ID Trigger Validation Requirements

Mark Sutton

University of Sussex

Preface

- In order to make use of the same framework for Trigger and Offline performance measurements, the framework must be able to properly handle all use cases
  - Also if any aspect for either Offline or Trigger is more complicated ion pone case than the other, then the framework has to be **natively designed** to handle t**he more complex** use case, with the more simple case handled using wrappers and specialisations.
  - For example:
    - Offline matching - generally more complex than that used in the trigger, ie the trigger generally is based on matching of track parameters, η, ϕ, z_0, can use hit multiplicity etc, but generally **does not** use matching for specific hits
    - Trigger navigation - Orders or magnitude more complex than Offline analyses - requires analyses for different trigger chains, different instance of the tracking, different Rois, single leg triggers, tag-and-probe triggers etc.

- So how should the matching and event storage be handled ?
- Clearly cannot use an offline based model, since that hs no concept of different chains, Rois, different reference and test track collections, used within a run etc.
- How complex are we talking about for the typical Tier 0 analysis for any given run ? …

Typical Tier 0 analysis …

- Analyses for chains produced using the TrigInDetAnalysis code ….

```
HLT_e14_idperf_tight_L1eEM9/HLT_IDTrack_Electron_FTF_HLT_Roi_FastElectron
HLT_e14_idperf_tight_L1eEM9/HLT_IDTrack_Electron_GSF
HLT_e14_idperf_tight_L1eEM9/HLT_IDTrack_Electron_IDTrig
HLT_e14_idperf_tight_nogsf_L1eEM9/HLT_IDTrack_Electron_FTF_HLT_Roi_FastElectron
HLT_e14_idperf_tight_nogsf_L1eEM9/HLT_IDTrack_Electron_IDTrig
HLT_e14_lhtight_e4_idperf_tight_probe_1invmAB5_L1JPSI-1M5-EM12/HLT_IDTrack_Electron_FTF_HLT_Roi_FastElectron
HLT_e14_lhtight_e4_idperf_tight_probe_1invmAB5_L1JPSI-1M5-EM12/HLT_IDTrack_Electron_GSF
HLT_e14_lhtight_e4_idperf_tight_probe_1invmAB5_L1JPSI-1M5-EM12/HLT_IDTrack_Electron_GSF_1
HLT_e14_lhtight_e4_idperf_tight_probe_1invmAB5_L1JPSI-1M5-EM12/HLT_IDTrack_Electron_GSF_1_el_probe
HLT_e14_lhtight_e4_idperf_tight_probe_1invmAB5_L1JPSI-1M5-EM12/HLT_IDTrack_Electron_IDTrig
HLT_e14_lhtight_e4_idperf_tight_probe_1invmAB5_L1JPSI-1M5-eEM15/HLT_IDTrack_Electron_FTF_HLT_Roi_FastElectron
HLT_e14_lhtight_e4_idperf_tight_probe_1invmAB5_L1JPSI-1M5-eEM15/HLT_IDTrack_Electron_GSF
HLT_e14_lhtight_e4_idperf_tight_probe_1invmAB5_L1JPSI-1M5-eEM15/HLT_IDTrack_Electron_GSF_1
HLT_e14_lhtight_e4_idperf_tight_probe_1invmAB5_L1JPSI-1M5-eEM15/HLT_IDTrack_Electron_GSF_1_el_probe
HLT_e14_lhtight_e4_idperf_tight_probe_1invmAB5_L1JPSI-1M5-eEM15/HLT_IDTrack_Electron_IDTrig
HLT_e20_idperf_loose_lrtloose_L1eEM18L/HLT_IDTrack_ElecLRT_FTF_HLT_Roi_FastElectron_LRT
HLT_e20_idperf_loose_lrtloose_L1eEM18L/HLT_IDTrack_ElecLRT_IDTrig_HLT_Roi_FastElectron_LRT
HLT_e26_idperf_tight_L1eEM26M/HLT_IDTrack_Electron_FTF_HLT_Roi_FastElectron
HLT_e26_idperf_tight_L1eEM26M/HLT_IDTrack_Electron_GSF
HLT_e26_idperf_tight_L1eEM26M/HLT_IDTrack_Electron_IDTrig
HLT_e26_lhtight_e14_idperf_tight_nogsf_probe_50invmAB130_L1eEM26M/HLT_IDTrack_Electron_FTF_HLT_Roi_FastElectron
HLT_e26_lhtight_e14_idperf_tight_nogsf_probe_50invmAB130_L1eEM26M/HLT_IDTrack_Electron_FTF_HLT_Roi_FastElectron_1
HLT_e26_lhtight_e14_idperf_tight_nogsf_probe_50invmAB130_L1eEM26M/HLT_IDTrack_Electron_FTF_HLT_Roi_FastElectron_1_el_probe
HLT_e26_lhtight_e14_idperf_tight_nogsf_probe_50invmAB130_L1eEM26M/HLT_IDTrack_Electron_IDTrig
HLT_e26_lhtight_e14_idperf_tight_nogsf_probe_50invmAB130_L1eEM26M/HLT_IDTrack_Electron_IDTrig_1
HLT_e26_lhtight_e14_idperf_tight_nogsf_probe_50invmAB130_L1eEM26M/HLT_IDTrack_Electron_IDTrig_1_el_probe
HLT_e26_lhtight_e14_idperf_tight_probe_50invmAB130_L1eEM26M/HLT_IDTrack_Electron_FTF_HLT_Roi_FastElectron
HLT_e26_lhtight_e14_idperf_tight_probe_50invmAB130_L1eEM26M/HLT_IDTrack_Electron_GSF
HLT_e26_lhtight_e14_idperf_tight_probe_50invmAB130_L1eEM26M/HLT_IDTrack_Electron_GSF_1
HLT_e26_lhtight_e14_idperf_tight_probe_50invmAB130_L1eEM26M/HLT_IDTrack_Electron_GSF_1_el_probe
HLT_e26_lhtight_e14_idperf_tight_probe_50invmAB130_L1eEM26M/HLT_IDTrack_Electron_IDTrig
HLT_e28_idperf_tight_L1eEM28M/HLT_IDTrack_Electron_FTF_HLT_Roi_FastElectron
HLT_e28_idperf_tight_L1eEM28M/HLT_IDTrack_Electron_GSF
HLT_e28_idperf_tight_L1eEM28M/HLT_IDTrack_Electron_IDTrig
```

