PAT Tutorial

Sudhir Malik

Fermilab/University of Nebraska-Lincoln

JTerm@Fermilab 3-5 August 2009

Physics Analysis Toolkit - Concepts

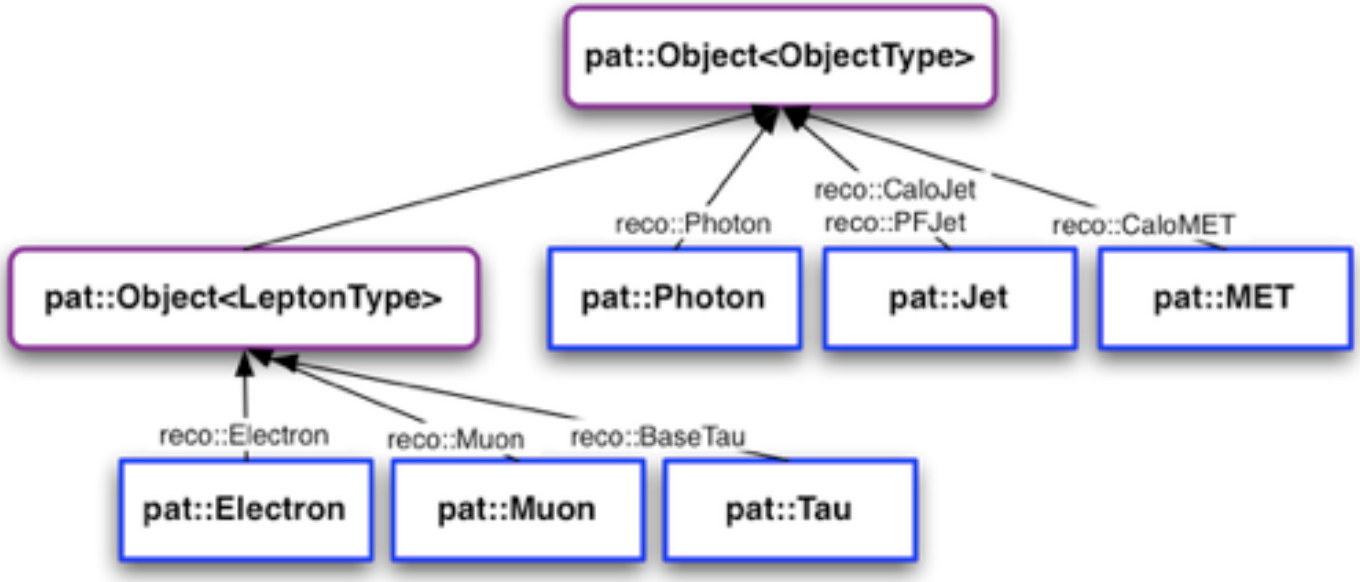
- **Interface**
  - between reconstruction and analysis
  - simplifies access via data formats
  - crossing point between
    - POGs (Physics Object Groups)
    - PAGs (Physics Analysis Groups)
  - channelizes expertise - POGs and PAGs contacts
- **Common Tool**
  - Common Standards, sensible defaults, approved algorithms
  - quick start into analysis
- **Common Format**
  - facilitates transfer and comparisons
  - PAG common configurations
  - sustained provenance

