A journey to the center of ROOT's log files

Alexandra-Maria Dobrescu

Who am I ?

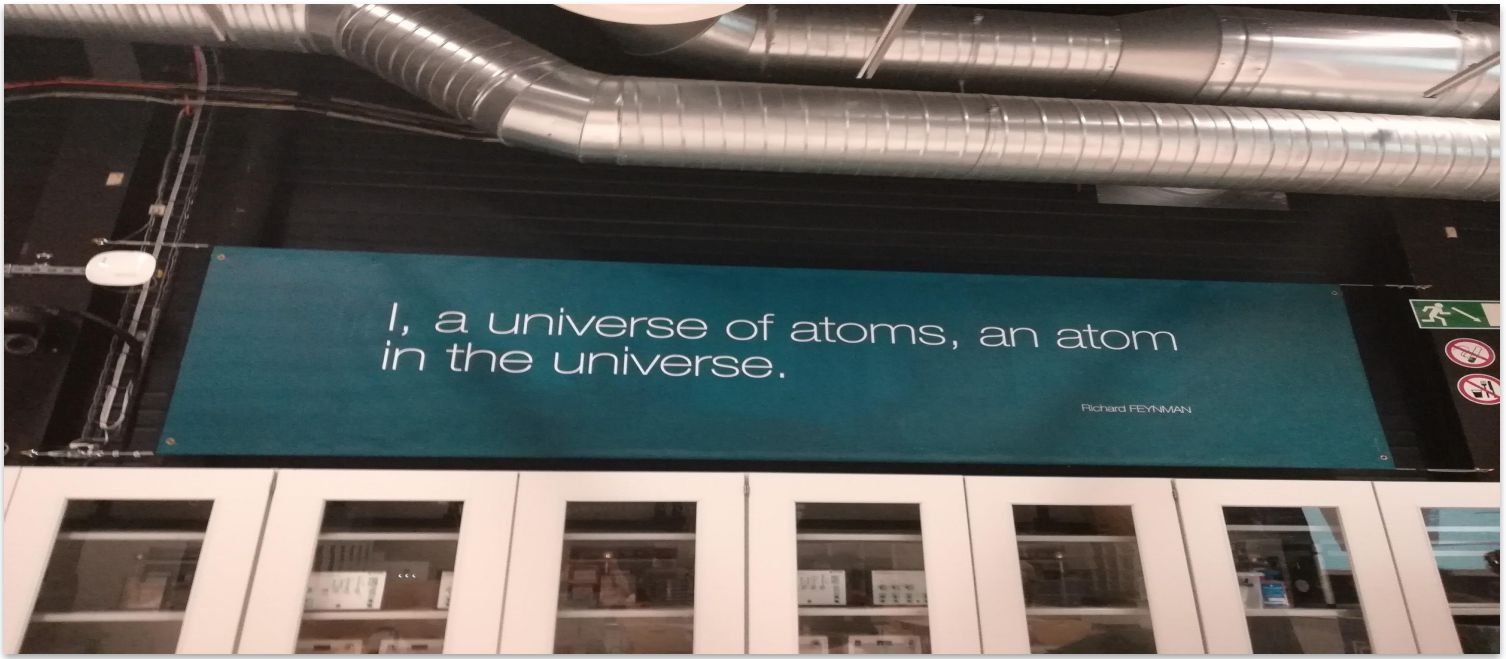
Maybe a **curious** Particle

https://scool.web.cern.ch/content/particle-identities-quiz

According to your answers, the particle which fits your personality

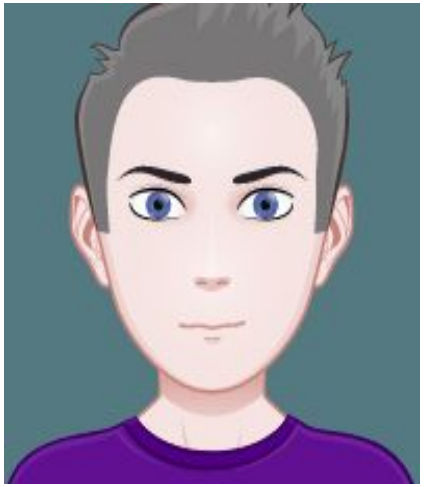
The Higgs boson is the most recently discovered elementary particle. It was predicted by physicists in 1964. However, it took almost 50 years and the most powerful particle accelerator in the world to discover the Higgs Boson in 2012 at CERN.

Or "a universe of atoms, an atom in the universe"...



