.EDFilter("PATCandViewCountFilter",  
    minNumber = cms.uint32 (3),  
    maxNumber = cms.uint32 (999999),  
    src      = cms.InputTag("goodPatJets"),)  
process.countElecs = cms.EDFilter("PATCandViewCountFilter",  
    minNumber = cms.uint32 (1),  
    maxNumber = cms.uint32 (999999),  
    src      = cms.InputTag("goodElecs"),)
```

```
process.nanoAOD_selection =  
cms.Sequence(process.primaryVertexFilter*process.goodMuons*process.goodElecs*pro  
cess.goodPatJets*process.countJets*process.countElecs)  
process.nanoAOD_step =  
cms.Path(process.nanoAOD_selection*process.nanoSequenceMC)
```

+Add in output Module:

```
SelectEvents = cms.untracked.PSet(  
    SelectEvents = cms.vstring('nanoAOD_step')  
)
```

→ **CAREFUL:** NanoAOD will still have objects with lower pt cut

→ **ONLY** selection of event is made, not of objects (→ I think we want to keep it this way for JEC etc)

Currently one crash when no events are selected:

<https://hypernews.cern.ch/HyperNews/CMS/get/physTools/3822/1/2.html>



To be defined

IP2I

Which basic selection do we want to apply?

- **Not HLT as we want to compute trigger efficiency, data are already skimmed by HLT**
- **Lower cut than the analysis ones → to have room for applying systematics (JEC etc)**
 - **Need to figure out if Filter are AFTER new JEC applied or BEFORE**
- **Standard CMS filter (Vertex + Noise filters) but not sure to find easily how to configure them**
- **JSON file (by definition within processing)**
- **Basic final state request (5 jets, 1lep+3jets etc)**
- **We could run all selection in one go (ie for each analysis [as HLT])**
 - **Not so sure it makes sense, except for commun ttbar (data could be on different dataset etc...)**