PAT Code in CVS

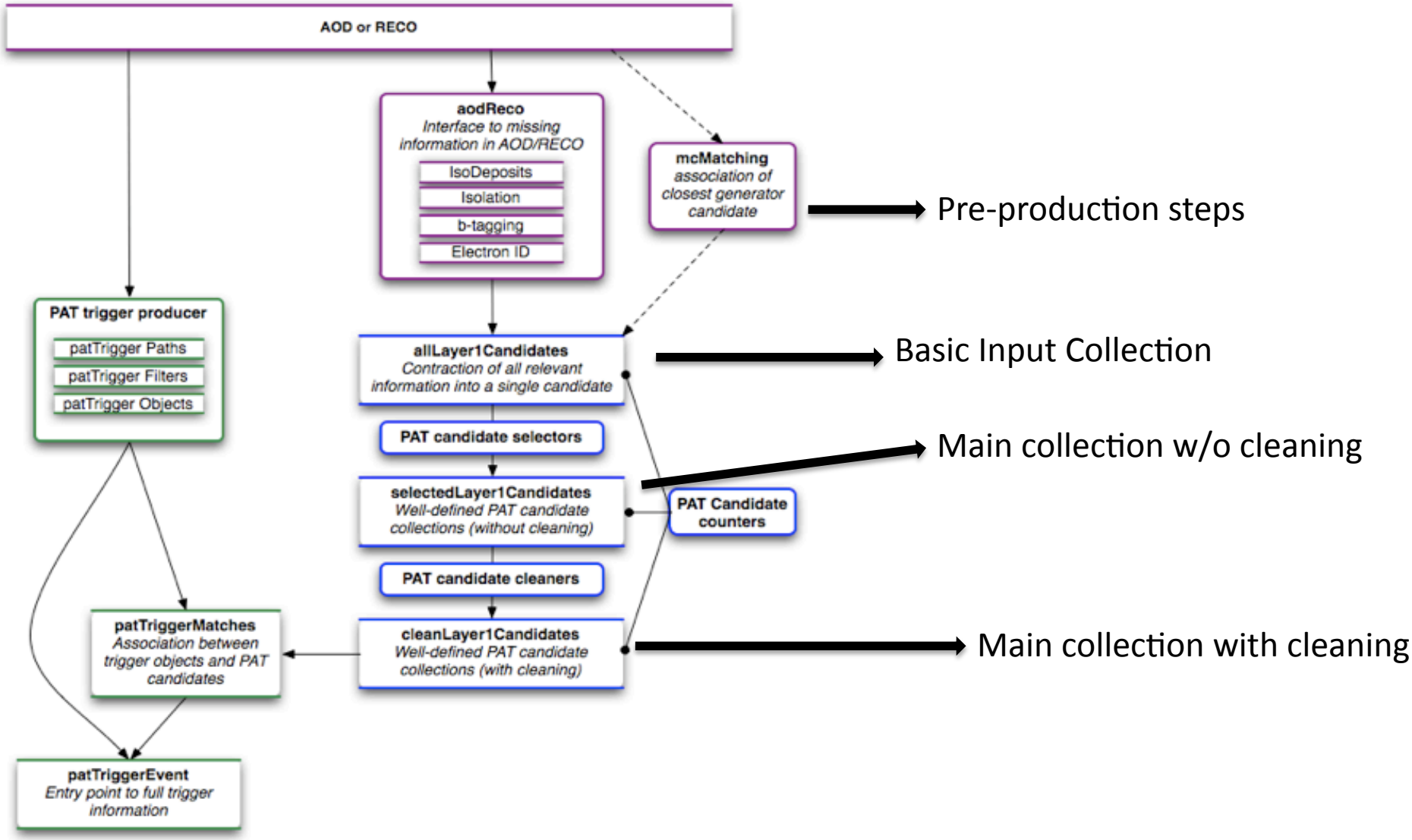
- **All code located in CMSSW, distributed in two systems**
  - **DataFormats/PatCandidates**
    - Structure and Candidate Classes
      - pat::Photon,pat::Electron,pat::Muon,pat::Jet,pat::MET, …
  - **PhysicsTools**
    - PatAlgos - Classes for pat::Candidate creation, algorithms
    - PatUtils – common utilities, isolation, object disambiguation,
    - PatExamples – Example analyzers built up for tutorials
- **Link to CVS**
  - http://cmssw.cvs.cern.ch/cgi-bin/cmssw.cgi/CMSSW/

Data Formats

- All **pat::Objects** inherit from their corresponding **reco::Candidates**
- A **pat::Object** is a **reco::Object** (plus more)