...From what I've seen at the cloud chamber workshop

Being more serious:

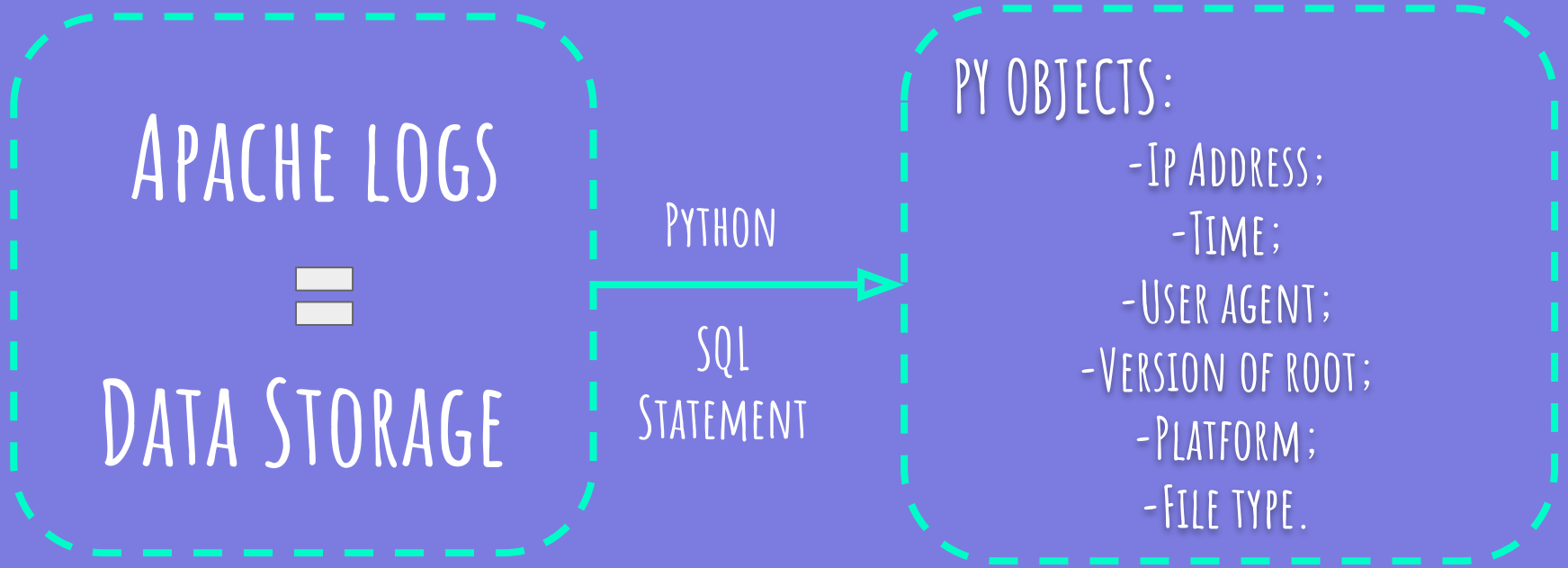


Supervisor : Axel Naumann, Olivier couet

Data collection

First trip:

