

第三章：实验统计分析工具ROOT

授课人：董燎原
中国科学院高能物理研究所

Introduction to ROOT

ROOT的主页，由CERN维护。

ROOT Data Analysis Framework

About Install Get Started Forum & Help Manual Blog Posts Contribute For Developers

ROOT: analyzing petabytes of data, scientifically.

An open-source data analysis framework used by high energy physics and others.

[Learn more](#) [Install v6.24/06](#)

基于面向对象设计的数据分析软件框架！



Start



Reference



Forum

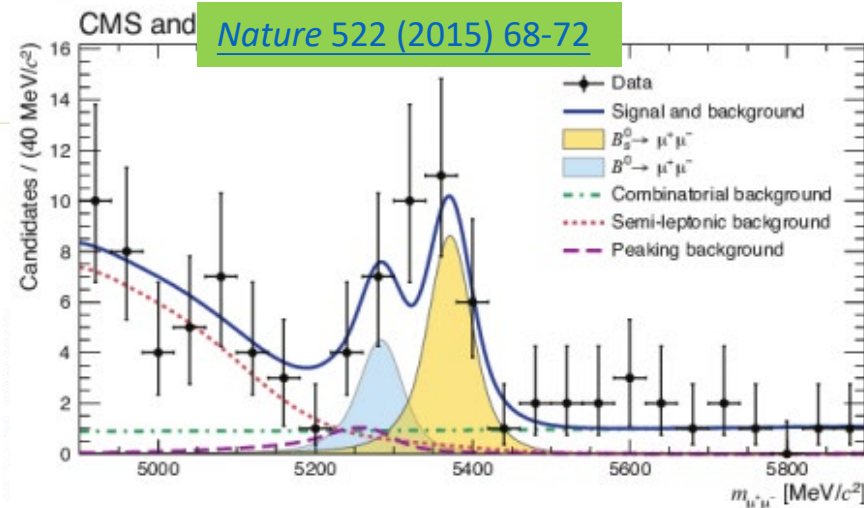


√-1

ROOT enables *statistically sound* scientific analyses and visualization of large amounts of data: today, more than 1 exabyte (1,000,000,000 gigabyte) are stored in ROOT files. The Higgs was found with ROOT!



As *high-performance* software, ROOT is written mainly in C++. You can use it on Linux, macOS, or Windows; it works out of the box. ROOT is open source: use it freely, modify it, contribute to it!

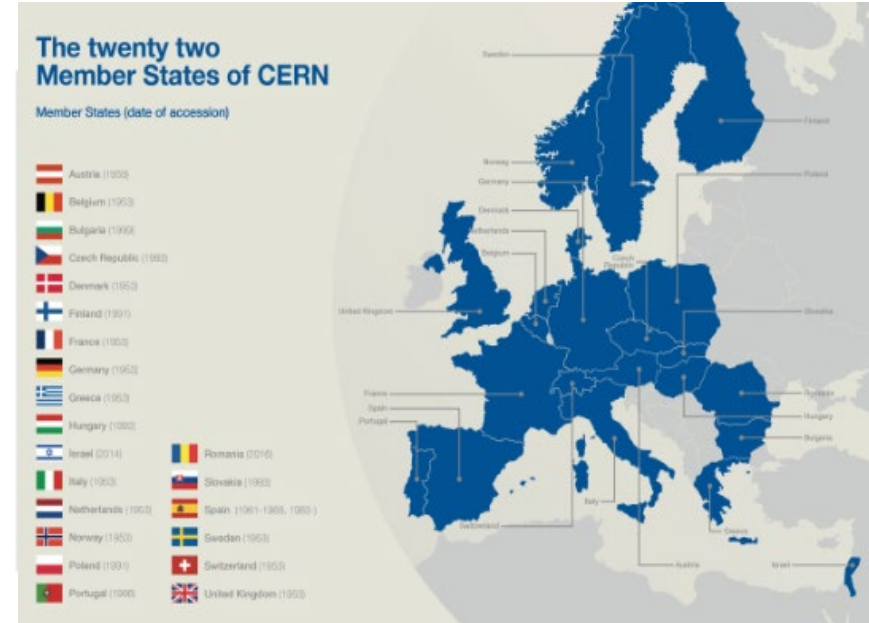


What is CERN ?



- An international research organization „*Conseil Européen pour la Recherche Nucléaire*”
欧洲核子研究中心

- Founded in 1954 by 12 members
- These days there are 22 member states



- One of the world's largest research labs at the Franco-Swiss border near Geneva
瑞士日内瓦
 - more than 10,000 (visiting) scientists work here
- Finding out what the Universe is made of and how it works.

