DPDs and Trigger

Plans for Derived Physics Data
Follow up and trigger specific issues

Ricardo Gonçalo and Fabrizio Salvatore RHUL

Derived Physics Data

Why do we need them?

- Disk space at Tier2s
  - Tier2 sites will provide data access to most of ATLAS – this is not enough for many studies
  - Only a fraction of the space can be taken with ESD
  - DPDs can have some ESD-level data if on the whole they're small enough
- Centralized production – the train model
- Skimming should make analysis faster
  - Primary DPDs produced centrally at Tier1s to optimize resources

Skimming – remove un-interesting events
Slimming – remove un-needed containers/objects
Thinning – remove un-necessary object details

DPD task force

- Created 22 July to establish the requirements and address the design of the DPD
  - Led by Gustaaf Brooijmans and David Côté
- First prototype is working in the FDR-2c starting today
  - Needs release 14.2.20.3 + PATJobTransforms-00-00-22 + PrimaryDPDMaker-00-00-28
  - This will be a first prototype to be improved for initial running
- More details:
  - https://hypernews.cern.ch/HyperNews/Atlas/get/primaryDPDcontent/34.html
  - http://indico.cern.ch/conferenceDisplay.py?confId=39971
- Trigger represented by Xin and me; Fabrizio also part of DPD Task Force

The story/plan so far…

- Only concerned with primary DPD
  - D^nPD really means ntuples obtained from DPD
- Produce 8 “Performance DPDs” at Tier0 and Tier1 for detector/reconstruction performance studies
  - Physics DPDs to be defined later – not immediate priority and would benefit from more experience
- Design based on (trigger) streams
  - Rely somewhat on stream overlap ≈ 10%
  - not producing 8 DPDs × N streams! Count on 1 DPD/stream + few extra requests
- Data formats in Tier2 disk:
  - 25% Simulation
  - 25% AOD
  - 25% DPD
  - 25% the rest: BS, ESD, user data (job output) – requested by each group
- Sizes:
  - AOD ≈ ⅕ ESD
  - Performance DPD ≈ 10% ESD ≈ ½ AOD
  - AOD : around 50 kB/event for ttbar with 10^{31} no Bphysics menu – assume 30 kB/streamed event
  - Perfrormance DPD: bottom line is around 15 kB / (event × stream efficiency × skimming efficiency) assuming 1 DPD per trigger stream

Performance DPDs

- Strategies:
  - "RoI DPD" – keep heavy objects only from regions of interest
    - E.g. calorimeter cells in a cone around an electron
    - Typical "efficiency" wrt full ESD found to be ~20%
  - Skim events to keep signal enriched sample (W/Z signal/standard candles, high-p_T, tight PID)
  - Prescale: keep only a fraction of events in some DPDs

- Heavy and Light performance DPDs
  - Heavy:
    - ESD->DPD at Tier0
    - ESD-level information for detailed performance studies
    - 8 heavy DPDs ≅ 80% AOD volume
  - Light:
    - AOD->DPD at Tier1
    - AOD-level information for quick exploration
    - 8 light DPDs ≅ 10% AOD volume
    - DPD Light intended as basis for Physics DPD

- DPD making:
  - New package `PhysicsAnalysis/PrimaryDPDMaker` to organise code for primary DPDs
  - Use `Tau/Egamma/XXXDPDMakers` for D^2PD, D^3PD only

- Bookkeeping:
  - Use EventBookkeeper being deveoped within PAT to keep track of skim efficiencies, etc
  - Need to address event duplication from stream overlap

Performance DPD Proposal

	Trigger Stream	Event Selection	Heavy Content	Usage
#1	EGamma	high p_T e/γ OR pre-scale	ROI ESD	e/γ reco+ID, τ fake
#2	Muon	high p_T μ OR pre-scale	ESD of μ+τ	μ reco+ID, τ fake
#3	Jet	high p_T jet OR pre-scale	full ESD	clustering, jets, MET, lepton fakes, b-tag
#4	EGamma	Z/W/JPsi → e medium	full ESD	e/τ signal, MET
#5	Muon	Z/W/JPsi → μ medium	full ESD	μ/τ signal, MET
#6	Jet	Z→τ med OR W→MET	ROI ESD	τ signal
#7	EGamma	γ tight OR Z→ee	full ESD	jet calibration (γ/Z+jet)
#8	Min Bias	pre-scale	full ESD	jet calibration, MET, LArg

➡ Should DPD #3 be split in two (tracking-oriented & calo-oriented) ?