Extracting the data from the Apache logs

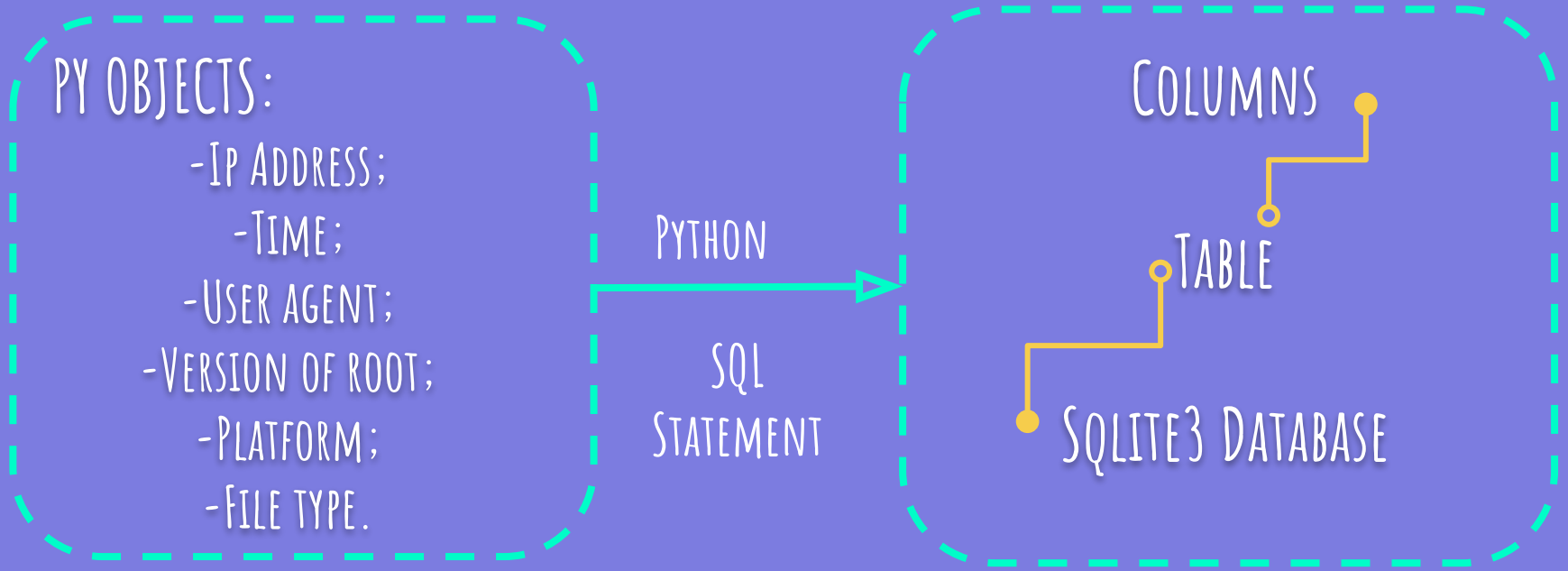


It filters robots and resolves GeoIP.

How can it get more interesting?

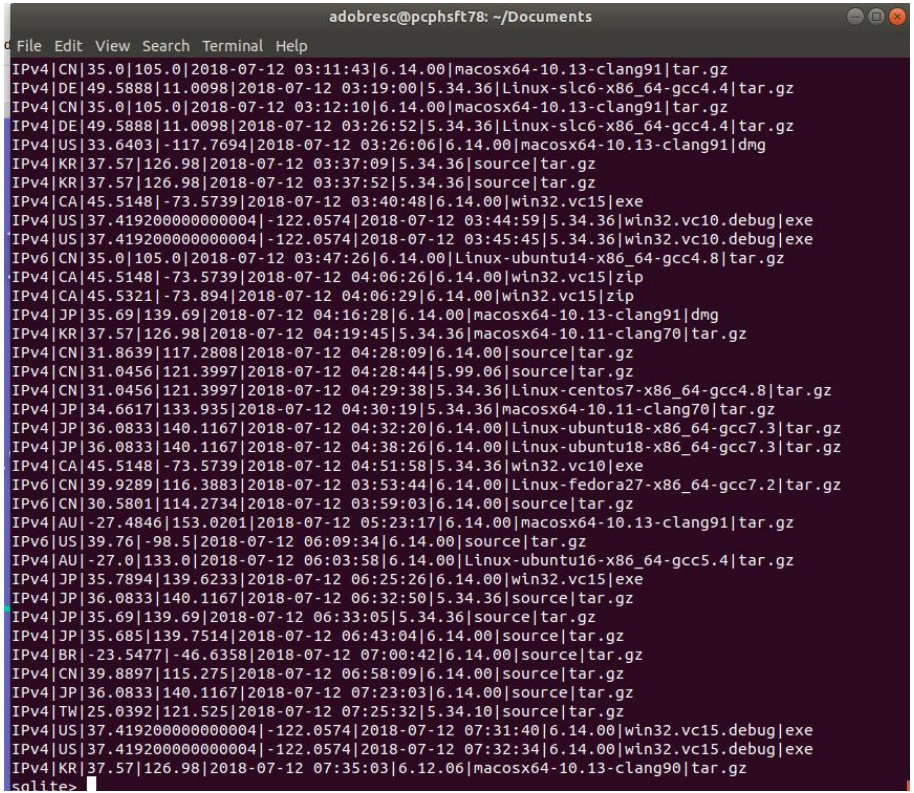
Second step:

From PY Objects to columns.



It creates a sqlite3 file.

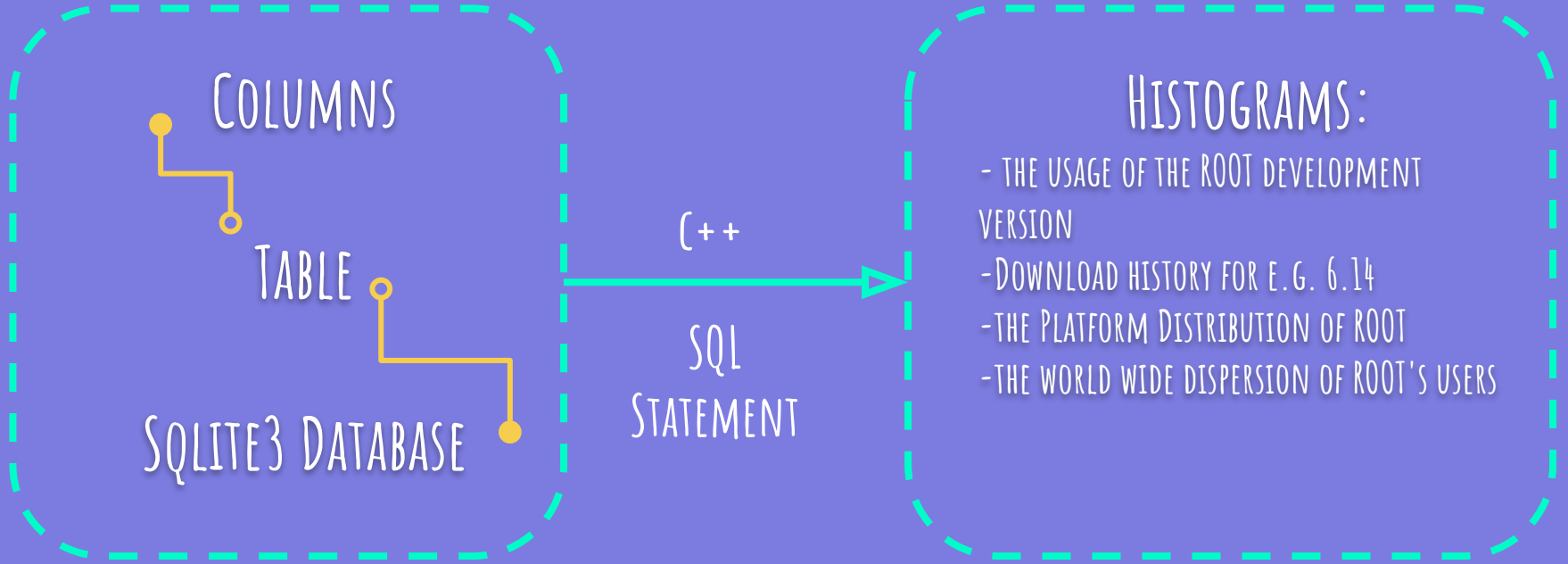
Who wants to see a terminal?



