TDR software and production

David Rousseau (CPPM)

A bit early to draw lessons from the big TDR effort. Here is what happened in the last months in DICE and ATRECON and around:

- Simulation
- Single detector reconstruction
- Combined reconstruction
- Production

Note: not an attempt to summarize the full Phys. TDR Volume I (detector performance and combined performance)!

Simulation

- **DICE was mostly frozen in February 1998 after the various detector TDR's were submitted**
  - **Notable exception of the muons with new layout, matter and digitization (M2.8 database) available in October 1998**
    ⇒a number of production was done with obsolete layout (notably H→ xx for b-tagging studies, some B physics studies)
    ⇒parametrised muon reconstruction efficiencies had to be used in some cases
    ⇒but a (sufficient) number of muon channel have been correctly simulated
  - **Bug fixes in calorimeter and improved TILE digitisation**
- **Detector changes since then have not been implemented in production version**
- **CVS version has not been used**

Pile-up

- New pile-up method has been used in the calorimeters (Stefan Simion)
  - use full electronic shaping
  - use pre-computed calorimeter matrices
- Allow also electronic noise with or without optimal filtering
- Implemented together with Inner Detector pile-up (S.S+Monika Wielers) to take correctly into account correlated noise in calorimeters and pile-up tracks in tracker
- Muon pile-up done in standalone

Recent updates in ATRECON

No big change in single detector reconstruction since ~1 year. Improved outputs in RECB bank and Combined Ntuple (CBNT). Brief summary:

- **Inner Detector reconstruction:**
  - **common clustering algorithm used by IPATREC+PIXLREC+XKALMAN**
  - **IPATREC and XKALMAN widely used, less IPATREC vs XKALMAN studies than in ID TDR (notable exception of b-tagging)**
  - **PIXLREC used for some specific studies**
  - **XKALMAN++ tested in atrecon framework (ZEBRA input and output), same performance as XKALMAN, used to study tracking in non-uniform B field (no surprise). Will be moved into srt/cvs.**
- **Calorimetry:**
  - **Implement JetFinder library to study different jet algorithm**
  - **More detailed outputs:**
    - **TILE calorimeter cells**
    - **calorimeter projected matrix (possibility to run different jet reconstruction algorithm on RECB output)**

Recent updates in ATRECON (II)

- **MUONBOX:**
  - **better pattern recognition and fitting implemented (improvement at low p_T)**
  - **muon parameters given at Muon System entry, Calorimeter entry, vertex,**
  - **with correct covariance matrix, correct parametrization of energy loss and multiple scattering in calorimeters.**
- **Timing of modules ($\rightarrow$ Ntuple)**
- **Atrecon ported to Linux (fully) and to Windows NT (fortran code only)**

