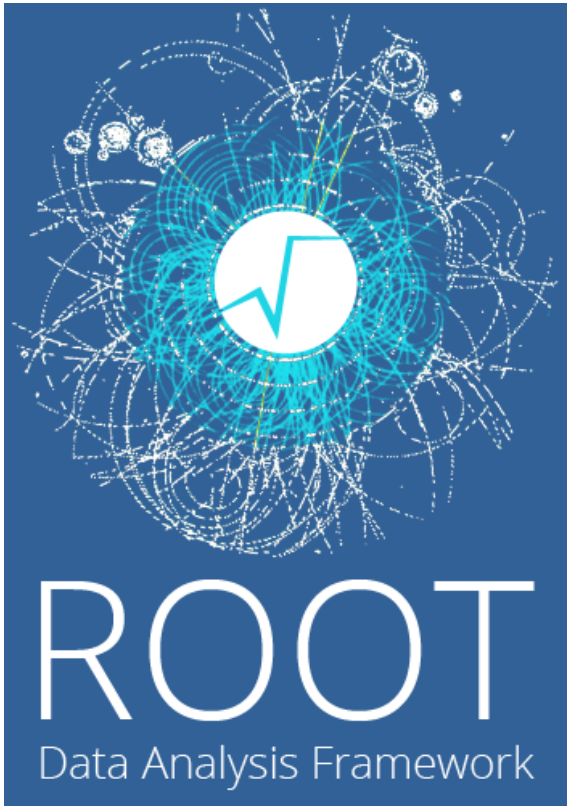
Kolja Kauder



Analyzing Simulated EIC Data in ROOT

Source

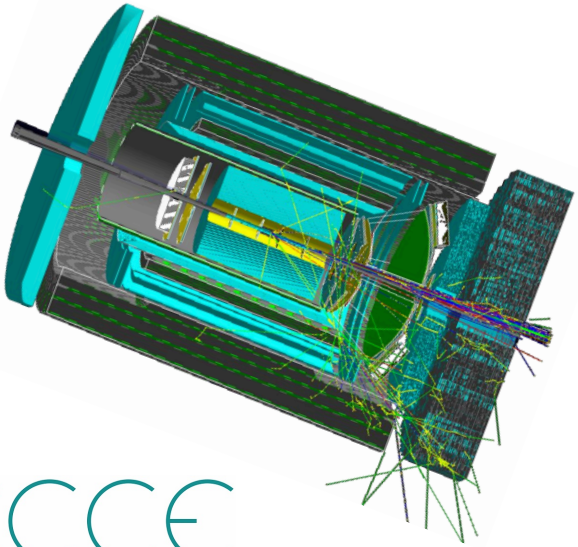
Center for Frontiers in Nuclear Science

CFNS Summer School
Online, August 20, 2021

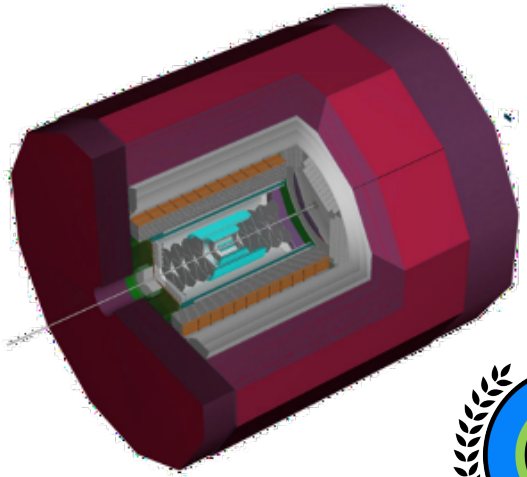
Preamble

Have to focus on a proto-collaboration

- Here → ECCE; Wouter → ATHENA
- There is no correlation. **Either approach works, modulo details, for any proto-collaboration!**



ECCE

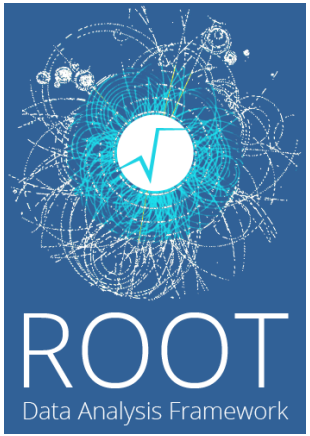


ATHENA

ROOT

ROOT is THE standard library in HEP

- Histogramming, fitting, statistical analysis, …
- **Trees** – the main data structure used to store and efficiently work with Terabytes and Petabytes of collision events
- Based on C++, includes C++ interpreter (cling) for a MATLAB/python-like experience
- Additional interfaces (PyRoot, Ruby, uproot, …)
- The initial learning curve can be steep…