No one. :(

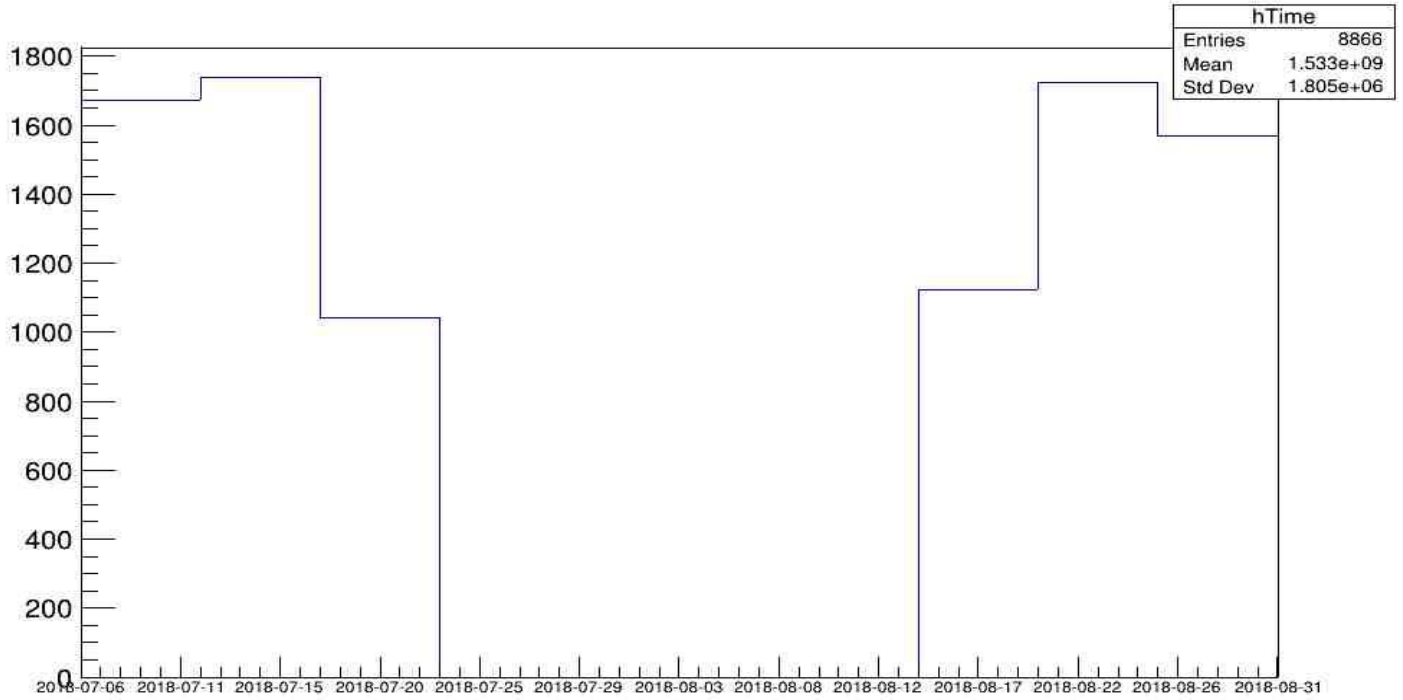
Third trip:

From SQliteDB to Histograms.