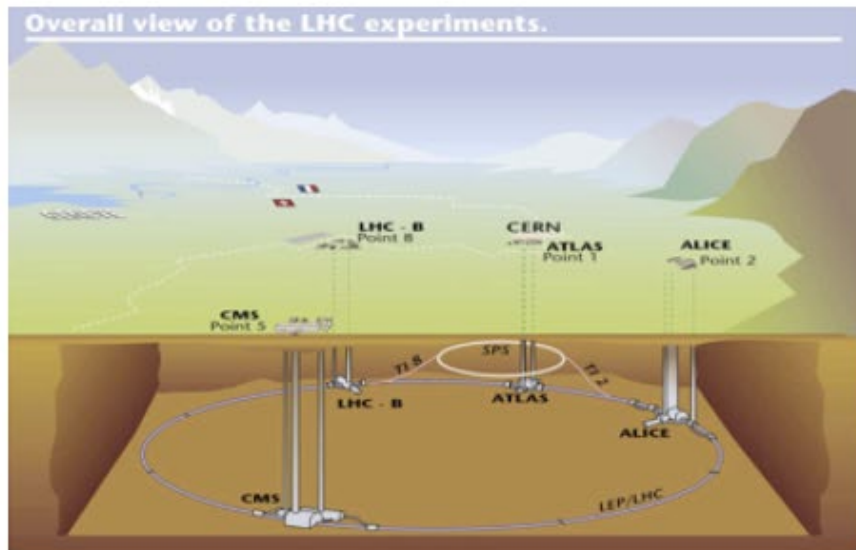
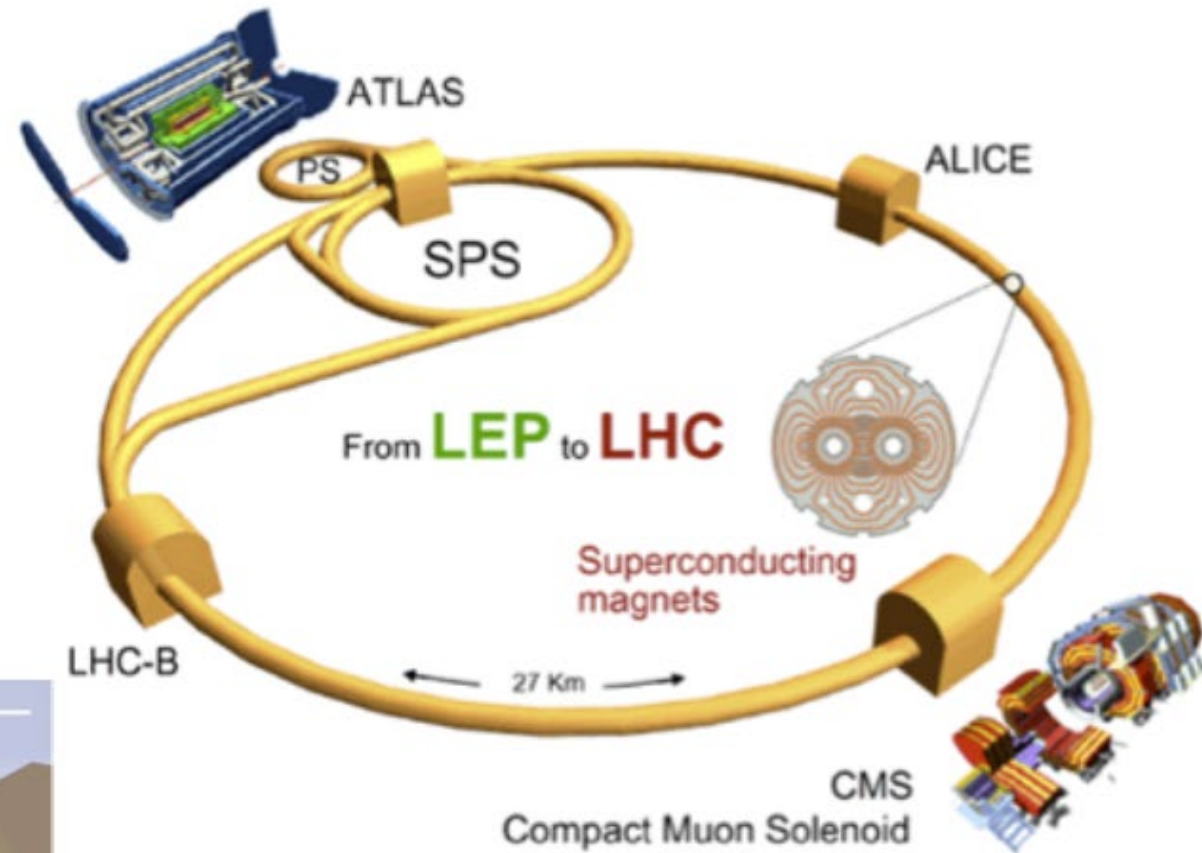


The LHC at CERN

LHC — the Large Hadron Collider — is a 27-kilometre collider ring.

Four large experiments use the LHC as a giant microscope

Started in 2009 and will run until approx. 2030



100 m below surface

ATLAS和CMS是交叉确认任何新发现的关键！
如 2013年3月两个实验宣布发现Higgs玻色子。

Computer Science at CERN

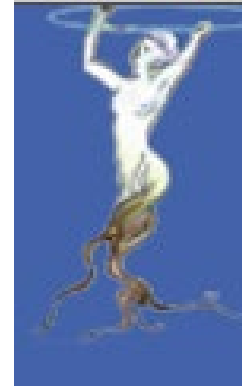
- To fulfill its goals, CERN has always been pushing the limits of computing technology.
- ✓ While searching for a solution for remote information exchange between physicists Tim Berners-Lee invented the web.