Beginner's Guide
ROOT Tutorials
Lecture from René Brun (2007)

My Credentials

- I am an experienced analyzer in STAR and EIC until ~Yellow Report
- Now primarily EIC software engineer
- I am NOT a member of ATHENA or ECCE
- I have NO experience with ATHENA's framework, and (practically) NO experience with fun4all

→ I started on Wednesday from other people's tutorials

It may take more than two days, but you too can get from zero to a working analysis quickly

ECCE Workshop
ECCE Tutorials
ATHENA Tutorials

Acquire Data

ATHENA S3
ECCE S3

```
$ setenv S3_ACCESS_KEY [redacted]; setenv S3_SECRET_KEY [redacted]
```

- Browse and copy with MinIO mc:

For tests or on a plane

```
$ mc ls eic/eictest/ECCE/MC/ana.14/5f210c7/SIDIS/pythia6
$ mc cp eic/eictest/ECCE/MC/ana.14/5f210c7/SIDIS/pythia6/ep_18x100highq2/
ST_SIDIS_pythia6_ep_18x100highq2_039_0030000_01000.root .
```

- Or open (stream) directly:

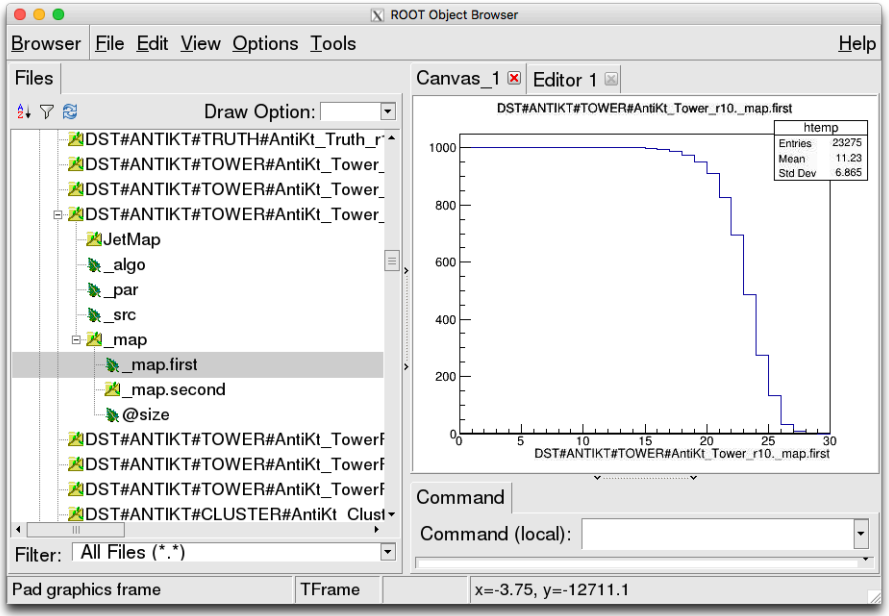
```
$ root
root [0] auto f = new
TS3WebFile("s3://dtn01.sdcc.bnl.gov:9000/eictest/ECCE/MC/ana.14/5f210c7/SIDIS
/pythia6/ep_18x100highq2/DST_SIDIS_pythia6_ep_18x100highq2_039_0030000_01000.
root");
```