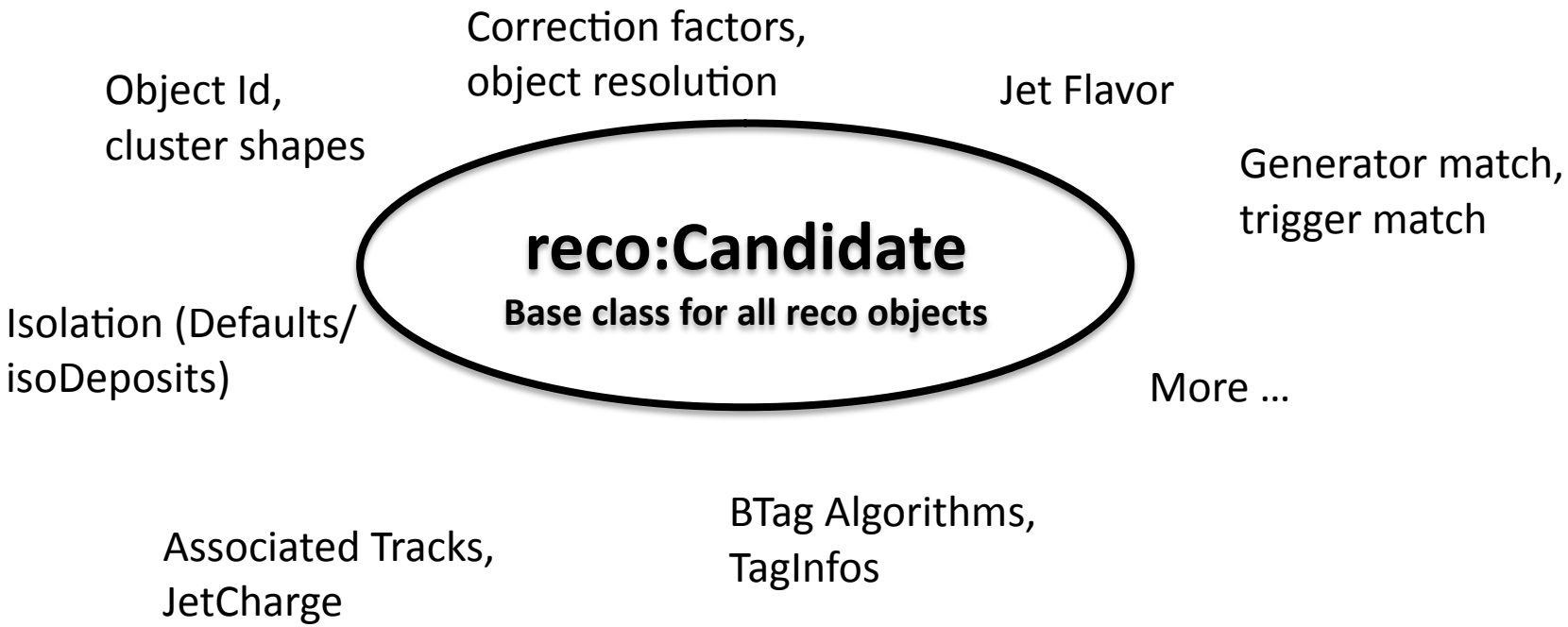


Workflow


AOD or RECO
aodReco
Interface to missing information in AOD/RECO
IsoDeposits
Isolation
b-tagging
Electron ID
mcMatching
association of closest generator candidate
Pre-production steps
PAT trigger producer
patTrigger Paths
patTrigger Filters
patTrigger Objects
allLayer1Candidates
Contraction of all relevant information into a single candidate
Basic Input Collection
PAT candidate selectors
selectedLayer1Candidates
Well-defined PAT candidate collections (without cleaning)
Main collection w/o cleaning
PAT Candidate counters
PAT candidate cleaners
patTriggerMatches
Association between trigger objects and PAT candidates
cleanLayer1Candidates
Well-defined PAT candidate collections (with cleaning)
Main collection with cleaning
patTriggerEvent
Entry point to full trigger information

pat::Candidate creation

- **Each pat::Candidate is a reco::Candidate + more..**
  - All info is folded into pat::Candidate



pat::Candidate creation

- **Information can be made persistent (*embedded*) or kept as reference**
  - All persistent data is easily accessible in FWLite
    - Today's beginner's tutorial talks about that in detail
  - Flexible size management of a single **pat::Candidate** class
  - Fully transparent for the user
- **The size of the pat::Candidate depends on the choice of the user**