世界第一个网站于1990年12月20日上线，
当属的网页属于CERN。直到1991年8月才
正式公开，成为互联网的基础！



- ✓ The immense data collected at CERN required the LHC Computing Grid for distributed processing and storage, anticipating many cloud concepts.
- ✓ In general, strong do-it-yourself attitude.

What's ROOT ?



Brief overview of ROOT and its functionality
before starting to use it

ROOT in a nutshell

- Framework for large scale data handling
- Provides, among others,
 - an efficient data storage, access and query system (PetaBytes)
 - advanced statistical analysis: histogramming, fitting, minimization and multi-variate analysis algorithms
 - scientific visualization: 2D and 3D graphics, Postscript, PDF, LaTeX
 - geometrical modeler
 - PROOF parallel query engine
- An Open Source Project

ROOT: An Open Source Project

- Started in 1995 by *R. Brun* and *F. Rademakers*.
- Available (including the source) under GNU LGPL.
- 7 full time developers at CERN, plus 2 at Fermilab (Chicago).
- Many contributors from high energy physics experiments which uses ROOT as base for their software frameworks.
- Large number of part-time developers.
- Several thousands of users giving feedback and a very long list of small contributions.

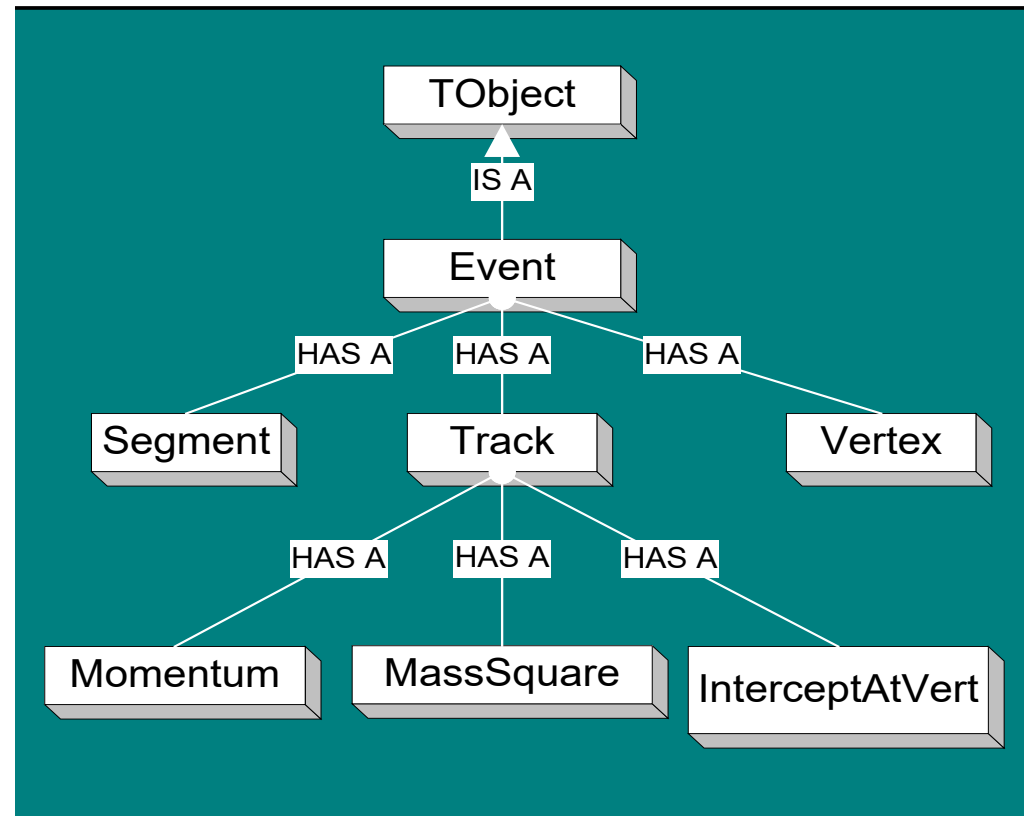
Why ROOT ?

- The analysis of data coming from LHC experiments and not only requires a powerful and general toolkit
 - Visualisation
 - Statistical studies
 - Data reduction
 - Multivariate techniques
- A scalable and reliable persistency method is needed to write the data on disks and tapes.