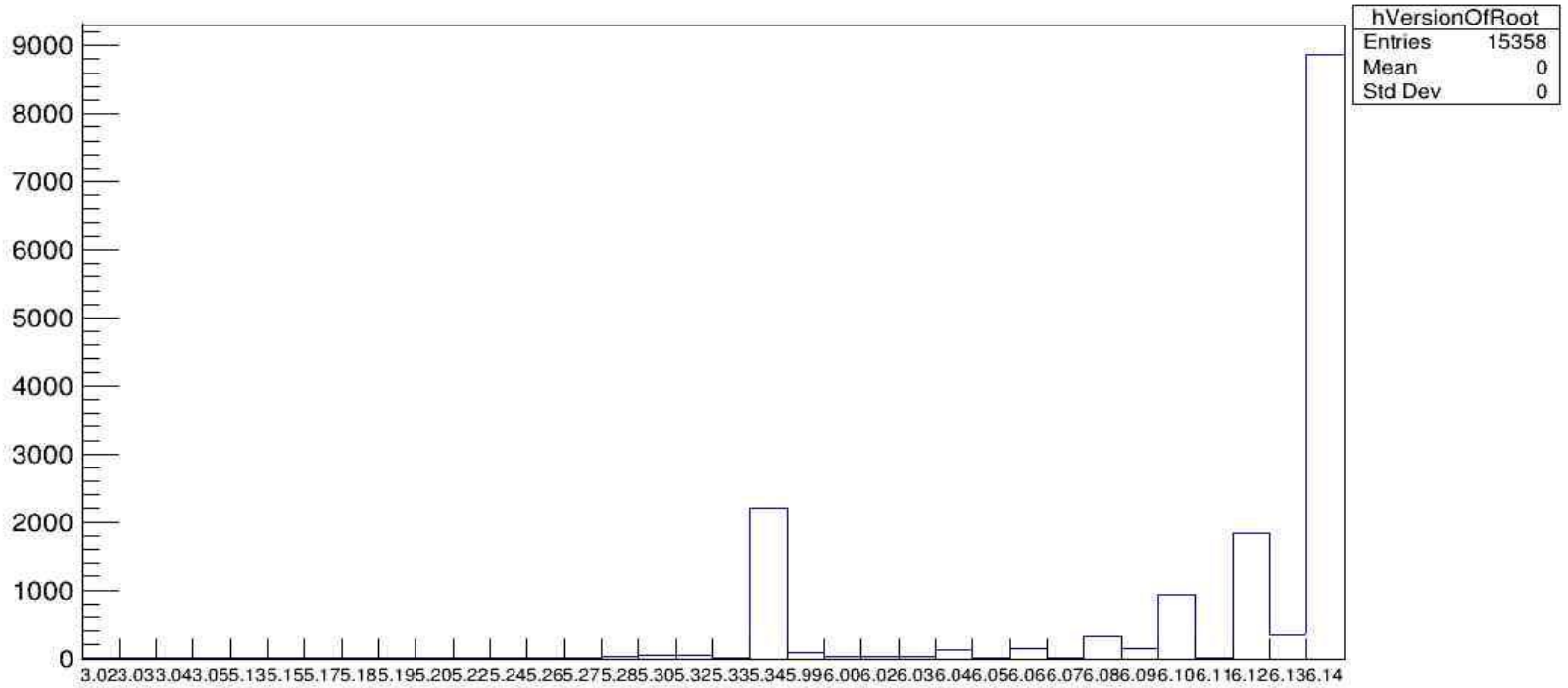


Create sqlite Tutorials for log statistics

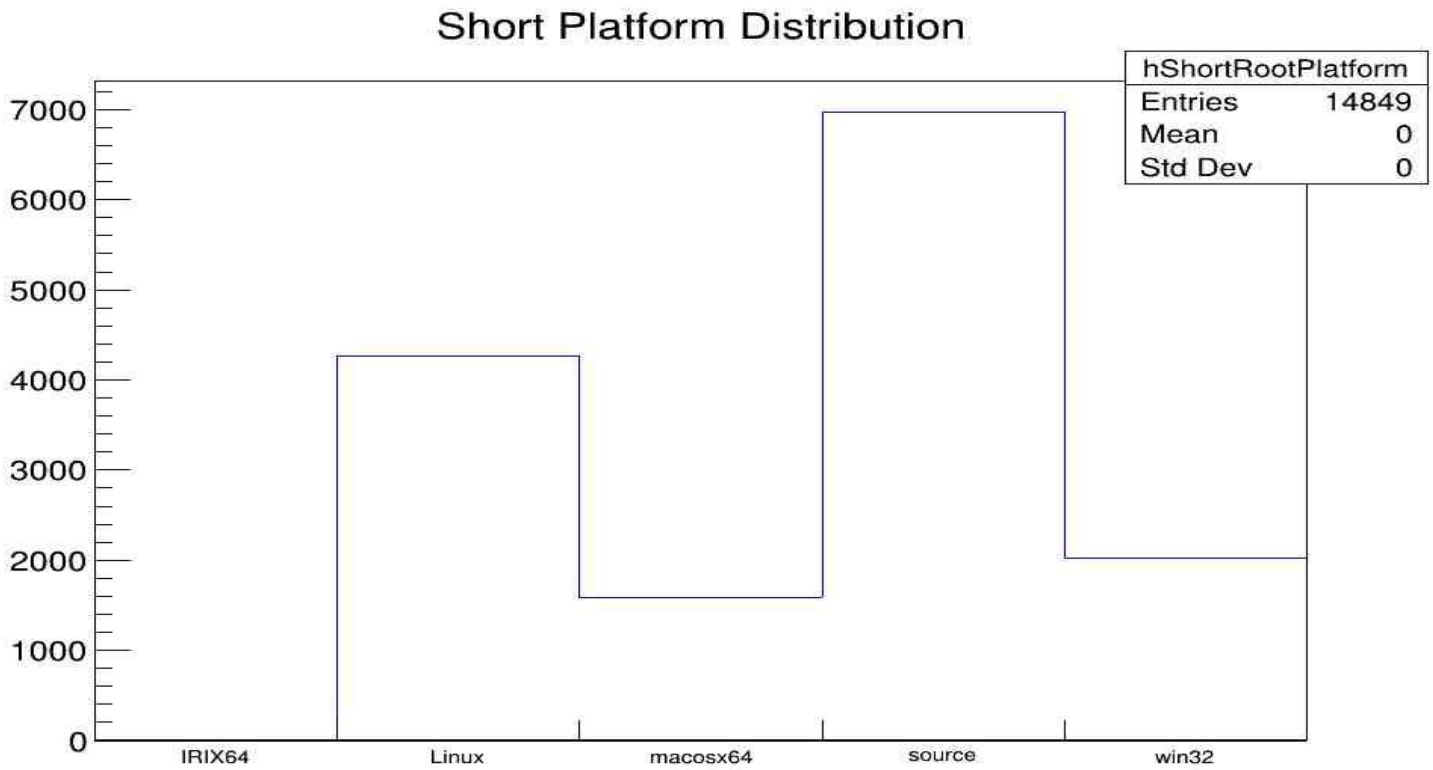