RECO ~ 500 kb/ev	AOD ~ 100 kb/ev	PAT 6-60 kb/ev

- Results in allLayer1Candidates (basic candidate collection of PAT)

EventContent

- **The patAODExtraReco -**Produce external information to be put in PAT objects
  - this is all stuff you retrieve from reco, no meta-physics in between.
  - PAT completes RECO and not a parallel world

- **Compose the pat::EventContent to your will:**

```
patEventContentNoLayer1Cleaning = [
    'keep *_selectedLayer1Photons_*_*',
    'keep *_selectedLayer1Electrons_*_*',
    'keep *_selectedLayer1Muons_*_*',
    'keep *_selectedLayer1Taus_*_*',
    'keep *_selectedLayer1Jets_*_*',
    'keep *_layer1METs_*_*',
    'keep *_selectedLayer1PFParticles_*_*'
]
```

OR

```
patEventContent = [
    'keep *_cleanLayer1Photons_*_*',
    'keep *_cleanLayer1Electrons_*_*',
    'keep *_cleanLayer1Muons_*_*',
    'keep *_cleanLayer1Taus_*_*',
    'keep *_cleanLayer1Jets_*_*',
    'keep *_layer1METs_*_*',
    'keep *_cleanLayer1Hemispheres_*_*',
    'keep *_cleanLayer1PFParticles_*_*'
]
patExtraAodEventContent = [
    # GEN
    'keep recoGenParticles_genParticles_*_*',
    'keep *_genEventScale_*_*',
    'keep *_genEventWeight_*_*',
    'keep *_genEventPdfInfo_*_*',
    # RECO
    'keep recoTracks_generalTracks_*_*',
    'keep *_towerMaker_*_*',
    'keep *_offlineBeamSpot_*_*',
    'keep *_offlinePrimaryVertices_*_*',
    # TRIGGER
    'keep edmTriggerResults_TriggerResults_*_HLT',
    'keep *_hltTriggerSummaryAOD_*_*'
]
```

- **Add EventContent to the output module:**

```
# Output module configuration
from PhysicsTools.PatAlgos.patEventContent_cff import *
process.out = cms.OutputModule("PoolOutputModule",
    fileName = cms.untracked.string('PATLayer1_Output.fromAOD_full.root'),
    # save only events passing the full path
    SelectEvents   = cms.untracked.PSet( SelectEvents = cms.vstring('p') ),
    # save PAT Layer 1 output
    outputCommands = cms.untracked.vstring('drop *')
)
process.out.outputCommands += patEventContent
```

Workflow Tools

- There are tool available to estimate the disksize of events and their components.
  - **diskSize.pl**

Collection	items/event	kb/event	kb/item	plot	%
patJets_cleanLayer1Jets__PAT	1.00	6.68	6.68		29.9%
patElectrons_cleanLayer1Electrons__PAT	1.00	5.00	5.00		22.4%
patPhotons_cleanLayer1Photons__PAT	1.00	4.27	4.27		19.1%
patMuons_cleanLayer1Muons__PAT	1.00	4.04	4.04		18.1%
patTaus_cleanLayer1Taus__PAT	1.00	1.13	1.13		5.1%
patMETs_layer1METs__PAT	1.00	0.92	0.92		4.1%
patHemispheres_cleanLayer1Hemispheres__PAT	1.00	0.30	0.30		1.4%