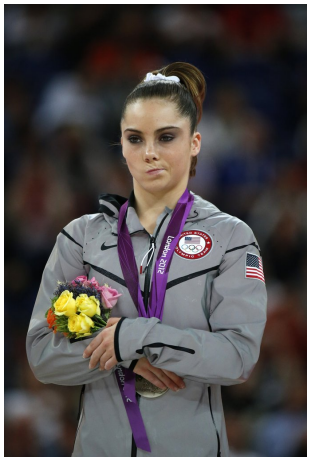
Much better – there's a LOT of data

Note: xrdcp/XRootD are very similar for data at JLab

Inspect the Data

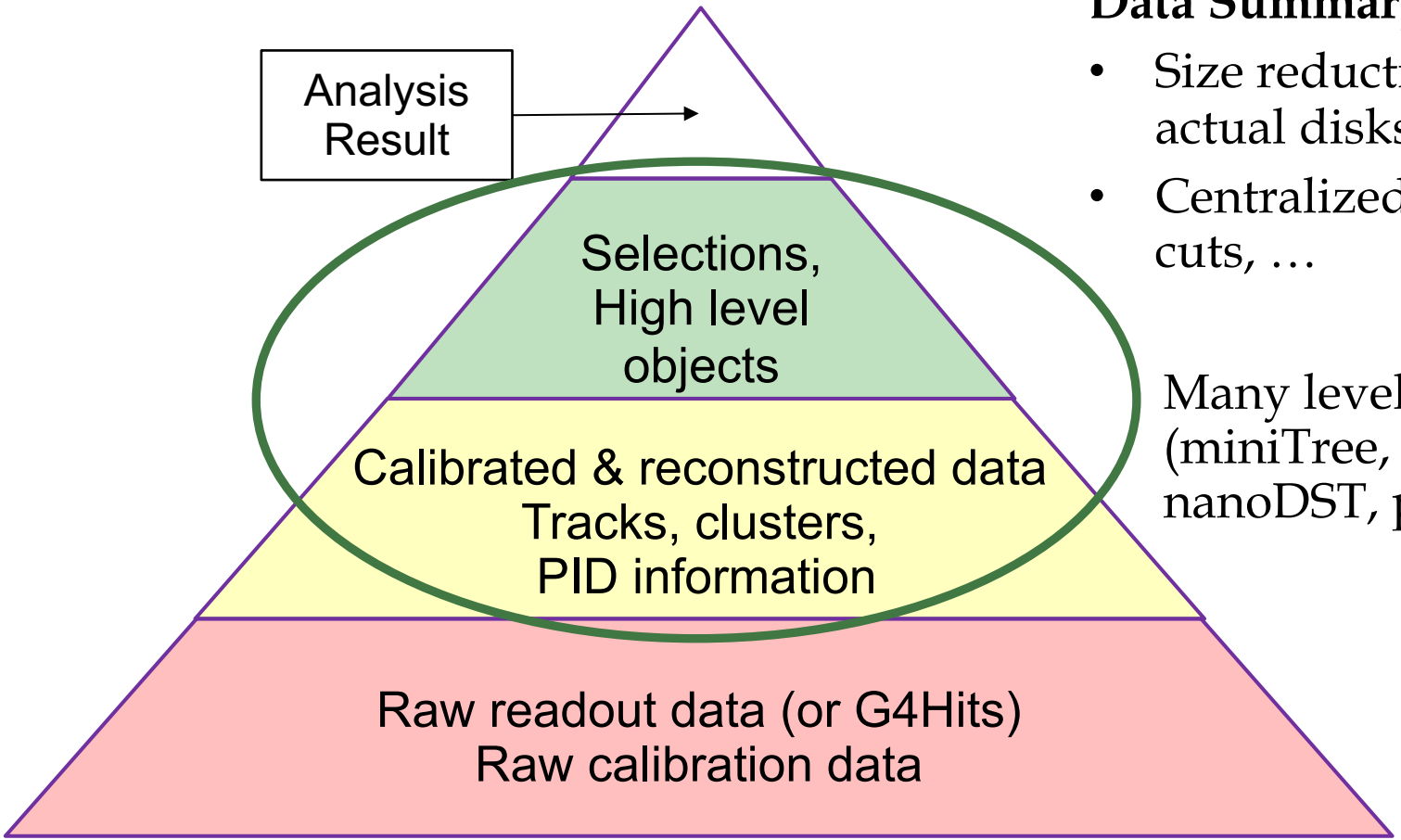
```
root [0] auto f = new
TS3WebFile("s3://dtn01.sdcc.bnl.gov:9000/eictest/ECCE/MC/ana.14/5f210c7/SIDIS
/pythia6/ep_18x100highq2/DST_SIDIS_pythia6_ep_18x100highq2_039_0030000_01000.
root");
root [1] new TBrowser
```





Working with raw trees is possible, but not recommended

The Data Pyramid



Data Summary Tape (DST)

- Size reduction (can fit on actual disks ☺)
- Centralized corrections, cuts, …

Many levels and names (miniTree, muDST, nanoDST, picoDST, …)

Load a Framework

- If you're not already logged into SDCC or JLab, use a container:

```
$ singularity shell -B /cvmfs:/cvmfs
/cvmfs/eic.opensciencegrid.org/singularity/rhic_sl7_ext.simg
```

- Source the setup script:

```
$ source
/cvmfs/eic.opensciencegrid.org/ecce/default/opt/fun4all/core/bin/ecce_setup.sh -n
# or
$ source
/cvmfs/eic.opensciencegrid.org/ecce/default/opt/fun4all/core/bin/ecce_setup.csh -n

## Note: You may have to source additional scripts
```

Ready to analyze!

