Some notes on ezTree and EMC data in MuDst

Marco van Leeuwen, LBNL

Introduction

Last collaboration meeting:

Jan Balewski introduced ezTree

(Developed with Piotr Zolnierczek)

Main features

- Fast I/O: quickly (~100 Hz) loop over EEmc raw data
  - Good for monitoring of data-integrity
  - Useful for status tables, pedestals etc
- Can be generated promptly from daq-files
  - Pseudo on-line (~20 minutes delay) checks possible

Question (Jerome): Do we really want another data-format in Star??
But also: - Can we improve the MuDst reading speed
- How are these things done for Barrel?
- Can we learn from each other? Unify procedures and formats?

To work

Activists
- Jan Balewski
- Alex Suaide
- Wei-Ming Zhang
- Jerome Lauret
- MvL

1) Increase reading speed of MuDst (MvL)
   - Factor of ~10 gain for reading EMC-only, speed close to ezTree

3) Integrate ezTree into MuDst (Wei-Ming, Jan)
   - Explore possibility to create ezTree fast off-line
   - Maybe replace existing Emc part of MuDst with ezTree?

5) Streamline analysis-path (Alex Suaide)
   - Mainly for Barrel: drop requirement to run StMuDst2StEventMaker

Plus: discussion in several phone-meetings about what exactly to store for analysis-level etc.

Results

- For analysis: keep StMuEmcCollection (with internal changes for speed)
  - Backward compatibility (important for Barrel)
  - Zero-suppressed SMD saves space
  - Reading speed now comparable to ezTree
- For fast-offline (calibration): special ezTree-mode of MuDstMaker
  - Will be used to extract trigger, Endcap, Fpd data on-line during run-5
  - No zero-suppression, so pedestals (& mapping) not needed on-line

Recap: What is stored on MuDst?

- All tower ADC
  (except for corrupted crates)
- SMD ADC only channels > 1.5 σ above pedestal
  (all for endcap, for now)

This is as it used to be, except:
- Rudimentary support for clusters and points was dropped (data are not filled)
- Small modification of TTree structure
  New form allows split reading of e.g. Endcap tower only
- Internal optimisation of StMuEmcCollection: factor 5 speed gain
  (When looping over SMD hits)
- Addition of crate status flags to signal crate corruption (set by StEmcRawMaker)

Note: if you do not use (global) track info, do not read it! Saves factor 5 or so in speed

ezTree mode

Special ‘ezTree’ mode to store all raw (E)Emc & trigger detector data

- No zero-suppression for Smd

During the upcoming run: fast-track procedure to extract all Fpd, EEmc and trigger data to ezTree MuDst

Main purposes:
- Fast-offline QA (checking data integrity)
- Determining pedestals, gain, status

ezTree mode has different data format, not compatible with MuDst-analysis code

Note: Existing barrel on-line QA based on daq files, so not strictly needed in fast-extract

(Wei-Ming & Jan)

More on processing speed

It is no longer necessary to convert the full MuDst into StEvent (StMuDst2StEventMaker) for Emc analysis

The StEvent structures (StEmcCollection) are still used for storing clusters and points, but this is now directly filled from MuDst (by StEmcAdcToEMaker)

Saves memory and processing time

Thanks to Alex Suaide

The End

More on filesize

Numbers of channels

	Endcap	Barrel
Towers	720	4800 (2400)
SMD 1&2	6912 (2304)	36000 (18000)
Pre, post	2160 (720)	4800 (2400)

(installed FY04 in parenthesis)

Many channels! (dominated by SMD)
Total raw size per event: 83kb+headers
All tracks in p+p: ~1kb

Zero-suppression for barrel-SMD very effective: only storing few 100 channels in p+p event

More on filesize

E-mail about endcap filesize:

Here is the size of the endcap ezTree branches:

branch: EztETow 40835
branch: EztESmd 204799

The same info is in the following branches in my private format (which is similar to the bralle format and backward compatible with the existing mudst):

branch: EEmcPrs 28175
branch: EEmcSmdu 59535
branch: EEmcSmdv 59073

*Br 4 :EmcTow.mEndcapTowerADC[1080] : mEndcapTowerADC[EmcTow_] *
*Entries : 73 : Total Size= 78926 bytes File Size = 30470 *
*Baskets : 1 : Basket Size= 65536 bytes Compression= 2.14 *

As you can see, the tower info in my file is 30kb, compared to 40kb for your format. I assume that the preshower is included your EztESmd, so all other stuff together is 200kb in your case and 150kb in my case

Introduction

- Main features of Emc analysis:
  - Need large sample of data to calibrate towers
  - Emc tower+smd clustering algorithms still being optimised

All data analysis basically starts from (raw) ADC values

SMD has many channels (9000 per barrel half similar number for endcap)

- On-line (fast off-line) monitoring of channels needs large data samples

Somewhat separate topic, but will discuss ‘ezTree’

ezTree II

- ezTree mode only implemented for Endcap
  - Allows for raw data inspection
  - Some existing Endcap tools are based on ezTree (Panitkin plots)
- Barrel data in MuDst format can easily be added (not the default)
  - Barrel on-line QA mostly done with tools reading raw data directly
  - Barrel pedestals are determined on-line