

# PAT Intro/Status

Amir Farbin  
Karsten Koeneke

# Overview

- Goal give a general picture of what we have been/plan to focus on recently/near future
- A few active projects/areas:
  - DPD Production
  - Tools
  - TAG
  - D3PD Making
  - Post Athena- D3PD Reading
  - Core
  - Storage of time-dependent information

# DPD Production

- Goal: Single DPD Producing Train simultaneously producing (via Production System) all group DPDs (POOL-based and D3PD) every 2-4 weeks.
- Motivation:
  - Reduce CPU usage
  - Better organization: well defined time-line for sw cache, validation, and production
  - Reduce required manpower
  - Establishing path to larger number of smaller/more targeted DPDs.
- Plan Outlined in: <https://twiki.cern.ch/twiki/pub/AtlasProtected/PhysicsAnalysisTools/DPDProductionPlan.pdf>
- Short-term steps in this direction:
  - Establishment of a single AtlasPhysicsAnalysis Cache with PAT coordinator (Annick Lieres)
    - See Karsten's talk on Tuesday (<https://indico.cern.ch/getFile.py/access?contribId=36&sessionId=7&resId=0&materialId=slides&confId=119169>)
  - D3PD Tag submission/validation procedure (defined by Attila): <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/D3PDMakerTagSubmission>
  - SUSY DPD: working towards the DPD Train "cart" (David Cote, Louise, ...)
- Next step: Common Object Selection Sequence/Containers (see Karsten's talk today)
- Related issue: Removing cross talk between D3PD Makers (Attila)

# Exotics Survey

- Due to large number of physics signatures the Exotic group has been parasitically relying on various DPDs from other groups (eg W/Z D3PD for W'/Z' search)
- Problem: since DPD developers not in Exotics group, they aren't motivated by the Exotic group timelines... so Exotics analysis don't get the DPDs they need on time.
- Discussion Between Junji, Beate, Jamie, Karsten, Amir... conclusions:
  - The organizational structure behind the DPD Train will eventually mitigate these issues.
  - But, we cannot fix this problem from top-down.
  - Each analysis needs to be proactive in making sure their DPDs are ready.
- Exotics suggestion: centralized space for all D3PDs
  - We didn't think this was a good idea... the group space quotas are an important pressure leading to the evolution of the DPDs.

# Tools

- First focus has been cleaning and selection tools.
- Stage 1 (last month): Louise gathered info on all cleaning and selection tools from Liaisons:
  - <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/CPObjectSelection>
  - <https://indico.cern.ch/getFile.py/access?contribId=3&resId=0&materialId=slides&confId=130277>
  - Liaisons: Let's keep the wiki page up-to-date.
- Stage 2 (now): First design of the Selector Tools settled this week (Karsten, Attila, Amir)
  - Very simple tools that will work in Athena, ROOT, and ARA
  - See Karsten's Talk this afternoon
- Stage 3 (next month): Combined Reco Groups should rewrite/wrap their tools to be consistent with stage.
  - Liaisons will need to help us make sure this work gets done.
- Stage 4 (summer): Use the standard tools to create Common Object Selection Sequence/Containers for DPD/TAG making.
- Stage 5 (???): Migrate tool inputs (constant, histograms, etc) into IOV-dependent structure which is distributed in a common way (eg via CVMFS)

# TAG

- PAT's Problem Child:
  - Lots of effort and promise... but effectively unused by physics community
  - Event Picking used by Data Prep and Combined Performance
  - Perpetually behind in validation, problematic for production, and inconsistent with newest object selections.
- Strategy, let's make the physics case:
  - New effort to demonstrate performance/disk benefits of TAG, using SUSY DPD making as concrete example (Jack Cranshaw, Louise, David C.)
    - Some groups are producing DAODs to speed up D3PD iterations... but then they fill disks with DAODs...TAGS could eliminate this problem.
  - Plan to simplify TAG building by using the D3PD framework.
    - Core elements will be implemented by Attila by May...
    - Need manpower for actual reimplementation of TAG via D3PD
      - Perhaps this person can then be the TAG content gatekeeper
  - New Manpower in TAG Validation (Wolfgang Ehrenfeld, James Dassoulas)
  - DPD Production Plan: Rebuild TAGs with the DPD so it is up-to-date with newest/greatest selections and DPD streams can be monitored/reproduced via the TAG
- Discussion about TAG today...

# D3PD Makers

- This activity is now much better organized thanks to Attila... more from him.
- Recent Developments:
  - New SW Tag submission procedure
  - New Configuration which makes D3PD content configuration look like POOL output stream creation
    - Unifies DPD/D3PD configuration
- Some Planned Release 17 Developments:
  - TAG building support
  - Associations (eg matching) via AssociationMaps (created by Athena Tools)
- More about status for Attila Today

# DP3D Reading

- Several Projects aimed at improving D3PD-based analysis:
  - D3PD Skeletons (MakeClass replacement) (Attila)
    - Provides common mechanism for accessing branches and event looping which take advantage of all known optimizations
    - Also presents simple object classes (eg Electron, Muon, ...) which can be used in the interface of ROOT based tools Tools (eg Selectors).
    - [https://twiki.cern.ch/twiki/bin/view/AtlasProtected/D3PDMakerReader#Generating\\_source\\_code\\_in\\_an\\_Ath](https://twiki.cern.ch/twiki/bin/view/AtlasProtected/D3PDMakerReader#Generating_source_code_in_an_Ath)
    - We hope to distribute skeletons for SUSY and Top DPDs via CVMFS soon.
  - D3PD Package Environment (Nils):
    - First step: Script for checking out/building all known ROOT-based tools used in analysis: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/D3PDLibs>
    - Next: Defining the layout of these packages so we can handle dependencies and build a single library
  - D[3]PD Book-keeper (Nils):
    - Scripts that check DPD vs AMI for consistency (eg NEvents) and relevant meta-data (eg x-section) and fills an local object that holds all this data.
    - Design settled (<https://indico.cern.ch/getFile.py/access?contribId=6&resId=0&materialId=slides&confId=128412>), waiting for Prototype.



# Core

- CompositeParticle now use IN4M rather than iParticle
- Cutflow service Book-keeping fixed in 16.6.3.4
- Karsten and Wim worked out (yesterday) how to get UserDataSvc to merge UserData when merging outout

# Time-dependent Information Storage

- Time-dependent = IOV (or LB)
- Examples: Efficiencies, Data/MC Scale Factors, Resolutions, Scale, Inputs to Selectors/Cleaning Tools (eg OTX maps)
- Most of these are now stored/passed around in non-ideal ways.
- Some Discussion Early in the year:
  - Agreement on Potential of InSituPerformance
    - Focus on the Storage/Book-keeping features (eg the classes representing efficiencies) and COOL interface not on the analysis (eg Tag Probe)
  - CVMFS Identified as the most appropriate means of sharing this sort of data
    - Already used for Distribution of COOL DB
  - Access to info outside of Athena:
    - Everything is in ROOT files, so just a matter of creating ROOT side tool to find the appropriate file
- Seems like the experts/advocates got very busy since early discussions... should follow up

# Moving the PAT Meeting

- Very difficult to find workable (ie afternoon) time which doesn't overlap with any the top-level Physics or Comb Reco WG meetings
- Some Liaisons can't attend both PAT and WG meetings as they are required.
- After doodle poll seems like Thursday 15:00 is the least problematic time.
- Let's discuss this afternoon and definitely decide.