ATHENA Containers

ECCE Containers

Start with a Tutorial

- Set up, clone, compile, run

```
$ mkdir install; setenv MYINSTALL $PWD/install;
$ source $OPT_SPHENIX/bin/setup_local.csh $MYINSTALL

$ git clone https://github.com/ECCE-EIC/tutorials

$ cd tutorials/AnaTutorialECCE
$ mkdir build; cd build
$ ../autogen.sh --prefix=$MYINSTALL
$ make install
```

```
$ cd ../../macro
$ root Fun4All_G4_EICDetector_AnaTutorial.C
```

Let's Look at the Result

- That created a lot of things!

```
[eickolja@eic0101 macro]$ ls *root
G4EICDetector_anaTutorial.root   G4EICDetector_g4hcalin_eval.root    g4fwdjets_full_05_eval.root
G4EICDetector_eventtree.root     G4EICDetector_g4hcalout_eval.root   g4fwdjets_tower_05_eval.root
G4EICDetector_g4cemc_eval.root   G4EICDetector_g4tracking_eval.root  g4fwdjets_tower_07_eval.root
G4EICDetector_g4eemc_eval.root   g4fwdjets_cluster_05_eval.root      g4fwdjets_tower_10_eval.root
G4EICDetector_g4femc_eval.root   g4fwdjets_cluster_07_eval.root      g4fwdjets_track_05_eval.root
G4EICDetector_g4fhcal_eval.root  g4fwdjets_cluster_10_eval.root
```

- What exactly? And how?

```
$ emacs Fun4All_G4_EICDetector_AnaTutorial.C &
$ root G4EICDetector_anaTutorial.root
root [1] new TBrowser
```

This macro creates a lot of the high-level objects you want to use → But somebody's already done that!

Never Start from Scratch

- Copy and clean up

```
$ cd ../..
$ cp -r AnaTutorialECCE/ MyDIS
$ rm -r MyDIS/src/build ; mkdir MyDIS/src/build
$ rm -r MyDIS/macro/*root
```

- Rename our starting point

```
$ cd MyDIS/src
$ mv AnaTutorialECCE.cc MyDIS.cc # [...]
$ emacs MyDIS.* Makefile.am configure.ac &
$
$ cd build
$ ../autogen.sh --prefix=$MYINSTALL
$ make install
```

Some of this looks like magic incantations – but all you need is search/replace

Now we can work

- Keep and adapt the wrapper macro we want

```
$ cd ../../macro
$ mkdir bak ; mv Fun4All_G4_EICDetector_AnaTutorial.C bak;
$ mv Fun4All_ReadDST.C MyDIS_Wrapper.C
$ emacs MyDIS_Wrapper.C & # Change names and input
```

- And run!

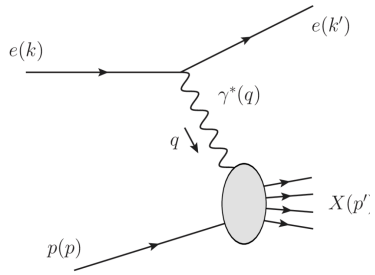
```
$ root MyDIS_Wrapper.C
$ root MyDIS.root
root [1] new TBrowser
```

We have successfully copied and renamed everything, and we're reading pre-generated DSTs from the cloud

Deep Inelastic Scattering

- Locate the scattered e-, calculate x, Q^2, …

```
$ cd ../src
$ emacs MyDIS.* &
```



$$
\begin{aligned}
s &= (p+k)^2\,, \\
q &= k - k'\,, \\
Q^2 &= -q^2\,, \\
W^2 &= p'^2 = (p+q)^2\,, \\
\nu &= \frac{p\cdot q}{m_p} = \frac{W^2+Q^2-m_p^2}{2m_p}\,, \\
x &= \frac{Q^2}{2m_p\nu} = \frac{Q^2}{W^2+Q^2-m_p^2}\,,
\end{aligned}
$$

What should we do?

- Get all tracks (= charged particles), PID them as electrons
- Get all EM towers, PID them as electrons
- If both infos match, combine for the best resolution
- Bonus: What about multiple electrons in one event?