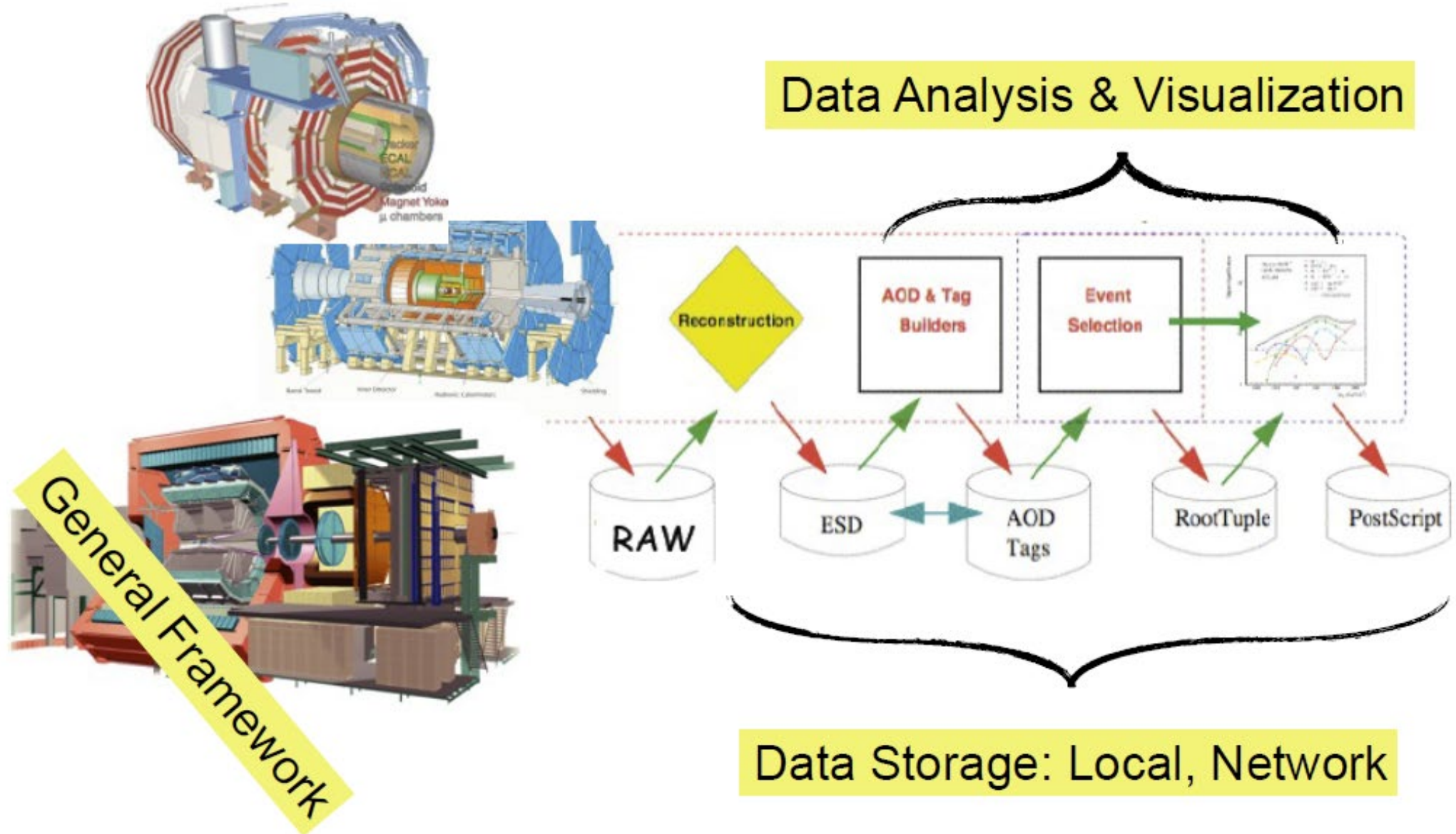
Object Oriented Concepts

- Class: the description of a “thing” in the system
- Object: instance of a class
- Methods: functions for a class

- Members: a “has a” relationship to the class.
- Inheritance: an “is a” relationship to the class.



