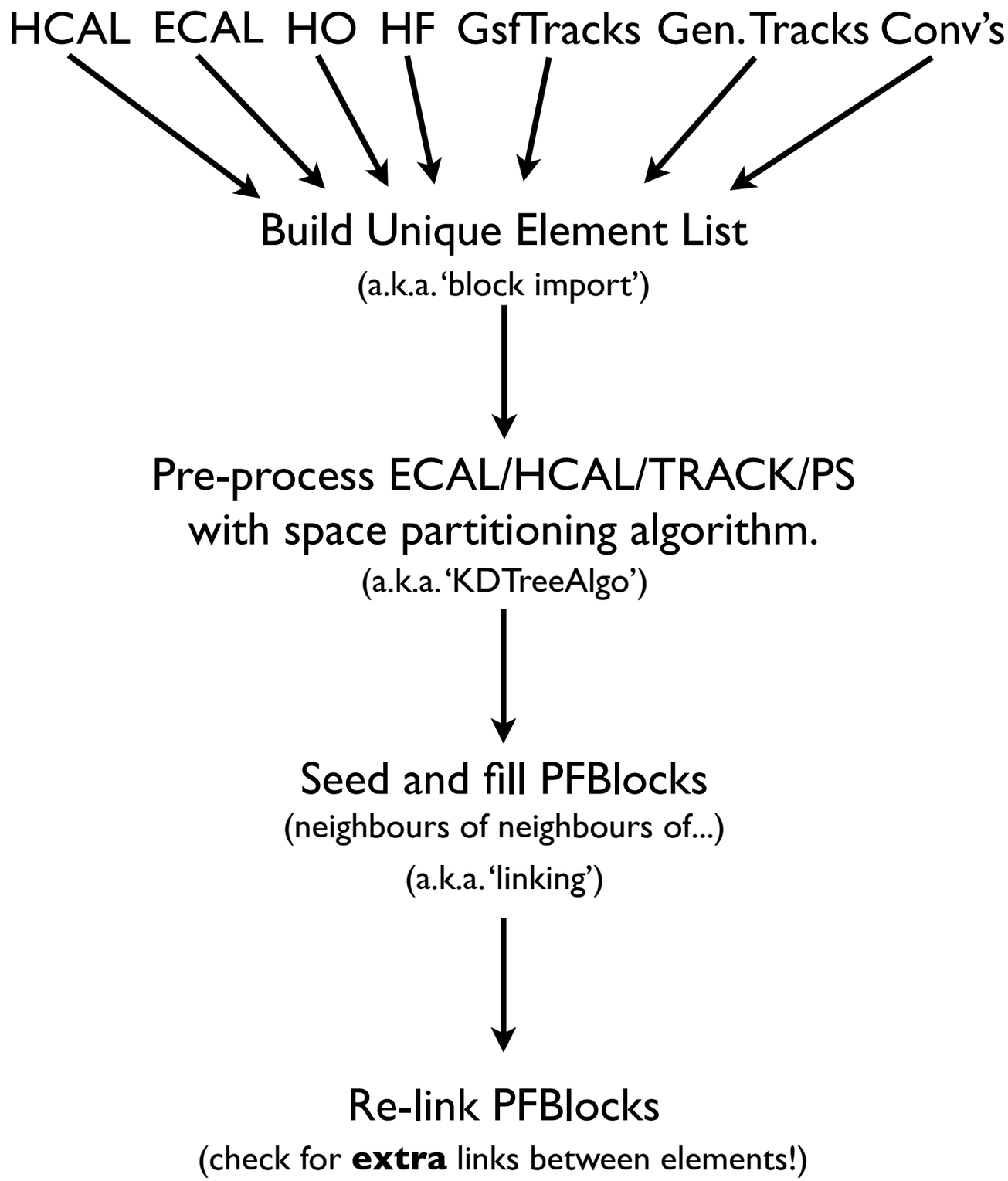
New PFBlockProducer Structure & Validation