What we will do: Tracks with matched EM tower == electron

What we will do

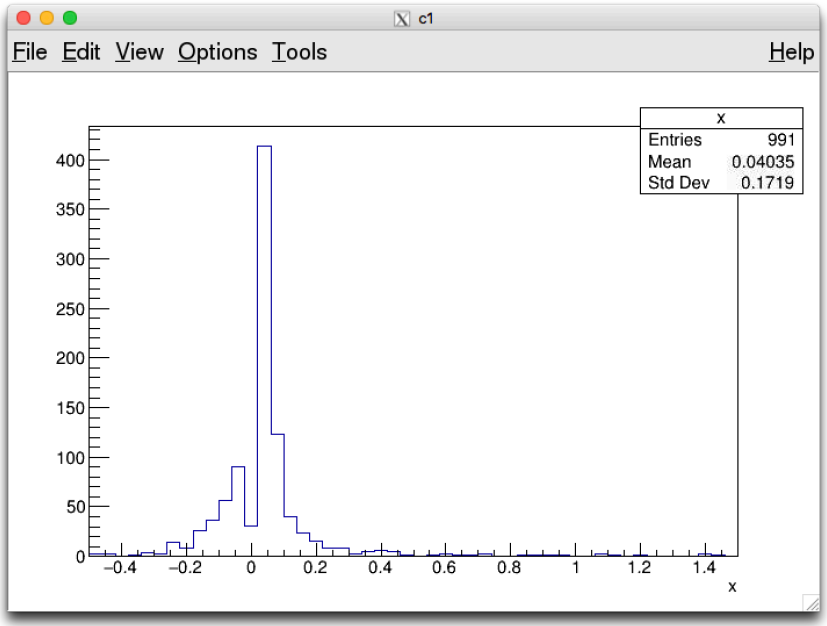
- Use truth PID info on tracks
- Ignore the EMC → need matching
- Use first found electron (cheat that works for MC)

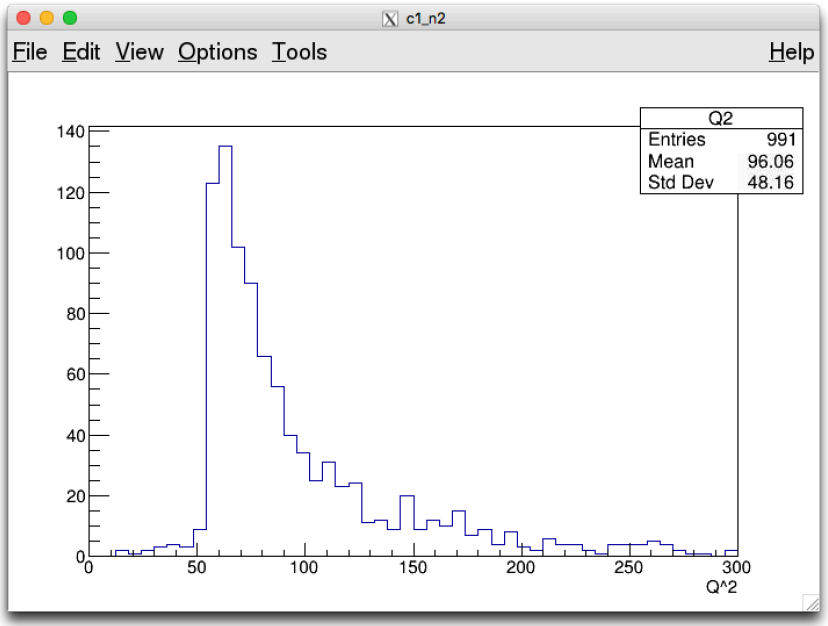
Relying heavily on tracks → Use `getTracks()` as a starting point

Google is the programmer's best friend: google SvtxTrack

First Result

- From zero to kinematics histograms, in an hour (I hope!)
- When I did it, something's fishy with negative x – let's do better together!





Time left?

- EMC?
- Comparison to truth?
- Jets: Either run your own FastJet or use pre-made jet objects
- Decays, like J/psi or Lambda → Probably need PID

Resources

- Mailing list: eicug-software@eicug.org, eic-bnl-soft-l@lists.bnl.gov
- Slack channel: eicug.slack.com/#software-support
- Contact: kkauder@bnl.gov

- Tutorials:
  https://indico.bnl.gov/event/11112/
  https://ecce-eic.github.io/tutorials_landing_page.html
  https://eic.phy.anl.gov/tutorials/eic_tutorial/getting-started/quickstart/

- Class documentation: https://eic.github.io/doxygen/

- HSF Training
  https://hepsoftwarefoundation.org/training/curriculum.html