ROOT Application Domains



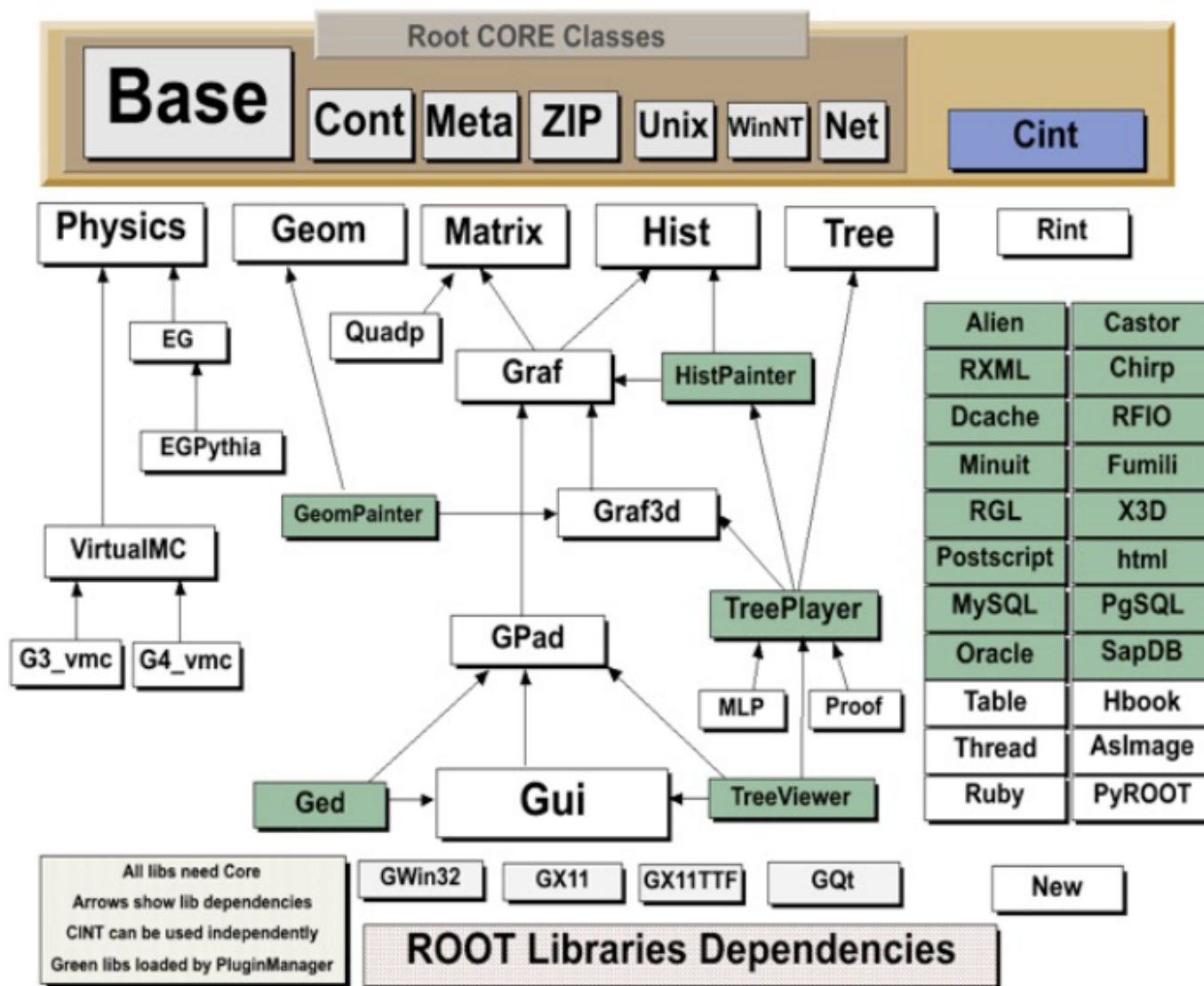
ROOT Services/Utilities

- Histogramming, Fitting, Statistics
- Graphics (2D, 3D)
- I/O to file or socket: specialized for histograms, Ntuples (Trees)
- Collection Classes and Run Time Type Identification
- User Interface
 - GUI: Browsers, Panels, Tree Viewer
 - Command Line interface: C++ interpreter CINT
 - Script Processor (C++ compiled \Leftrightarrow C++ interpreted)

ROOT Libraries

- Overview of ROOT libraries and their dependencies

- 1,700,000 lines of code.
- More than 100 shared libraries
- Fully cross-platform: Unix/Linux, MacOS and Windows.
- More than 500000 downloads (since 1997)



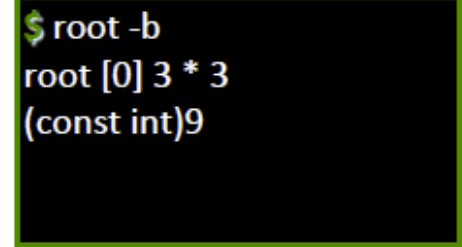
More ROOT Packages

The packages provided by ROOT include those for

- ⑩ [*histogramming*](#) and [*graphing*](#) to visualize and analyze [*distributions*](#) and [*functions*](#),
- ⑩ [*curve fitting*](#) (regression analysis) and minimization of [*functionals*](#),
- ⑩ [*statistics*](#) tools used for [*data analysis*](#),
- ⑩ [*matrix*](#) algebra,
- ⑩ [*four-vector*](#) computations, as used in [*high energy physics*](#),
- ⑩ standard [*mathematical functions*](#),
- ⑩ [*multivariate data analysis*](#), e.g. using [*neural networks*](#),
- ⑩ *image manipulation*, used e.g. to analyze [*astronomical*](#) pictures,
- ⑩ access to *distributed data* (in the context of the [*Grid*](#)),
- ⑩ [*distributed computing*](#), to parallelize [*data analyses*](#),
- ⑩ [*persistence*](#) and [*serialization*](#) of objects, which can cope with changes in class definitions of persistent data,
- ⑩ access to [*databases*](#),
- ⑩ [*3D visualizations*](#) (geometry)
- ⑩ creating files in various *graphics formats*, like [*PostScript*](#), [*JPEG*](#), [*SVG*](#),
- ⑩ interfacing [*Python*](#) and [*Ruby*](#) code in both directions,
- ⑩ interfacing [*Monte Carlo event generators*](#).

The Interpreter

- ROOT is shipped with an interpreter, CINT*
 - C++ not trivial to interpret and not foreseen in the language standard!
- Provides interactive shell.
- Can interpret “macros”
(not compiled programs)
 - Rapid prototyping possible.
- ROOT provides also Python bindings:
 - Can use Python interpreter directly after a simple import ROOT !!



```
$ root -b  
root [0] 3 * 3  
(const int)9
```

* Starting from ROOT 6, the interpreter will be Cling (based on Clang)

Persistency (I/O)