All AOD's are also available from all trigger streams.

David Côté - Physics Coordination - 20 August 2008 6/10

See PhysicsAnalysis/PrimaryDPDMaker/python/PrimaryDPD_OutputDefinitions.py for list of objects

DPDs and Trigger

- Trigger slices interested in using Combined Performance DPDs
- Important to have enough trigger info to also allow trigger studies
- Monte Carlo DPDs:
  - DPD relies on streaming
  - Streaming relies on Trigger
  - For Monte Carlo data rely on trigger simulation

Trigger information in DPDs

- Input from slices and physics groups collected: reported by Fabrizio in the July 23rd Menus meeting
  - Which triggers passed/failed, prescale, passthrough
  - Enough info to allow tag-and-probe (match trigger and offline objects)
    - Benchmark for minimal trigger info being requested!
  - Possibility: use for selection optimisation

- Trigger navigation:
  - For 2008 run: store as is, no thinning
  - For 2009 run:
    - Thin down to contain only requested chains
    - Tomasz and Harvard group interested
    - Not clear how to deal with trigger features attached to deleted chains
    - Configuration information to remain unslimmed

- Matching between offline and online objects
  - Quite some interest from physics groups
  - Existed in EventViewTrigger
  - Can we provide common solution? Carsten Hensel working on this

Preliminary list of trigger objects in DPDs

- Based on the experience of the SUSY DPD
  - Baseline for starting the discussion
  - Will grow once different WGs will agree on using same DPDs (e.g. SUSY+Exotics)

```
# Trigger objects
if WriteTrigDetails :
    Stream.AddItem( ["TrigDec::TrigDecision#TrigDecision*"] )
    Stream.AddItem( ["HLT::HLTResult#HLTResult_L2"] )
    Stream.AddItem( ["HLT::HLTResult#HLTResult_EF"] )

    if WriteTrigObjects :
        Stream.AddItem( ["LVL1_ROI#LVL1_ROI*"] )
        Stream.AddItem( ["CombinedMuonFeature#HLTAutoKey*"] )
        Stream.AddItem( ["TrigEMCluster#HLTAutoKey*"] )
        Stream.AddItem( ["TrigT2Jet#HLTAutoKey*"] )
        Stream.AddItem( ["TrigElectronContainer#HLTAutoKey*"] )
        Stream.AddItem( ["TrigPhotonContainer#HLTAutoKey*"] )
        Stream.AddItem( ["TrigTau#HLTAutoKey*"] )
        Stream.AddItem( ["JetCollection#HLTAutoKey*"] )
        Stream.AddItem( ["TrigMuonEFContainer#HLTAutoKey*"] )
        Stream.AddItem( ["egammaContainer#NoIDEF_RoI*"] )
        Stream.AddItem( ["egammaContainer#egamma_RoI*"] )
        Stream.AddItem( ["egDetailContainer#HLTAutoKey*"] )
        Stream.AddItem( ["Analysis::TauDetailsContainer#HLTAutoKey*"] )
        Stream.AddItem( ["Analysis::TauJetContainer#HLTAutoKey*"] )
        Stream.AddItem( ["TrigMissingET#HLTAutoKey*"] )
        Stream.AddItem( ["TrigRoiDescriptor#HLTAutoKey*"] )
```

TrigDetails

TrigObjects

Menus meeting
23/7/08

Trigger objects currently in DPD

Configuration and steering:
TrigConf::Lvl1AODPrescaleConfigData#AODConfig-0
TrigConf::Lvl1AODPrescaleConfigData#AODConfig* duplicate
TrigConf::HLTAODConfigData#AODConfig*
TrigConf::Lvl1AODConfigData#AODConfig*
TrigDec::TrigDecision#TrigDecision*
HLT::HLTResult#HLTResult_EF
HLT::HLTResult#HLTResult_L2
TrigRoiDescriptorCollection#HLT
TrigRoiDescriptorCollection#HLT_T2TauFinal
TrigRoiDescriptorCollection#HLT_TrigT2CaloEgamma
TrigRoiDescriptorCollection#HLT_TrigT2CaloJet
TrigRoiDescriptorCollection#HLT_TrigT2CaloTau
TrigRoiDescriptorCollection#HLT_forMS