  - **edmEventSize**

```
[malik@cmslpc05 src]$ edmEventSize -v myTuple.root]
File myTuple.root Events 1300
patJets_cleanLayer1Jets__PAT. 35140.5 5378.1
patElectrons_cleanLayer1Electrons__PAT. 17915.1 8460.48
patPhotons_cleanLayer1Photons__PAT. 6662.56 3428.32
patMuons_cleanLayer1Muons__PAT. 6365.38 3093.48
patTaus_cleanLayer1Taus__PAT. 5013.97 1709.78
patMETs_layer1METs__PAT. 1277.08 829.328
patHemispheres_cleanLayer1Hemispheres__PAT. 611.965 183.739
EventAuxiliary 138.373 12.5177
```

Workflow Tools

- **Tools will help you to configure the pat::Workflow:**
  - **addJetCollections: (patLayer1_fromAOD_jetSuite_full)**

```
############Exercise-6 Add KT4 Jet Collection ################
from PhysicsTools.PatAlgos.tools.jetTools import *
addJetCollection(process,cms.InputTag('kt4CaloJets'),'KT4',
                         doJTA=True,doBTagging=True,jetCorrLabel=('KT4','Calo'),doType1MET=True,doL1Counters=False,
                         genJetCollection=cms.InputTag("kt4GenJets"))
```

  - **switchJetCollections (patLayer1_fromAOD_sisCone_full)**

```
## ==== Example with CaloJets
switchJetCollection(process
        ,cms.InputTag('sisCone5CaloJets')  # Jet collection; must be already in the event when patDefaultSequence is executed
        ,doJTA=True            # Run Jet-Track association & JetCharge
        ,doBTagging=True       # Run b-tagging
        ,jetCorrLabel=('SC5','Calo') # example jet correction name; set to None for no JEC
        ,doType1MET=True       # recompute Type1 MET using these jets
        ,genJetCollection=cms.InputTag("sisCone5GenJets")
 )
```