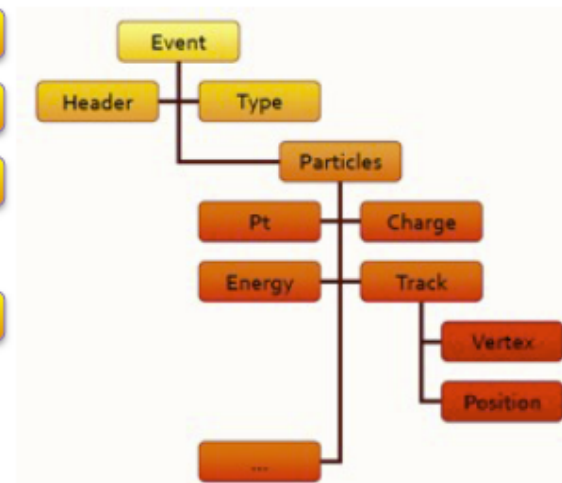
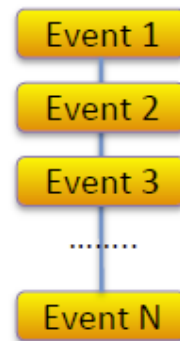
- ROOT offers the possibility to write C++ objects to file
 - Extraordinary: impossible in native C++!
 - Used for petabytes/year rates of LHC detectors.
 - Achieved with serialization of the object using the reflection provided by the CINT dictionary.
- Single objects, collections, object trees
 - Basically one method for all ROOT objects: `TObject::Write`

Cornerstone of the storage
of experimental data

x	y	z
-1.10228	-1.79939	4.452622
1.867178	-0.59662	3.842313
-0.52418	1.868521	3.766139
-0.38061	0.969128	1.084074
0.552454	-0.21231	0.350281
-0.18495	1.187305	1.443902
0.205643	-0.77015	0.635417
1.079222	-0.32739	1.271904
-0.27492	-1.72143	3.038899
2.047779	-0.06268	4.197329
-0.45868	-1.44322	2.293266
0.304731	-0.88464	0.875442
-0.71234	-0.22239	0.556881
-0.27187	1.181767	1.470484
0.866202	-0.65411	1.213209
-2.03555	0.527648	4.421883
-1.45905	-0.464	2.344113
1.230661	-0.00565	1.514559
		3.562347

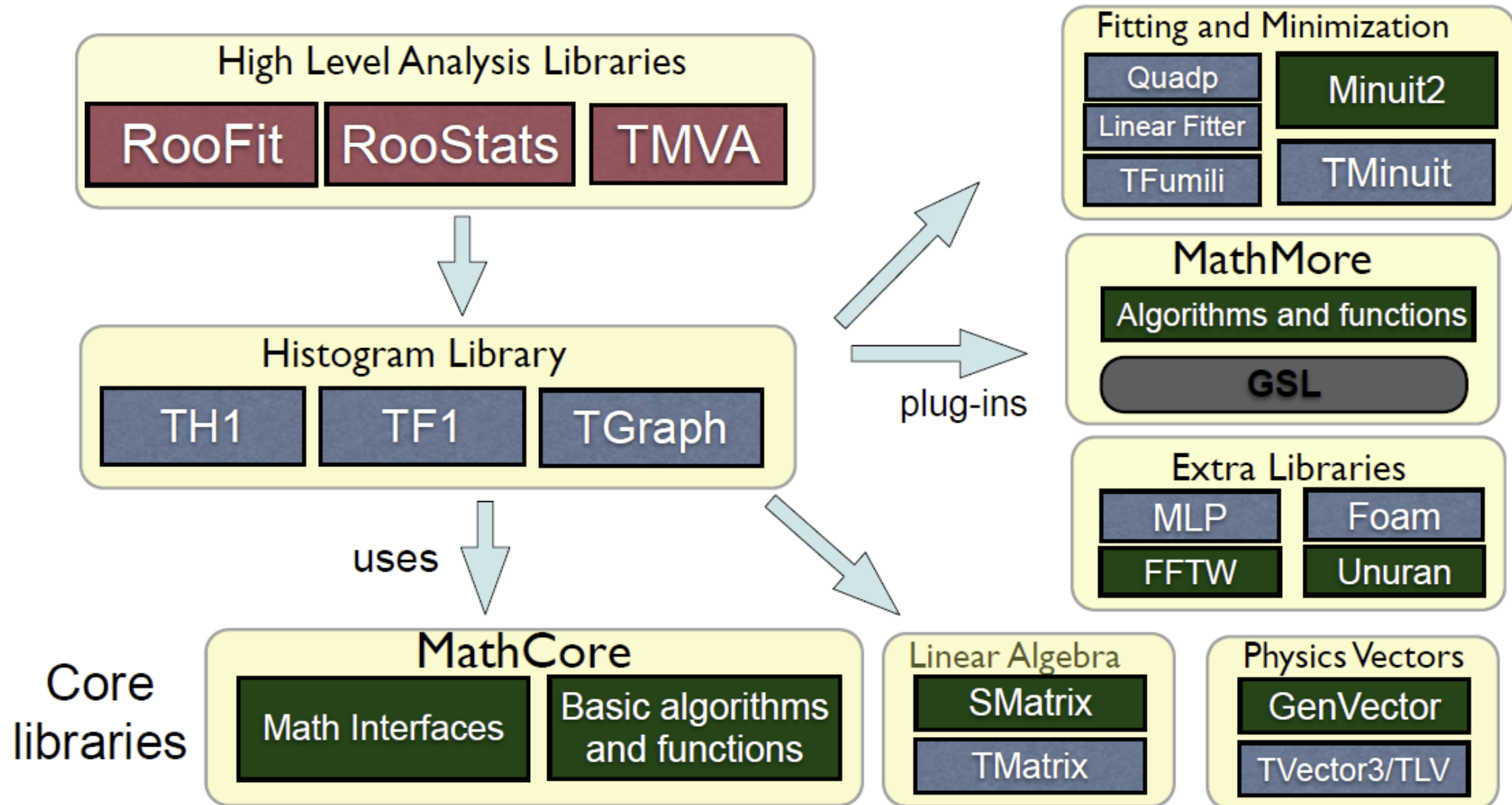


*LEP style flat n-tuples
evolved in more efficient
trees (fast access, read
ahead)*