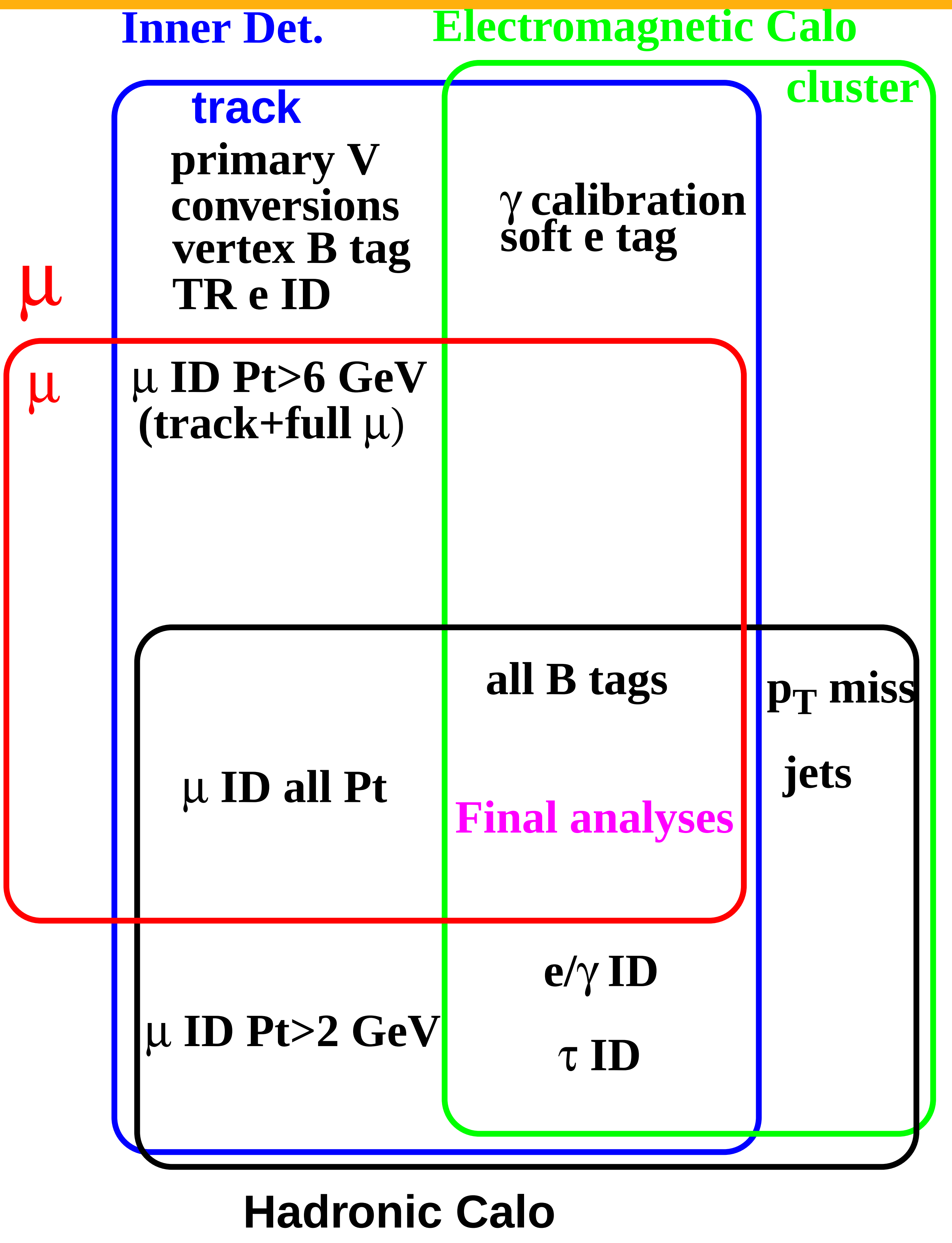
Timings

- **Timings obtained on shift, scaled by 7 to obtain SpecInt95 seconds**

Event	Muon	Calo	xKalman Full event	xKalman Seed	iPatRec Seed
$b\bar{b} \to \mu_6 X$	238	9	26	-	-
$b\bar{b} \to \mu_6 X$ low lumi	238	9	99	-	-
$WH(100) \to \mu b\bar{b}$	154	12	39	26	9
$WH(100) \to \mu b\bar{b}$ high lumi	170	93	2660	366	40
$H(130) \to e^+ e^- \mu^+ \mu^-$	242	10	24	7	4

- **WARNING: timing is very dependent on exact control parameters used (thresholds, window size)**
- **all algorithms have been optimised for efficiency and precision, not for CPU time**

Combined reconstruction



Combined reconstruction

- Ideally, the combined reconstruction tools studied by the combined performance groups would have been integrated in Atrecon to be used in physics studies.
- Given the parallel development of combined reconstruction and physics studies, this was only partially the case.
- When ready in time tools were integrated in Atrecon, with output stored in RECB Zebra banks and in Combined Ntuple
- Some tools were provided as routines to be run on Combined Ntuple
- Some tools were provided as standalone routines

Vertexing

- Conversion finding (xconver+xhourec for late conversions) implemented in ATRECON and widely used
  - Few % loss of efficiency at high luminosity
- K_S^0 and secondary vertex fitting (B-physics): CDF package CTVMFT interfaced to Atrecon (Francesco Tartarelli) and widely used in B physics community (good results)
- Primary vertex finding: CTVMFT plus algorithm from F.T., not implemented in Atrecon (simple parameterisation available)

e/γ identification

- Electron TR ID: available as a flag (80% or 90%efficient) available in CBNT or as a routine.
- E.m clusters basic quantities (energy and position) corrected from S-shape effect etcetera was widely used.
  - Some eta dependent and (to a less extent) p_T dependent calibration factors are still needed from the user
- Identification flag (ISEM) combining calorimeter quantities widely used.
  - Optimised for p_T>20 GeV and low luminosity, not optimal elsewhere
  - A single yes/no flag when different efficiency/ rejection optimization are needed, depending on the analysis
- Combined (track+calo) identification implemented but not really used
  - Difficult to provide ID flag and calibrated energy from a few GeV to infinity, for all rapidity and luminosity

e/γ identification (II)