Lindsey Gray
22 April, 2014

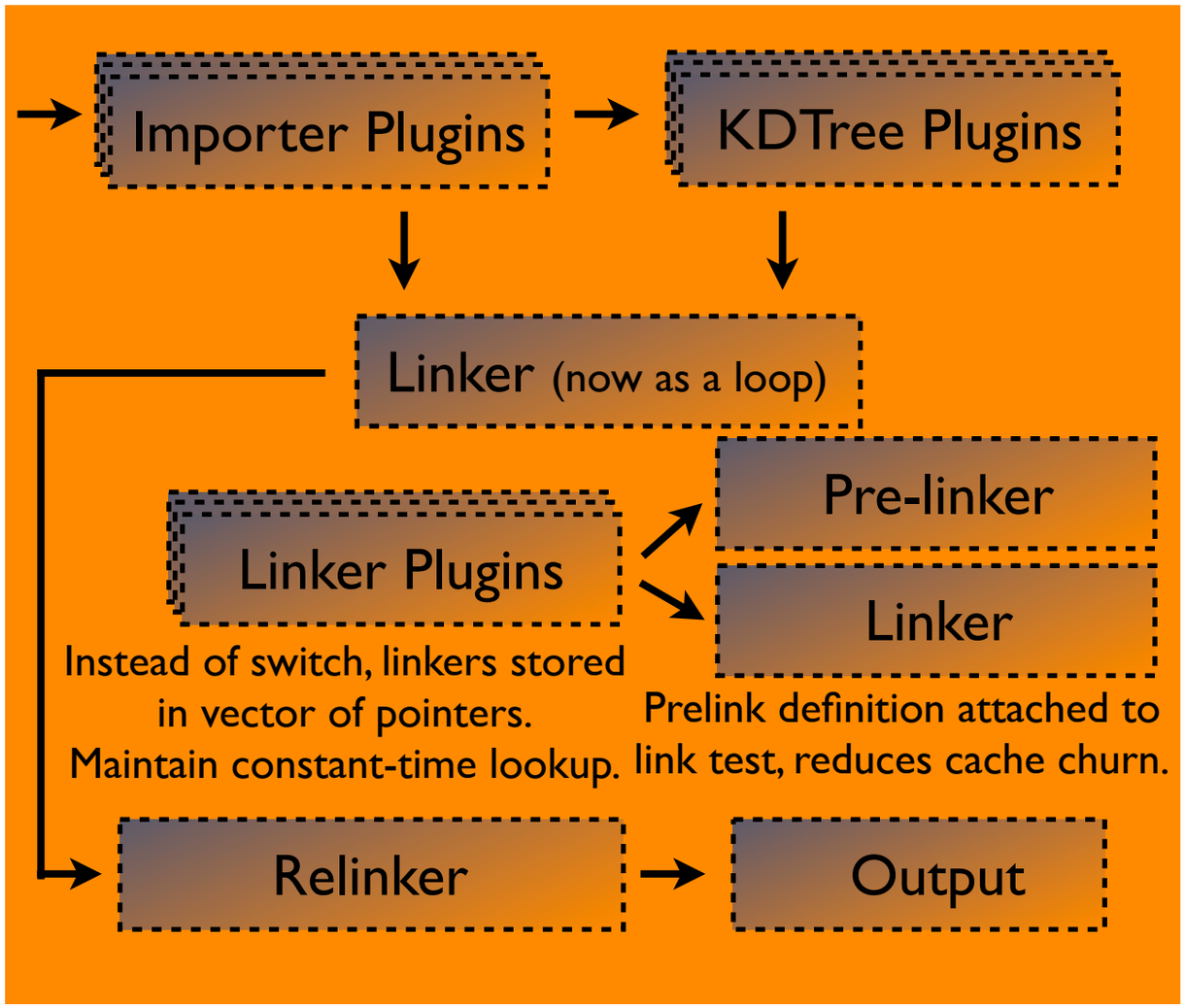
PFBlockProducer Workflow

- List of input objects different for HLT and Offline
  - Reduced list for HLT timing budget
- Most combinatorially expensive pairs are pre-processed
  - KDTreeAlgo gives quick access to closest neighbour
- Links found through iterative-neighbours approach
  - The same as topological clusters!
- Only one link tested during the first step
  - Need to check for additional links in block (can change final EFlow!)