- This is just a small subsample of the analyses that are configured for the ID Trigger Tier 0
  - In fact this is just for electrons, but it is **not even all the electron analyses**

Tier 0 continued …

- Why are there so many analyses ?
  - Perhaps monitoring too many different chains, but his is not really the issue …
  - We have many different signatures to monitor …
    - Electrons
    - Muons
    - Taus
    - B-jets
    - B-physics
    - Standard Jets
    - MinBias
    - Large Radius Tracking
      - Muons
      - Electrons
      - Taus
      - Fullscan Jets
- Many of these can use different reference collections:
  - Offline tracks or Offline large radius tracks, or combinations of both
  - Offline Electrons (offline electron tracks)
  - Offline Muons
  - Offline Taus tracks (1-prong, 3-prong)
  - …

- But don't forget, for Monte Carlo analyses, (uses the Tier 0 framework also for for the PhysVal monitoring) …
  - General truth particles
  - Truth Muons
  - Truth Electrons
  - Tracks from truth Tau decays
  - Exotic particles
  - …

- And above all, in the Trigger we have **several different track types in the trigger to evaluate**
  - Fast track finder tracks for each signature
  - Precision tracks for the Roi based triggers
  - GSF tracks for the electron chains
  - Full detector tracks for the jets, and b-jet preselection triggers
    - May be adding precision tracks for the b-jet preselection
  - MinBias tracks
  - Large radius tracks
  - Disappearing tracks

- Should not forget the vertices, but won't discuss them here

Tier0 analysis summary …

- IIn principle have
  - ~ 60 egamma analyses
    - 21 analyses of the FTF
    - 21 for the Precision tracking
    - 19 for the GSF tracking
  - ~ 142 muon based analyses
    - 40 Muon FTF
    - 40 Muon precision tracking
    - 2 MuonIso FTF - not muon tracking, but second stage muon isolation tracking (need more of these)
    - 2 MuonIso precision tracking
    - (~ around 40 are mu+tau tag and probe analyses, for TauCore_FTF, TauIso_FTF, Tau_IDTrig)
  - ~ 30 tau analyses
    - 10 TauCore_FTF - first stage tau reconstruction
    - 10 TauIso_FTF - second stage tau reconstruction (in wider Roi, but with restricted z range at the beam line)
    - 10 Tau_IDTrig
  - ~ 143 analyses of jets, either
    - FS jets
    - b-jet preselection
    - **b-jet FTF (Roi based)**
    - **b-jet Precision tracking (Roi Based)**
- The list goes on and on …