- **Soft electron identification (p_T>2 GeV) using calorimeter quantities around extrapolated track (Michal Seman)**
- **Available as a private code**
- **Used on B events (J/$\psi \rightarrow e^+e^-$) and inside B-jets**
- **Attempt output calorimeter cell around track in Combined Ntuple (Anna Kaczmarka), but too big$\Rightarrow$had to use another Ntuple, meanwhile allowed reoptimisation of the cuts (but very big ntuples (~150Mb) uneasy to handle)**
- **No attempt for a smooth transition around 6 GeV between soft electron (track based) and hard electron (cluster based)**

Combined muon measurement

- **Two approaches:**
  - **STACO (Igor Gavrilenko): "simple" statistical combination of MuonBox track and covariance matrix (propagated to the beam-line) and XKALMAN track**
  - **set of routines to be run on combined ntuple (fast optimization loop)**
  - **95% overall efficiency (average on $H \rightarrow \mu^+\mu^-\mu^+\mu^-$)**
  - **MUID (Alan Poppleton+Dimitrios Fassouliotis): refit of ID hits with MuonBox track and covariance matrix**
  - **need rerun of (slow) full reconstruction because it needs access to the hits**
  - **can use measured energy loss in calorimeters**
  - **92% overall efficient (worse at small p_T) but less tails than STACO**

  $\Rightarrow$Improved p_T resolution dominated by Muon System above 50 GeV, by ID below

- **$K/\pi \rightarrow \mu$ decay rejection studied with STACO: rejection of 2/3 of the decays is reachable**

Combined muon identification (II)

- **Low p_T (>2 GeV) muon identification:**
  - **using last segment of TILE calorimeter (Gerard Montarou, Dario Barberis):**
  - **used in B physics ($J/\psi \rightarrow \mu^+\mu^-$) and B-tagging (inside jet)**
  - **use of hits in Muon Inner Station not tried**
- **Combined muon identification and measurement from a few GeV to infinity not available in one given frame**

Jet identification

- **Jet τ identification (Donatella Cavalli, Silvia Resconi)**
  - available in CBNT (additional variables in jet block)
- **Jet B-tag (Dario Barberis, Eduardo Ros)**
  - vertex B-tag with IPATREC or XKALMAN reassessed. Better results than before (R_u~350 instead of R_u~90 for ε_b=50%).
  - soft lepton ID brings little improvement
  - code to compute B-tag variable from CBNT is available

Some (biased) lessons from CBNT

- **Was used more in some groups than others**
- **Was used both for optimization of combined performance tools (e.g. e/γ separation, STACO combined muon measurement) and for physics studies (e.g. $H \rightarrow l^+l^-l^+l^-$, top)**
- **For physics studies, one would have liked to write a new ntuple (with e.g. H candidate mass, and selected leptons) from CBNT. Not possible with PAW unless one rewrites an entirely new ntuple.**
- **50.000 hard-limit on number of variables per event was annoying because it needs:**
  - **careful tuning of the content of the ntuple to avoid reaching this limit**
  - **merged variables are not easy to use (lots of INT and MOD)**
- **small size of ntuple was appreciated for export**
- **combination of full and fast reconstruction**
- **$\Rightarrow$ interest of a combined *something* was demonstrated but *something* should be better than HBOOK+PAW**

Productions

- **Simulation (coo. by Maya Stavrianakou):**
  - **Lots of different channels (also some with pile-up) have been simulated at CERN and in outside institutes (see web page Atlas Computing$\rightarrow$ Production)**
  - **CPU time has not been a problem, but human time to look at the data has been**
  - **People who have made private production should make sure they are listed in the web page above for future reference**
- **Reconstruction:**
  - **Mostly done on a private basis, as combined reconstruction tools were getting ready.**
  - **Some "big" productions used by more than one person (e.g $Z\rightarrow e^+e^-$, muons)**
  - **CPU time has been a problem given the short delay**
  - **Data on CPU time, output files should be gathered for future reference**

Conclusion

- TDR software was ready on time.
- Salient holes were successfully filled in the last months.
- Big effort from (not so) many people.

Next (=before transition to OO):

- Collect and archive information on productions (simulation probably complete, reconstruction to be updated)
- Gather private but useful code
- Gather impressions and comments from TDR workers while they are still hot.
  - what they liked/disliked
  - remaining holes
- Check DICE CVS version
- Detector Geometry has evolved since Dice was frozen (e.g. pixel layout, TRT modular geometry, Muon CSC,...). Should it be updated now, or shift to GEANT4 before?
- Some bug fixes to Atrecon need be implemented