New PFBlock Producer Structure

- Remove all instances of hardcoding
  - Configuration is driven entirely by python cfi's
  - Importer plugins each import one variety of object
    - Still able to protect against double-import!
  - Linker plugins each define their pre-linking condition and link test
    - KDTree usage defined per-linker for link preprocessing
    - Possibility to add in new fast-linking in whatever variables you want
- Linker no longer recursive
  - Yields minor performance gain for long loops



Changes in Design Philosophy

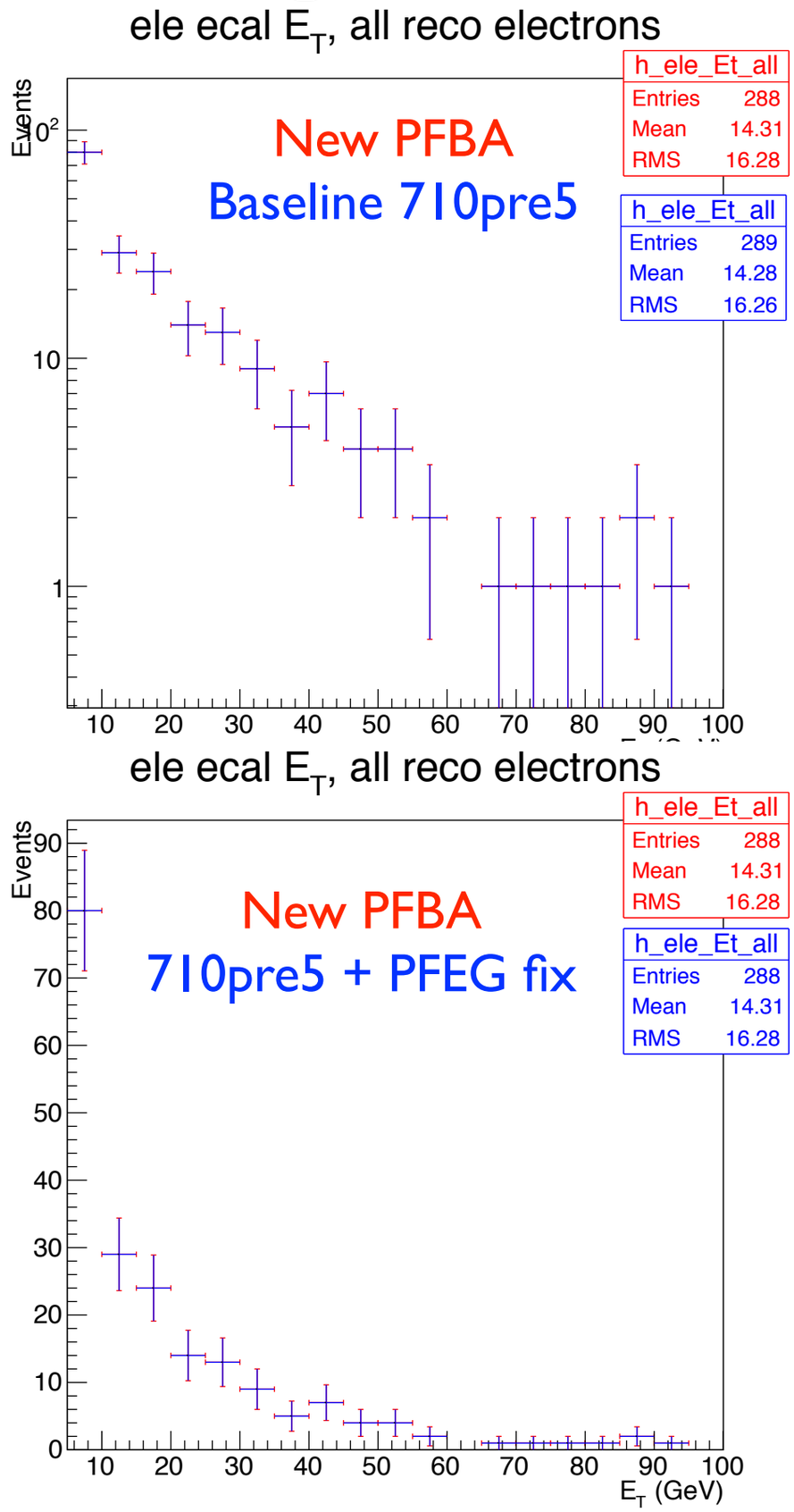
- Introduction of new importing or new linking should (and does) require minimal programming overhead
  - Just write new importing or linking module and stuff it in the python config
    - The time it takes to write a new module is the time it takes to have your idea in Particle Flow
    - The extra time to understand whatever behavior it introduces is entirely your own fault :-)
- Allow POGs/PAGs/free-agents with ideas, good or bad, to easily introduce new behavior to particle flow, perform studies, and have them reviewed by the community at large
  - Organizing PFBlockProducer into plugins lets developers focus on the physics impact of what they're doing by plainly exposing the relevant parts of the algorithm.
    - No one cares how you loop over elements, only how the elements are related!
  - This exposes PFlow to the whole collaboration, which is an absolute good for CMS.
  - Combined with recent PFClustering modularization this allows for quick adaptation to all of the upgrade calorimeter designs