6.14 Download history



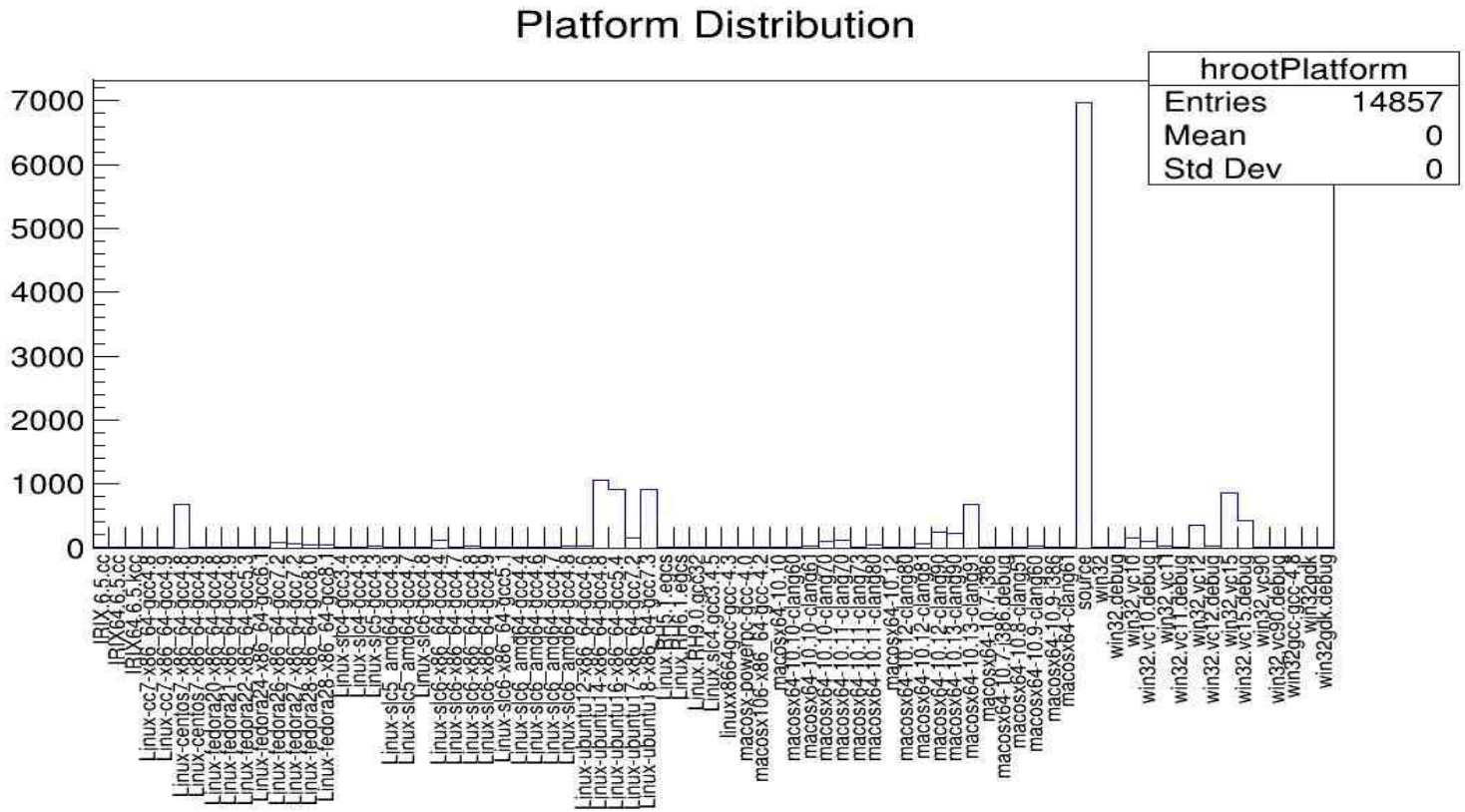
THE USAGE OF PAST ROOT VERSIONS