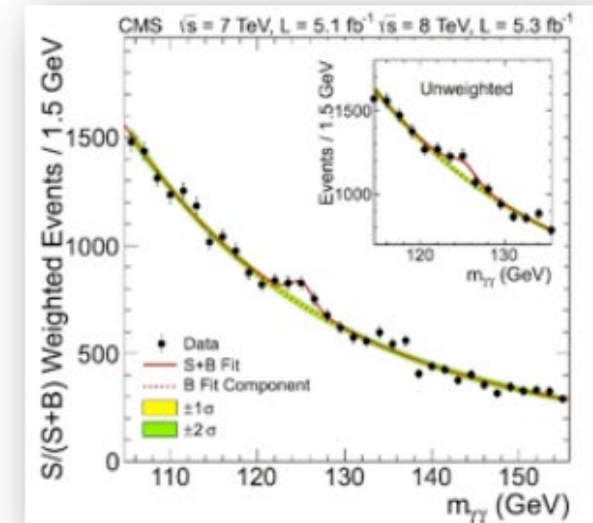
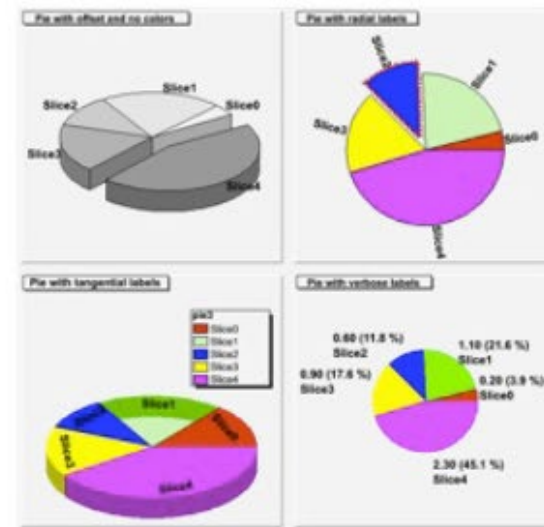
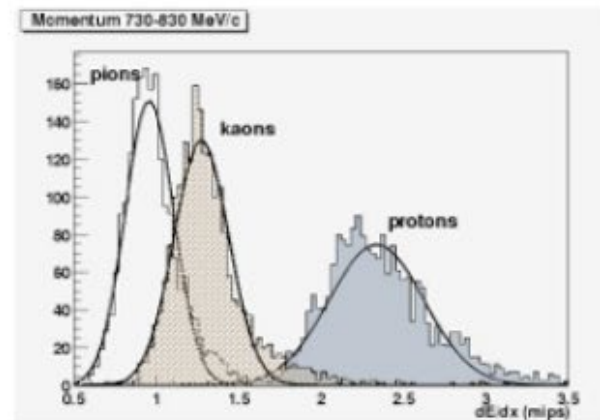
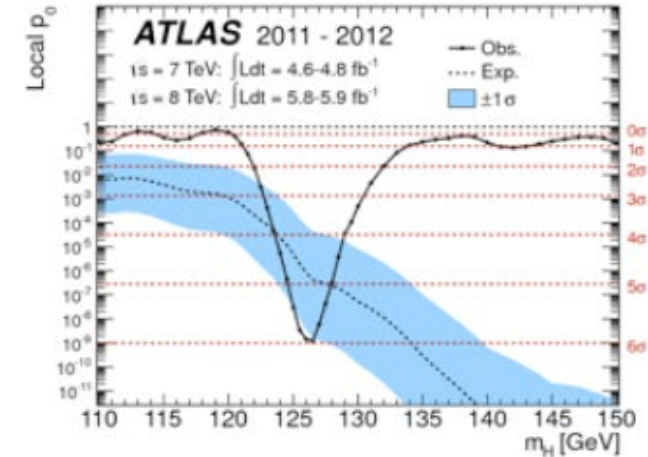
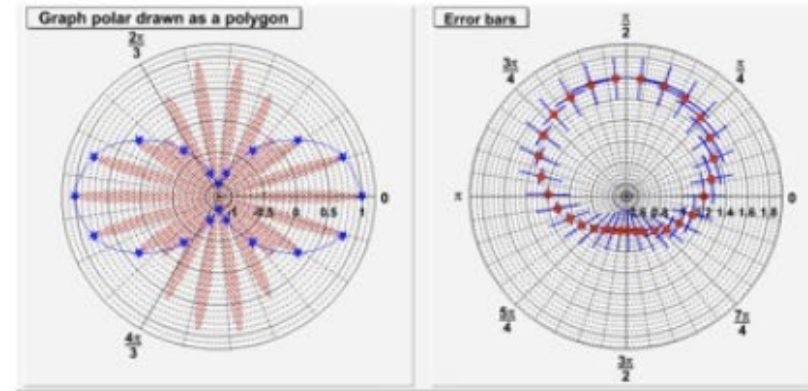
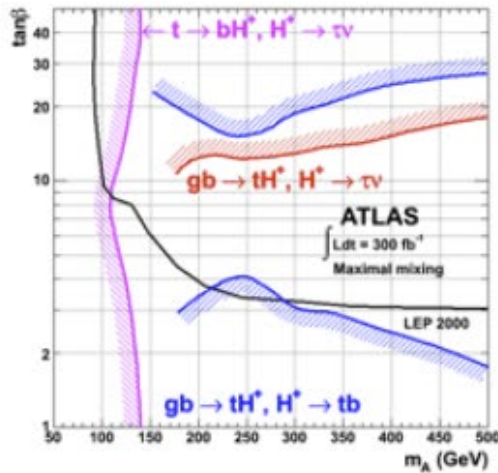
ROOT Math/Stat Libraries