How to proceed

- Clearly, for the ID Trigger validation we need to reduce the number of chains that we monitor
- However, this is not a consideration, as even with **only a single chain for each type of analysis** we are still talking around **30 separate, distinct analyses**
- But a single chain would not be enough for each signature, or type of track collection, we have eg
  - low PT and high PT triggers
  - Triggers with preselection, trigger without preselection
  - Tag and probe triggers for the electrons, muon and taus
  - …

- The **absolute minimum** number of analyses that we would need to support to have a full spectrum validation for the ID Trigger, **is around 90**

- Any matching / storage solution has to be able to natively, and efficiency store the matching information for this many analyses
- Even though we have too many analyses at the moment, the framework **still needs to be able to support this many**
  - We can not have a framework that does not scale with the number of analyses and imposes limits on the number of allowed analyses just because of easily avoidable sub-optimal design decisions
- As importantly, each of these analyses will potentially be running in their own threads, so any framework needs to be able to simultaneously read and write different analyses **at the same time**

Storage information

- Because of the large number of potential analyses, then a matching / storage solution that stores information on s**ingle global collections** is not workable
- Example:
  - Store a single offline / truth track collection - want to store the “matched” test tracks as “decorations” on this global collection
  - OK for any scenario with only a few separate analyses - eg with 5 analyses, could store a vector of 5 tracks links on each tracks, one for each analysis
  - For the trigger, this is a non-starter
    - Would need to have potential up to 90 track links for every track
    - In practice there would be far fewer per track, most tracks woulds not be in a give Roi never for a specific analysis, some analyses for a specific track type might have a link for trigger Chain A, but not trigger Chain B and so on,
    - However, to make it work for the trigger, we would need to store an entire trigger navigation tree as a decoration on a track, ie a complex, branching tree structure would need to be duplicated for every item in a simple vector
    - The tree structure could be simplified by compacting information such as trigger chain, Roi type, Roi number, track collection type into a single flag
      - But then decoding the flag would be non trivial
    - Could use a more complicated structure, but then need to store IDs of chains, Rois etc, all in the navigation “decoration”
    - …

- The number of potential issues is large, but the potential for the “decorations” to be orders of magnitude larger than the original collection is plain.

How to proceed

- So the general principle that has top be followed is that each analysis **will store its own matching information**
  - The structure of the different analyses, which track collections they need, and so on is **already stored** in each analysis, so this information does not need to encoded elsewhere
- So the matchhing information must be stored with the analysis

- If people want to keep using decorations on the track collections then these decorations will need to be on a **copy of the reference track collection** for each specific analysis
  - Clearly using an actual copy is a non-starter - with over 90 potential analyses, we do not want 90 copies of the offline track collections
  - Would potentially need to use a ViewContainer on to the reference tracks collection, with the matching track link information stored on the view container

- Alternatively can use a new structure consisting simply of a vector of pairs of element links
  - One link links back to the global Offline track collection
  - One link links to the matching track for the analysis

- It is more logical to store the simple matching information within the (already existing) complex structure storing each analysis, rather than have a complex structure for the analyses, and a yet more complex structure on the tracks themselves

- NB: The complexity is in the trigger use case, so the design needs to address the trigger use case in the first instance