Choose between supported jet algo's, pflow, JPT, calo, user defined

  - **More Tools: switch τ collections, tcMET, trigger configuration, ..**

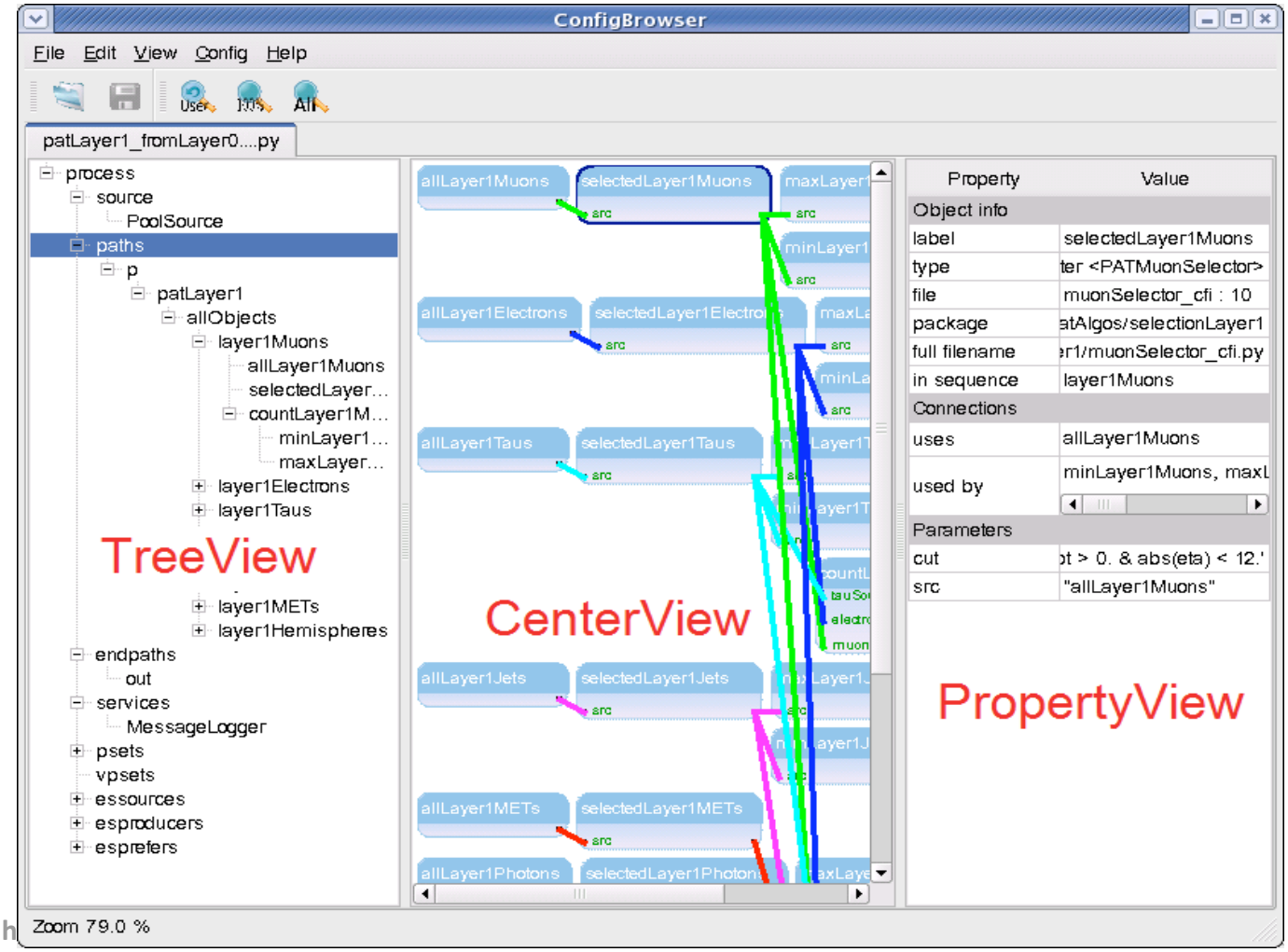
ConfigBrowser for Python Configs

•The Config Browser is a tool for inspection of the structure of Python Config files in CMSSW. It allows to:

- Process configuration via click and menu
- Less error prone / much improved user friendliness
- Visualize the complete structure of a Config File and all included config files (via import)
- Inspect the parameters of modules
- Track which modules use input from which other modules
- Track in which file certain modules can be found
- Open the definition of certain modules in the user's favorite editor
- Edit a config file and save a config file which lists all changes to the orignal config file