Level 1:
LVL1_ROI#LVL1_ROI
LVL1::JEMRoI#JEMRoIs
LVL1::TriggerTower#TriggerTowers
LVL1::JEMEtSums#JEMEtSums

L2 muons:
CombinedMuonFeature#HLT_egamma
CombinedMuonFeature#HLT
MuonFeatureContainer#HLT
TrigMuonEFContainer#HLT_MuonEF

L2 Jets:
TrigT2Jet#HLTAutoKey* obsolete
TrigT2JetContainer#HLT_TrigT2CaloJet

L2 Missing ET:
TrigMissingETContainer#HLT_T2MissingET

L2 tracks:
TrigInDetTrackCollection#HLTAutoKey*
TrigInDetTrackCollection#HLT
TrigInDetTrackCollection#HLT_TRTSegmentFinder
TrigInDetTrackCollection#HLT_TRTxK
TrigInDetTrackCollection#HLT_TrigIDSCAN_eGamma*
TrigInDetTrackCollection#HLT_TrigSiTrack_eGamma*
TrigInDetTrackCollection#HLT_TrigIDSCAN_Tau
TrigInDetTrackCollection#HLT_TrigSiTrack_Tau
TrigInDetTrackCollection#HLT_TrigIDSCAN_Muon
TrigInDetTrackCollection#HLT_TrigIDSCAN_Jet
TrigInDetTrackCollection#HLT_TrigIDSCAN_eGamma

L2 egamma:
TrigEMCluster#HLTAutoKey* obsolete
TrigEMClusterContainer#HLT*
TrigEMClusterContainer#HLT_TrigT2CaloEgamma
TrigElectronContainer#HLTAutoKey* obsolete
TrigPhotonContainer#HLTAutoKey* obsolete
TrigElectronContainer#HLT_L2IDCaloFex
TrigPhotonContainer#HLT_L2PhotonFex

L2 taus:
TrigTau#HLTAutoKey* obsolete
Trigtauclustercontainer#HLT_TrigT2CaloTau
TrigTauContainer#HLT
TrigTauTracksInfoCollection#HLT

EF taus:
Analysis::TauJetContainer#HLT_TrigTauRecMerged
Analysis::TauJetContainer#HLT_TrigTauRecCalo
Analysis::TauDetailsContainer#HLT_TrigTauDetailsCalo
Analysis::TauDetailsContainer#HLT_TrigTauDetailsMerged
Analysis::TauDetailsContainer#HLTAutoKey* duplicate
Analysis::TauJetContainer#HLTAutoKey* obsolete
Analysis::TauJetContainer#HLT_TrigTauRecMerged
Analysis::TauJetContainer#HLT_TrigTauRecCalo
TrigTauTracksInfoCollection#HLT

EF Missing ET:
TrigMissingET#HLTAutoKey*
TrigMissingETContainer#HLT_TrigEFMissingET

EF Jets:
JetCollection#HLTAutoKey* obsolete
JetCollection#HLT
JetCollection#HLT_TrigJetRec

EF Muons:
TrigMuonEFContainer#HLTAutoKey* obsolete

EF egamma:
egammaContainer#HLTAutoKey* obsolete
egammaContainer#NoIDEF_RoI*
egammaContainer#egamma_RoI*
egammaContainer#HLT_egamma
Egdetailcontainer#HLTAutoKey*
egDetailContainer#HLT_egamma*

EF tracks:
Rec::TrackParticleContainer#HLTAutoKey* obsolete?
VxContainer#HLT_PrimVx
Rec::TrackParticleContainer#HLT_InDetTrigParticleCreation_Tau_EFID