Changes in Functionality

- No longer able to run in FWLite mode
  - Not necessary any longer due to PFRootEvent overhaul-in-progress
- Easy customization of linking for HLT or other situations
  - Define in python what you import and link
  - If an operation isn't in the python it doesn't happen
    - ... and if something is there it does!
    - Strong guarantee ;-)
- Configuration information is compartmentalized
  - Linking/importing modules provide structure for organizing the parameter of the algorithm
  - Makes it easier to know exactly how you're putting all the ingredients into particle flow
- Ordering of elements within a block is entirely different
  - Links within and content of a block are exactly the same

Validation & Exposed Bugs I

- New PFBlockProducer is submitted on github
  - http://goo.gl/6reTlo
  - Presently waiting for updated confdb setup from Martin G.
  - Already ported to 620_SLHC
- First round of validation found bugs in PFEGammaAlgo
  - Different relative ordering of elements produced different e/γ candidates
  - This is fixed in the PR mentioned above



Validation & Exposed Bugs 2

- Further validation performed event by event comparing contents
  - PFBlocks produced in different order
    - Elements within each block also in different order
  - Content and links found to be exactly the same
- Still, found different lists of PFCandidates!
  - Also checked changing the ordering of the elements in the blocks in the **old** PFBlockProducer
  - Got different answer there too!
    - There is an order dependency in PFAlgo that we need to find and fix

Standard ordering:

```
Begin processing the 11th record. Run 1, Event 311, LumiSection 4 at 16-Apr-2014 19:43:28.744 CEST
Begin processing the 12th record. Run 1, Event 312, LumiSection 4 at 16-Apr-2014 19:43:54.435 CEST
Begin processing the 13th record. Run 1, Event 313, LumiSection 4 at 16-Apr-2014 19:44:14.055 CEST
Begin processing the 14th record. Run 1, Event 314, LumiSection 4 at 16-Apr-2014 19:44:36.556 CEST
Begin processing the 15th record. Run 1, Event 315, LumiSection 4 at 16-Apr-2014 19:44:50.463 CEST
Begin processing the 16th record. Run 1, Event 316, LumiSection 4 at 16-Apr-2014 19:45:12.781 CEST
%MSG-w InvalidHelix:  SeedGeneratorFromRegionHitsEDProducer:tobTecStepSeedsPair  16-Apr-2014 19:45:14 CEST Run
: 1 Event: 316
PixelClusterShapeSeedComparitor helix is not valid, result is bad
%MSG
%MSG-w InvalidHelix:  SeedGeneratorFromRegionHitsEDProducer:tobTecStepSeedsPair  16-Apr-2014 19:45:14 CEST Run
: 1 Event: 316
PixelClusterShapeSeedComparitor helix is not valid, result is bad
%MSG
Begin processing the 17th record. Run 1, Event 317, LumiSection 4 at 16-Apr-2014 19:45:30.110 CEST
%MSG-e TooManySeeds:   SeedGeneratorFromRegionHitsEDProducer:regionalCosmicTrackerSeeds  16-Apr-2014 19:45:40
CEST Run: 1 Event: 317
Found too many seeds (10002 > 10000), bailing out.

%MSG
140416 19:45:57 9464 Xrd^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@
^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@140416 19:45:57 9468 Xrd^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@
^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@140416 19:45:58 8356 Xrd^@^@^@^@^@
^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@
^@^@140416 19:45:58 9461 Xrd^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@
^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@140416 19:45:58 9466 Xrd^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@
^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@^@Begin processing the 18th reco
rd. Run 1, Event 318, LumiSection 4 at 16-Apr-2014 19:46:07.789 CEST
Begin processing the 19th record. Run 1, Event 319, LumiSection 4 at 16-Apr-2014 19:46:28.051 CEST
Begin processing the 20th record. Run 1, Event 320, LumiSection 4 at 16-Apr-2014 19:46:50.187 CEST
Begin processing the 21st record. Run 1, Event 321, LumiSection 4 at 16-Apr-2014 19:47:05.853 CEST
Begin processing the 22nd record. Run 1, Event 322, LumiSection 4 at 16-Apr-2014 19:47:47.161 CEST
Begin processing the 23rd record. Run 1, Event 323, LumiSection 4 at 16-Apr-2014 19:48:08.840 CEST
Begin processing the 24th record. Run 1, Event 324, LumiSection 4 at 16-Apr-2014 19:48:23.863 CEST
Begin processing the 25th record. Run 1, Event 325, LumiSection 4 at 16-Apr-2014 19:48:40.080 CEST
Begin processing the 26th record. Run 1, Event 326, LumiSection 4 at 16-Apr-2014 19:49:08.773 CEST
Begin processing the 27th record. Run 1, Event 327, LumiSection 4 at 16-Apr-2014 19:49:38.439 CEST
Begin processing the 28th record. Run 1, Event 328, LumiSection 4 at 16-Apr-2014 19:49:59.521 CEST
```

