

Experiences in CAF

A. Abrahantes
CEADEN, Cuba

Beginner Experience (2008)

- Problems I had while learning how to use CAF
 - My analysis did not work
 - Tutorial files worked on certain data-sets
 - Not connected to CAF
 - Problems while staging: none.
 - Main suggestion follow documentation page
 - <http://aliceinfo.cern.ch/Offline/Activities/Analysis/CAF/>

From user side, what have changed.

- GRID validation of users.
- Access data staged on the GRID. (I did not tested myself... yet)
- In ROOT an environment variable needs to be set.
- Likely there are much more but, hopefully, invisible to final users.

Single Variable Setting in ROOT

- If not existing create in your \$HOME directory a .rootrc file
- Put the following line.
- XSec.GSI.DelegProxy: 2

I have noticed...

- CAF is more stable (except in period right after new variable setting for ROOT)
- A considerable less amount of post asking to reset sessions or cluster.

Useful?

- Yes, very useful.
- CAF allows to use par files (which is discouraged in GRID analysis)
- CAF is a perfect playground to quickly test own analysis trains in recently produced data (depending on the willing of PWG to stage data-sets for testing).
- I'd like more disk space and CPU, but... not my call

My Train scores

- In three weeks 40 or more useful calculations (ESD + MC)
- My train runs over three different data sets, around 1 millions events (make life easier using batch).
 - ESD analysis 350 events/second, (<1 hour)
 - MC 100 events/sec. (bit more time but only after ESD data is useful, also reported for GRID)
- Accounting analysis tool tuning and other miscellanea, at least 50% more full analysis.
- For me it is a reasonable time.

Speeding up User's Analysis

My Suggestion to “Stagers”

- State clearly relevant information of a dataset. It will allow other users to spot which kind of data is staged.
- Job Production In GRID catalog (i.e., LHC09a3, LHC09a4...)
- User get full info by visiting:
http://pcalimonitor.cern.ch/job_details.jsp