The platform distribution of ROOT



The platform distribution of ROOT



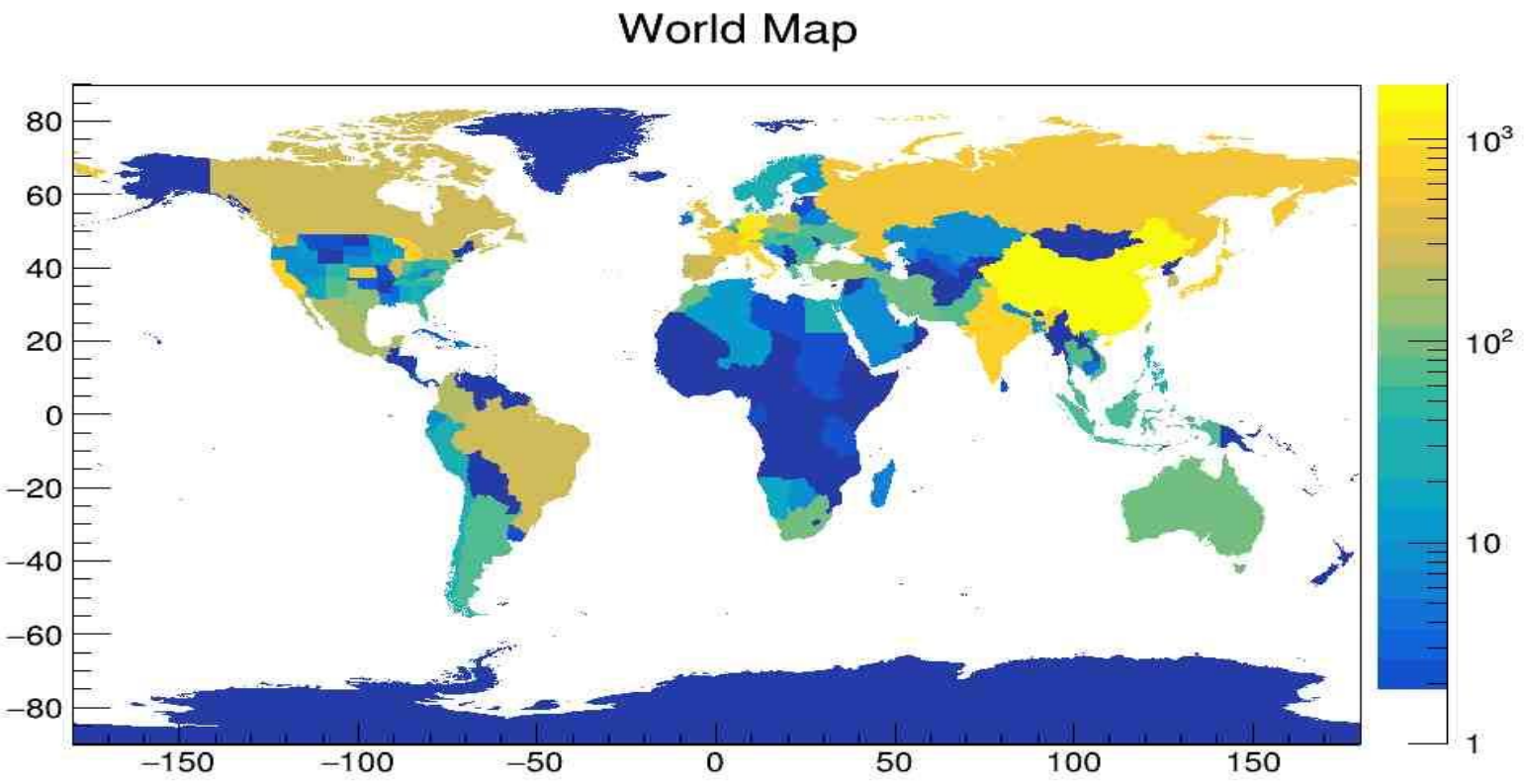
Is the journey ending?