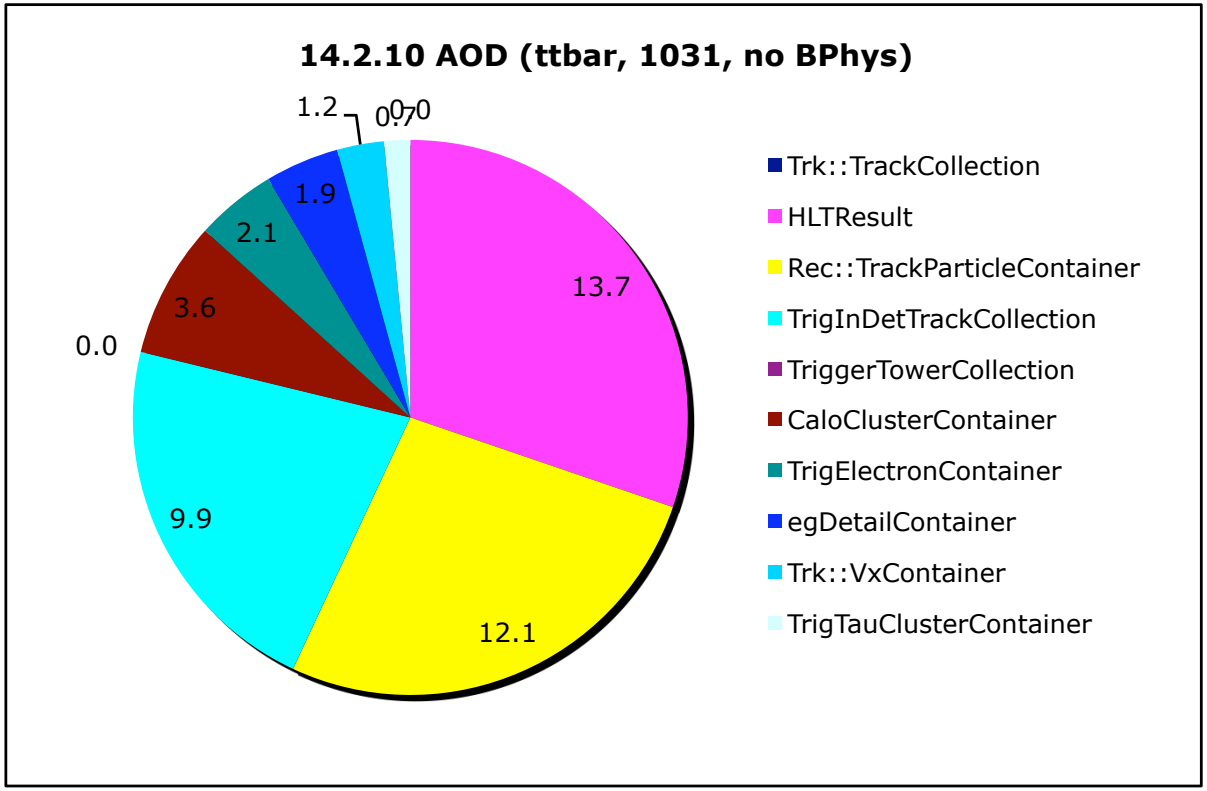
See PhysicsAnalysis/PrimaryDPDMaker/python/PrimaryDPD_OutputDefinitions.py for more details and offline objects
Please let us know what's missing and what's obsolete (we need confirmation)

Conclusions

- First prototype of Performance DPDs being tried out right now in FDR-2c
- All trigger information being kept
  - We need to know what is missing and what is old

- EDM migration allowed improvements in data size on file
- Overall size of trigger data depends strongly on data type and on menu
- Running menu for L=10^{31} $cm^{-2}s^{-1}$ (no Bphysics):
  - AOD total size: 48 kB/event
  - ESD total size: 86 kB/event
- The HLTResult may be reduced further to 1-2 kB/event by slimming out navigation information - (e.g. for inclusion in DPDs)

13.7 HLTResult
12.1 Rec::TrackParticleContainer
9.9 TrigInDetTrackCollection
3.6 CaloClusterContainer
2.1 TrigElectronContainer
1.9 egDetailContainer
1.2 Trk::VxContainer
0.7 TrigTauClusterContainer



Seth Caughron,
July 30 Trigger
Menu Meeting

EF rates by group

	13.0.40 (Full Min Bias)		14.2.10.7 (Enhanced Bias)	
	rate (Hz)	error (Hz)	rate (Hz)	error (Hz)
Jets	**36.87**	**0.02**	**24.7**	**0.045**
B-jets	14.12	0.06	11.7	0.054
electron	47.1	0.1	56.9	0.71
photon	9.13	0.02	10.4	0.023
tau	33.4	0.07	18.9	0.046
muon	23.6	0.4	77.2	1.58
XE	3.79	0.009	1.40	0.0069
TE	0.924	0.004	1.19	0.0037
JE	1.63	0.05	1.77	0.063
combined	48	1	53.0	0.64
TOTAL	186.74	0.08	229	0.56

Red: Rates are higher
Black: Rates are (relatively) stable
Blue: Rates are lower

Muon slice stable at L1 & L2, but rate triples at EF

EF rate appears to have fluctuated high