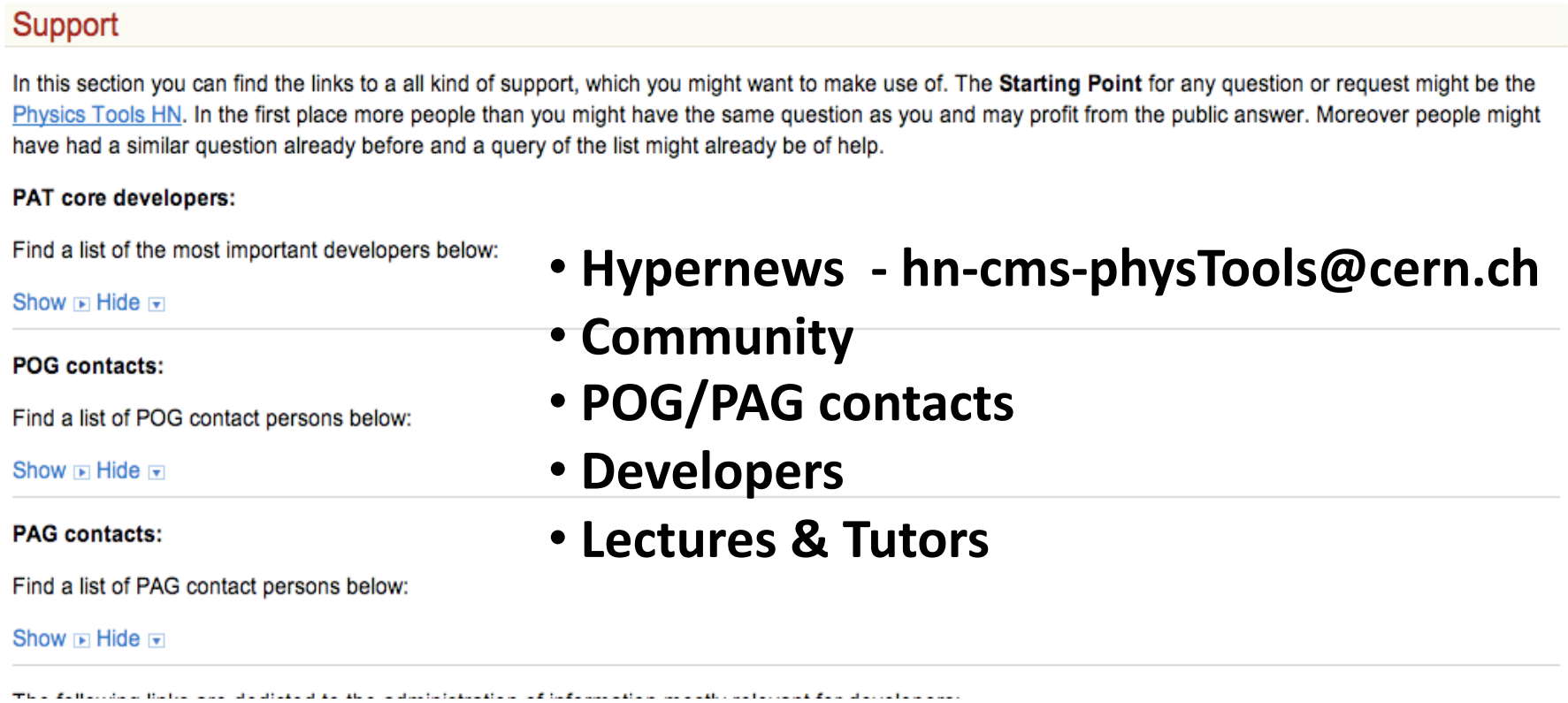
ConfigBrowser Snapshot

- **How to install and use this tool see:**
  - https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideConfigBrowser



Support

- **For more information on support have a look at SWGuidePAT**



Support

In this section you can find the links to a all kind of support, which you might want to make use of. The **Starting Point** for any question or request might be the Physics Tools HN. In the first place more people than you might have the same question as you and may profit from the public answer. Moreover people might have had a similar question already before and a query of the list might already be of help.

PAT core developers:

Find a list of the most important developers below:

Show ▸ Hide ▾

POG contacts:

Find a list of POG contact persons below:

Show ▸ Hide ▾

PAG contacts:

Find a list of PAG contact persons below:

Show ▸ Hide ▾

- **Hypernews - hn-cms-physTools@cern.ch**
- **Community**
- **POG/PAG contacts**
- **Developers**
- **Lectures & Tutors**

Documentation

- SWGuidePAT - Main documentation page
- SWGuidePATRecipes - Installation recipes
- SWGuidePATExamples - Tutorials and examples to get started
- SWGuidePATDataFormats - pat::Candidate description
- SWGuidePATConfiguration - Module configuration
- SWGuidePATEventSize - Tools for event size estimate
- SWGuidePATWorkflow - PAT workflow description
- SWGuidePATTools - Description of workflow configuration tools

Please prefix this to the SWGuides below - https://twiki.cern.ch/twiki/bin/view/CMS/

Over to Tutorial Exercises

- **The PAT tutorial twiki is here**
  - https://twiki.cern.ch/twiki/bin/view/Sandbox/JTermPAT
  - Use release CMSSW_3_1_2
  - 8 Exercises and links to many other