From the center of ROOT's log files

TO

the Wide World.

The world wide dispersion of ROOT's users



What are the histograms hiding?

- 1 set of macros using **TSQLServer**
- 1 set of macros using **RDataFrame**

TSQLServer

Abstract base class defining an interface to SQL Server

What I’ve learned and used:

1. Create macros with ROOT
2. Connect to SQLite3 database
3. Launch Queries
4. Generate a Histogram
5. Fill the Histogram with the specific data stored in the columns.

The inconvenient: the database should be downloaded locally.

RDataFrame

The Swiss army knife for ROOT Data Analysis

What I've learned and used:

1. Create macros with ROOT
2. Connect to SQLite3 database
3. *Launch a general Query to select the entire table*
4. Generate a Histogram
5. Fill the Histogram with the specific data stored in the columns
6. *Simplify the procedure with a lambda expression*

Advantage: More interesting task to come

Other travel expenses:

- Learn how to use UNIX
- Improve Regular Expressions
- gdb

Last but not least - Let the others know about the journey:

- Submit a PR on GitHub

And make new friends @jakob, @Enrico

99.9 %

One last visit :

Improving the RDataFrame data source class for reading **remote** SQlite files over http(s).

Aha!

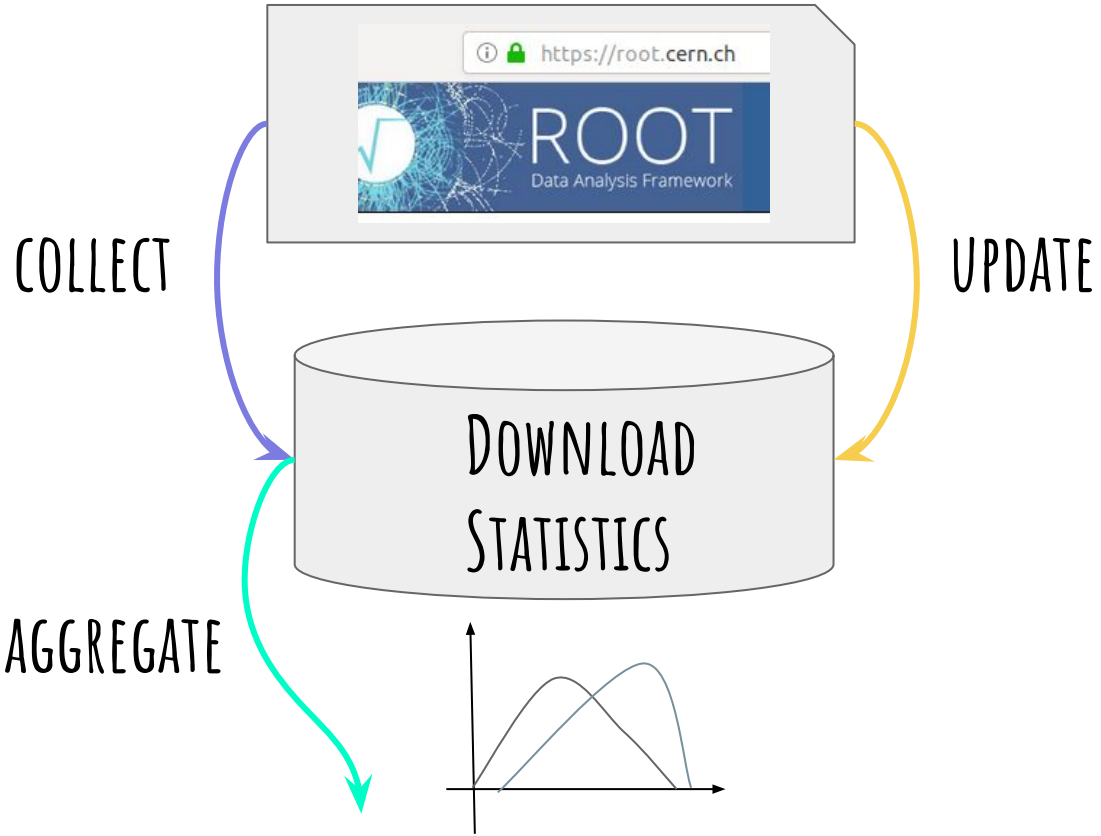
Purpose: To update on root.cern.ch the sqlite3 database. In this way, the user can benefit of a real time experience with the log files. Like a journey in real time.
Meaning: the local file becomes a remote file.
"The ingredients" :

- A customised sqlite VFS layer;

"Recipe secret":

- access to the Davix library in Jakob's brand new RSqliteDS.cxx file.

Conclusion



- Displayed by Tutorials
- Using **Remote** SQlite DataSource access

Thank you !