Different ordering:

```
Begin processing the 11th record. Run 1, Event 311, LumiSection 4 at 21-Apr-2014 15:30:09.032 CEST
+++WARNING+++ PFCandidate size changed for entry 10 !
 - RECO    size : 1640
 - Re-RECO size : 1641
+++WARNING+++ PFCandidate 3 changed  for entry 10 !
 - RECO     :   PFCandidate type: 4 E/pT/eta/phi 30.125/9.524/1.819/-1.175, blocks/iele: (0|34), (0|35),  sour
ce:3:1119/118
 - Re-RECO  :   PFCandidate type: 4 E/pT/eta/phi 13.275/11.745/0.505/-2.582, blocks/iele: (1|169), (1|210), (1
|180), (1|181),  source:4:755/401
 DeltaE   = : -0.388
 DeltaEta = : -1.31
 DeltaPhi = : -1.41

+++WARNING+++ PFCandidate 4 changed  for entry 10 !
 - RECO     :   PFCandidate type: 1 E/pT/eta/phi 29.921/9.373/1.828/-1.105, blocks/iele: (0|108), (0|38), (0|1
09),  source:3:1119/36
 - Re-RECO  :   PFCandidate type: 4 E/pT/eta/phi 30.125/9.524/1.819/-1.175, blocks/iele: (1|97), (1|98),  sour
ce:4:755/353
 DeltaE   = : 0.00339
 DeltaEta = : -0.00968
 DeltaPhi = : -0.0704

+++WARNING+++ PFCandidate 5 changed  for entry 10 !
 - RECO     :   PFCandidate type: 5 E/pT/eta/phi 42.803/7.928/-2.371/-0.302, blocks/iele: (3|37), (3|148), (3|
0), (3|120), (3|125), (3|116), (3|117), (3|142), (3|34), (3|86), (3|88), (3|87),  source:3:1119/309
 - Re-RECO  :   PFCandidate type: 1 E/pT/eta/phi 29.921/9.373/1.828/-1.105, blocks/iele: (1|87), (1|177), (1|8
8),  source:4:755/413
 DeltaE   = : -0.177
 DeltaEta = : 4.2
 DeltaPhi = : -0.803

+++WARNING+++ PFCandidate 6 changed  for entry 10 !
 - RECO     :   PFCandidate type: 1 E/pT/eta/phi 16.450/7.516/-1.420/-2.901, blocks/iele: (7|0),  source:3:111
9/463
 - Re-RECO  :   PFCandidate type: 5 E/pT/eta/phi 42.803/7.928/-2.371/-0.302, blocks/iele: (0|37), (0|148), (0|
0), (0|120), (0|123), (0|116), (0|117), (0|142), (0|34), (0|86), (0|88), (0|87),  source:4:755/81
 DeltaE   = : 0.445
 DeltaEta = : -0.951
```

Conclusions / ToDo

- New PFBlockProducer is in the queue to be integrated with CMSSW
  - Already in 620_SLHC releases
  - Just waiting for HLT integration
- New PFBlockProducer reduces development overhead
  - Just make a importer or linker and plug it in
  - Reduces necessary learning curve, could get POGs/PAGs more involved
- Validation performed and a few bugs fixed or exposed
  - Only observe jitter in Jets and MET when ordering of elements induces different behavior in PFAlgo