Example

- In the Trigger we make us of our own structured ntuple, Storing everything as
  - Event
    - Chain
      - Rois
        - Tracks
- Offline (or other full scan) collections are just single Roi “chains”, then we run all matching as part of the analysis

```
Event run: 451936     event: 243631408        lb: 179         bc: 1382        time: 1683855081        mu: 31.4157
      Chain Offline   rois: 1
             Roi  z: 0 (-225 - 225) eta: 0 (-5 - 5) phi: 0 (-3.14159 - 3.14159) (fullscan) RoIid: 0 RoIword: 0 (size 0)
                    [       eta=-2.29081    phi=-2.50252    z0=-21.2959     pT=-2.39158 GeV         d0=0.0189982    hp=0xf8073      chi2=25.6459/21         algo=5  bl=2:t:f        id=0x2c3bfb(
                    [       eta=2.35076     phi=1.48694     z0=-78.1592     pT=-1.92857 GeV         d0=0.604137     hp=0xf8077      chi2=15.2842/19         algo=5  bl=4:t:f        id=0x2c3bfc8
                    [       eta=2.29501     phi=1.60211     z0=-78.3974     pT=-3.93523 GeV         d0=0.788655     hp=0x7c077      chi2=38.9437/18         algo=5  bl=2:t:f        id=0x2c3bfe(
                    [       eta=-2.30124    phi=0.163501    z0=33.2869      pT=1.22025 GeV          d0=-0.242336    hp=0xf8037      chi2=24.6198/18         algo=5  bl=2:t:f        id=0x2bc224(
                    [       eta=2.22997     phi=1.71154     z0=-22.5707     pT=2.40036 GeV          d0=0.688701     hp=0x7c037      chi2=19.6648/17         algo=5  bl=2:t:f        id=0x2bc226(
                    ...
      Chain HLT_tau35_idperf_tracktwoMVA_L1TAU20IM:HLT_IDTrack_TauCore_FTF:HLT_Roi_TauCore          rois: 1
             Roi  z: 0 (-180 - 180) eta: -0.289879 (-0.389879 - -0.189879) phi: -1.07275 (-1.17275 - -0.97275) RoIid: 0 RoIword: 3184533290 (size 0)
                    [       eta=-0.288919   phi=-1.13859    z0=-81.2384     pT=-3.11117 GeV         d0=-0.0257964   hp=0x78f        chi2=45.0065/11         algo=0  bl=2:f:f        id=0x2bcaed8
                    [       eta=0.0468482   phi=-1.03825    z0=-80.8669     pT=-1.93801 GeV         d0=-0.0083678   hp=0x687        chi2=5.50693/7  algo=0  bl=2:f:f        id=0x2bcaef00:0xfff:
                    vertices: 0
      Chain HLT_tau35_idperf_tracktwoMVA_L1TAU20IM:HLT_IDTrack_TauIso_FTF:HLT_Roi_TauIso   rois: 1
             Roi  z: -81.1912 (-88.1912 - -74.1912) eta: -0.288919 (-0.688919 - 0.111081) phi: -1.13862 (-1.53862 - -0.738619) RoIid: 0 RoIword: 3184533290 (size 0)
                    [       eta=-0.147445   phi=-1.43065    z0=-80.8515     pT=-2.72107 GeV         d0=0.0225279    hp=0x78b        chi2=33.1896/11         algo=0  bl=2:f:f        id=0x2bcaf08
                    [       eta=-0.112779   phi=-1.16351    z0=-80.8643     pT=2.6317 GeV           d0=0.0501033    hp=0x78f        chi2=8.1677/11  algo=0  bl=2:f:f        id=0x2bcaf200:0xfff:
                    [       eta=-0.288919   phi=-1.13859    z0=-81.2384     pT=-3.11117 GeV         d0=-0.0257964   hp=0x78f        chi2=45.0065/11         algo=0  bl=2:f:f        id=0x2bcaf38
                    [       eta=0.037213    phi=-1.12606    z0=-74.919      pT=0.904538 GeV         d0=0.0273292    hp=0x78f        chi2=13.9531/10         algo=0  bl=2:f:f        id=0x2bcaf50
                    [       eta=-0.239056   phi=-0.948416   z0=-80.9445     pT=8.02508 GeV          d0=-0.0224689   hp=0x78f        chi2=8.54364/10         algo=0  bl=2:f:f        id=0x2bcaf68
                    [       eta=0.0468479   phi=-1.03825    z0=-80.8668     pT=-1.938 GeV           d0=-0.00836786  hp=0x687        chi2=5.50759/7  algo=0  bl=2:f:f        id=0x2bcaf800:0xfff:
                    vertices: 0