My Suggestion to PWG

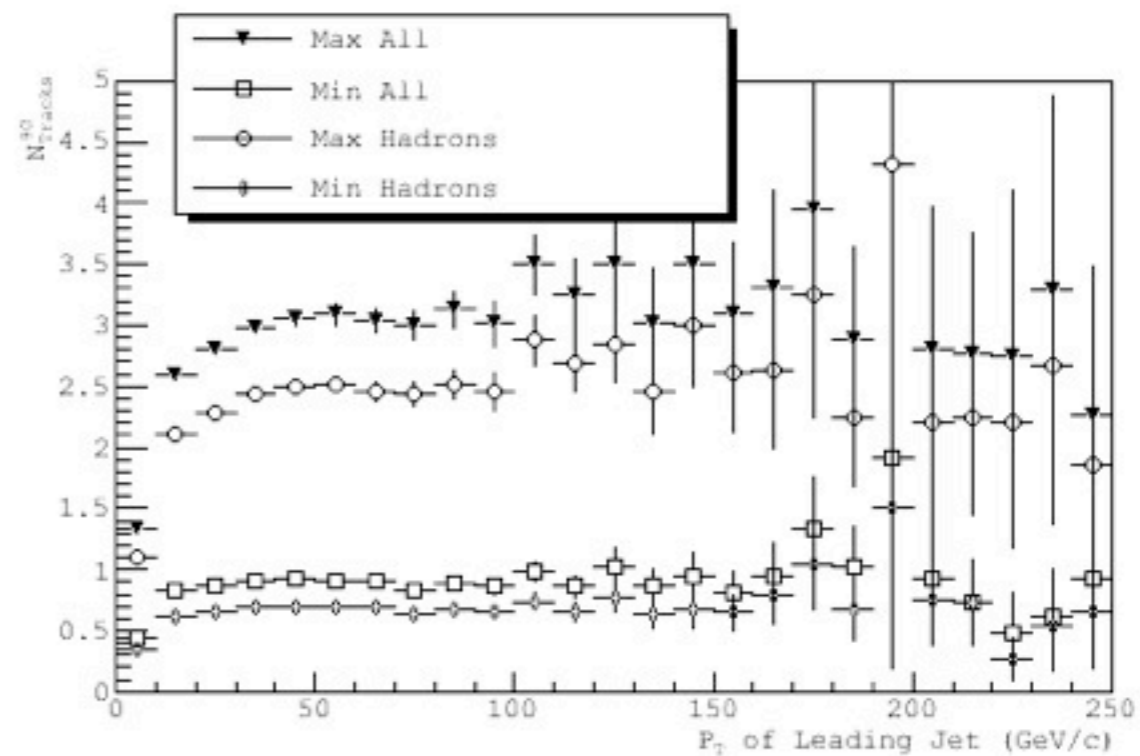
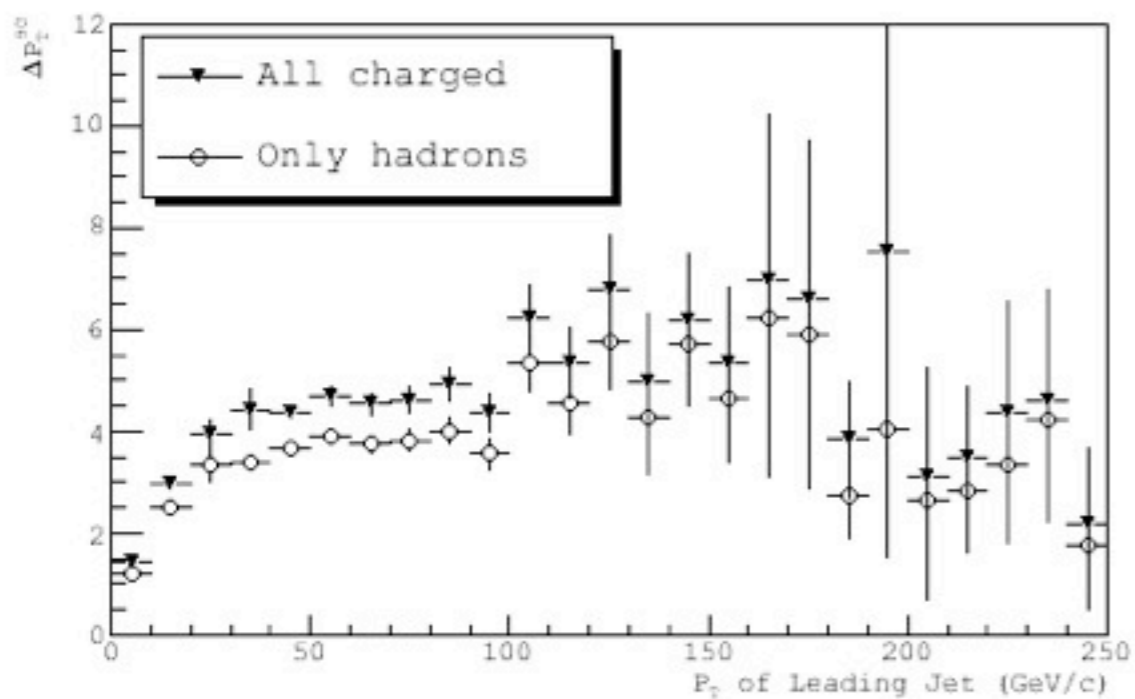
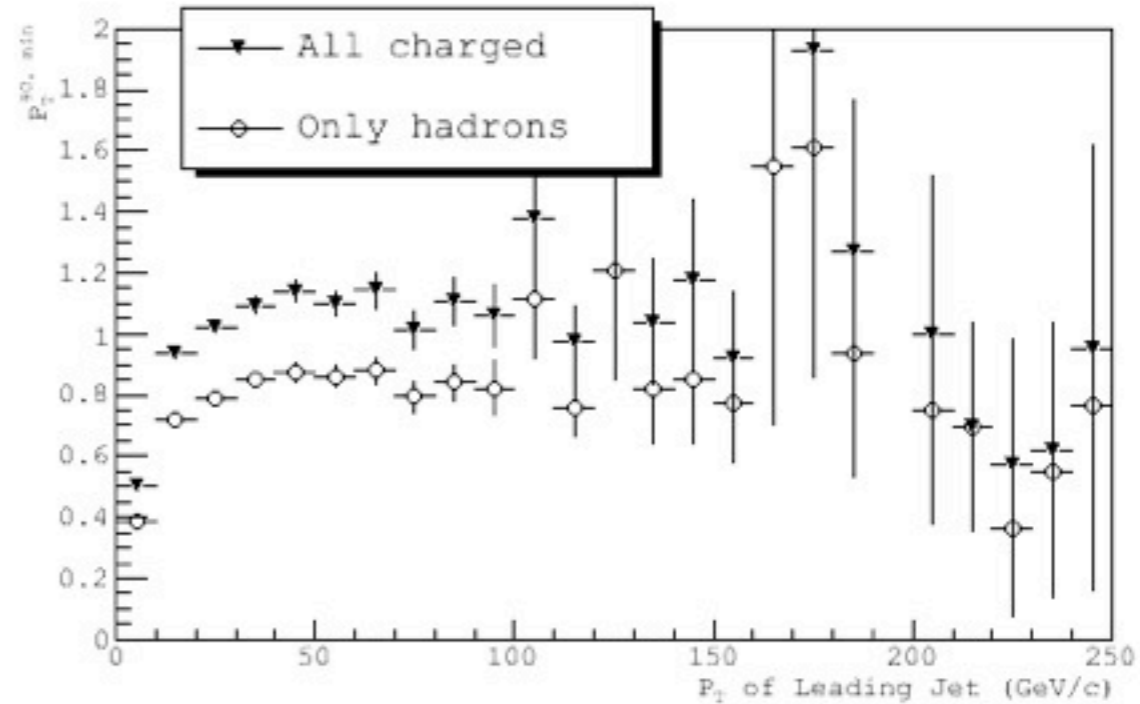
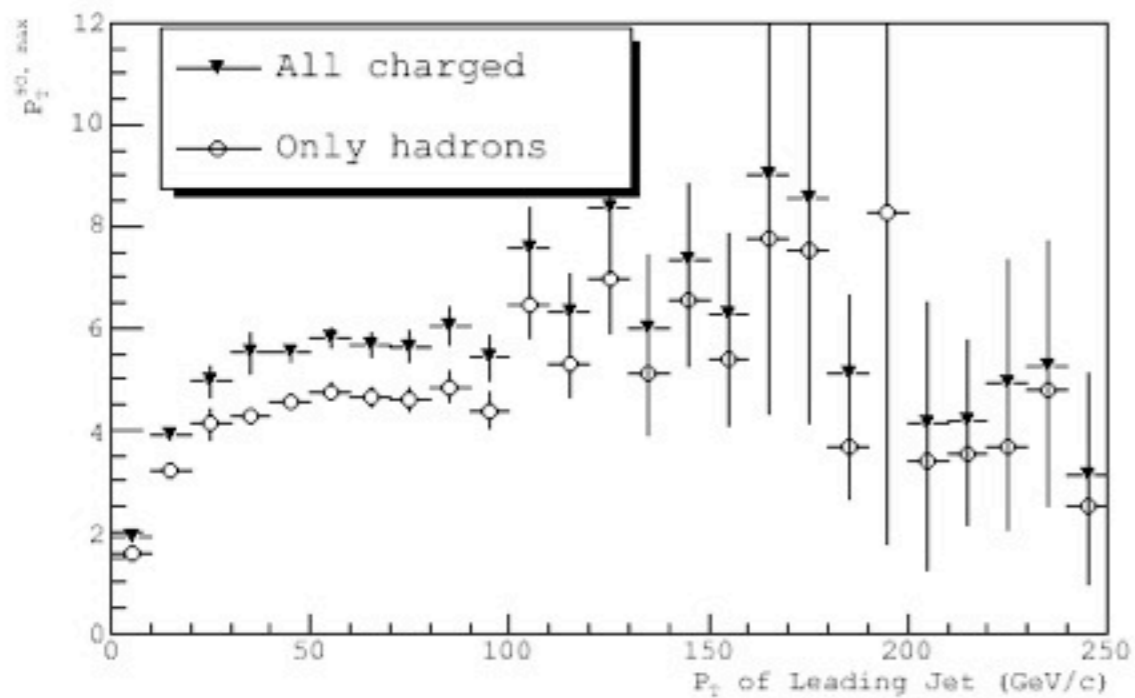
- Check PWG 4 strategy for analysis train (CAF example included in group's main macros)
- Showing your preferred user how to use already coded tasks in a train with a minimal code intervention and teaching him/her a non anarchic way to add their own classes.
- Such approach saved me a lot of time. Furthermore it can be ported to GRID analysis with a few changes (Thanks to alien plug-in).

CAF admins

- Keep nice and good work with cluster tools and new features.
- Keep short response time to users
- Keep well documented and updated CAF page
 - <http://aliceinfo.cern.ch/Offline/Activities/Analysis/CAF/>
 - And new users please use it...

What I'd like...

- Do not put constraints on users.
- CAF analysis are scattered and, at least for me, short computation time (data sets are short).
- Probably “clever” users can be spotted and warned to behave correctly (i.e. analyzing in CAF large collection from GRID). At least warn verbally in web page.
- Still CAF admins needs to put a top level on what can be a “large collection”



Run While waiting my turn