      Chain HLT_tau35_idperf_tracktwoMVA_L1TAU20IM:HLT_IDTrack_Tau_IDTrig:HLT_Roi_TauIso:HLT_IDVertex_Tau rois: 1
             Roi  z: -81.1912 (-88.1912 - -74.1912) eta: -0.288919 (-0.688919 - 0.111081) phi: -1.13862 (-1.53862 - -0.738619) RoIid: 0 RoIword: 3184533290 (size 0)
                    [       eta=-0.112448   phi=-1.16344    z0=-80.8762     pT=2.68571 GeV          d0=0.0482012    hp=0x78f        chi2=32.6678/37         algo=5  bl=2:t:f        id=0x2bcaf98
                    [       eta=0.0367936   phi=-1.12508    z0=-74.8799     pT=0.874409 GeV         d0=0.00916009   hp=0x78f        chi2=8.21185/10         algo=5  bl=2:t:f        id=0x2bcafb(
                    [       eta=-0.238882   phi=-0.94782    z0=-80.9532     pT=8.0649 GeV           d0=-0.0463545   hp=0x787        chi2=45.6846/40         algo=5  bl=2:t:f        id=0x2bcafc8
                    [       eta=-0.145276   phi=-1.42906    z0=-81.0089     pT=-2.86676 GeV         d0=-0.0248162   hp=0x78b        chi2=14.1884/18         algo=5  bl=2:t:f        id=0x2bcafe(
                    [       eta=-0.287456   phi=-1.13657    z0=-81.2024     pT=-3.39313 GeV         d0=-0.0913213   hp=0x783        chi2=48.3158/43         algo=5  bl=2:t:f        id=0x2b2380(
                    [       eta=0.0473358   phi=-1.03859    z0=-80.8587     pT=-1.88952 GeV         d0=-0.00156973  hp=0x687        chi2=35.6555/37         algo=5  bl=2:t:f        id=0x2b2381(
                    vertices: 1
                    [ x=-0.610656 +- 8.11367e-05    y=-0.455104 +- 9.46898e-05      z=-80.9565 +- 0.00199777;       chi2=16.0513/5  Ntracks=4       ]
                    ...
Event run: 451936     event: 243652967        lb: 179         bc: 2393        time: 1683855081        mu: 31.4157
      Chain Offline   rois: 1
             Roi  z: 0 (-225 - 225) eta: 0 (-5 - 5) phi: 0 (-3.14159 - 3.14159) (fullscan) RoIid: 0 RoIword: 0 (size 0)
                    [       eta=2.38643     phi=-3.0461     z0=-105.773     pT=-1.49582 GeV         d0=0.313794     hp=0xf8077      chi2=29.8194/17         algo=5  bl=2:t:f        id=0x2bc449(
                    [       eta=1.8597      phi=1.30561     z0=38.9517      pT=1.99475 GeV          d0=0.475224     hp=0x1f017      chi2=37.4124/45         algo=5  bl=2:t:f        id=0x2bc440(
                    [       eta=-2.09749    phi=-1.97597    z0=-99.251      pT=-1.06332 GeV         d0=-0.529055    hp=0x3c073      chi2=12.6338/15         algo=5  bl=2:t:f        id=0x2bc43e(
                    [       eta=2.06983     phi=2.43767     z0=65.28        pT=1.1787 GeV           d0=0.568255     hp=0x3c037      chi2=28.128/15  algo=5  bl=2:t:f        id=0x2bc43c80:0xfff
                    [       eta=-1.50637    phi=-1.86226    z0=-101.425     pT=-1.1358 GeV          d0=-0.415735    hp=0x780f       chi2=39.8166/50         algo=5  bl=2:t:f        id=0x2bc43b(
```

So …

- For the Filter offline tracks to within the Rois, match Roi tracks with the (filtered) offline tracks etc)
- If you want to store the actual matching information, then we would clearly need to encapsulate this information into the structure of the storage,. Ie Store information by
  - Event
    - (Global collections)
    - Analysis
      - Reference tracks View (with decorations to the matched test tracks)
      - Test tracks

- Or …
    - Analysis
      - Test tracks
      - Matching structure (pointing to global collection and Test tracks)

- How this should best be done is up for discussion, but any other proposal would need to efficiently handle the Trigger use case as the primary consideration, with offline as a simplified case, and not the other way around