- ROOT provides a reach set of mathematical libraries and tools needed for event reconstruction, simulation and statistical data analysis



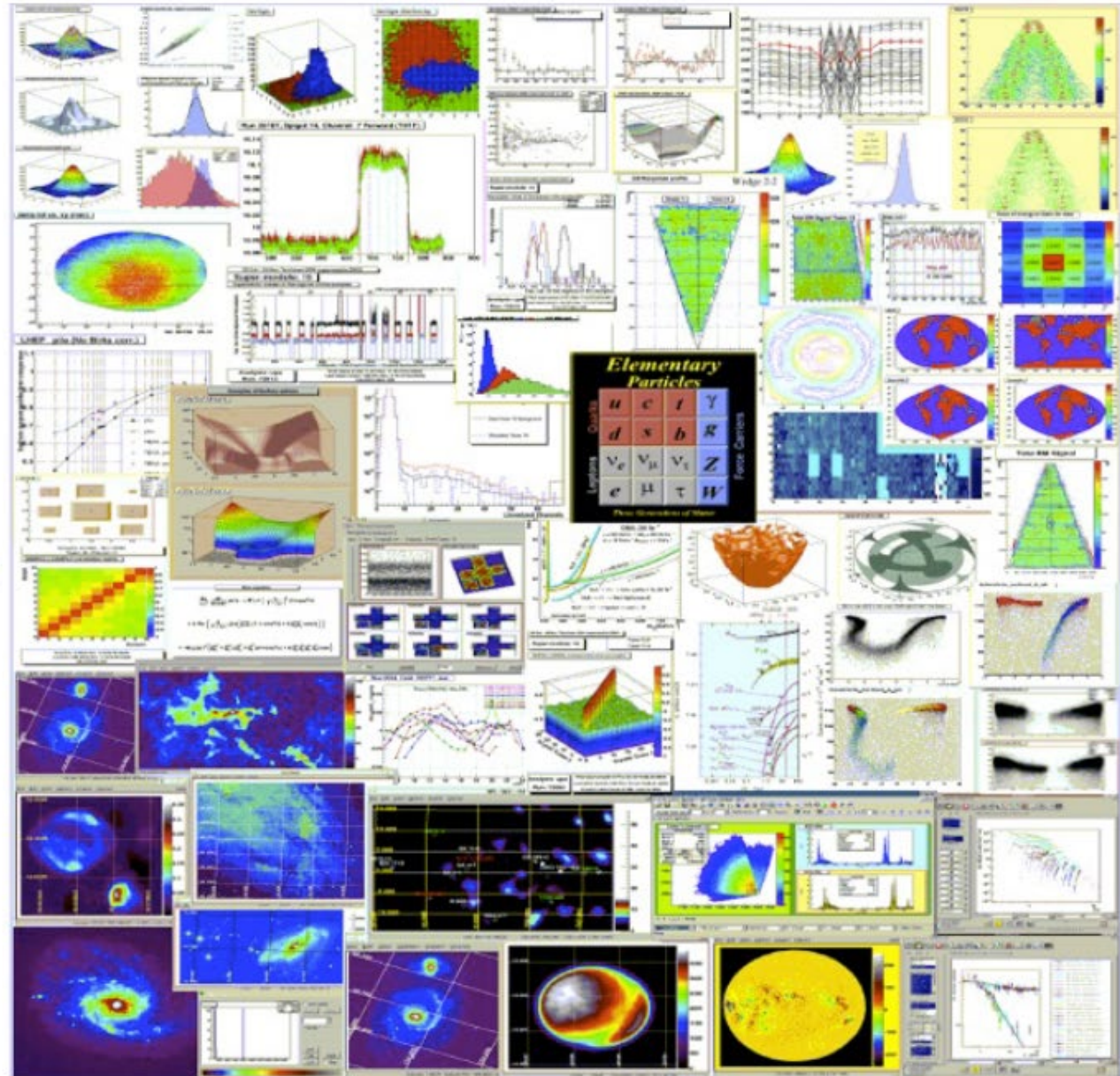
Graphics in ROOT

- Many plots available for data analysis:

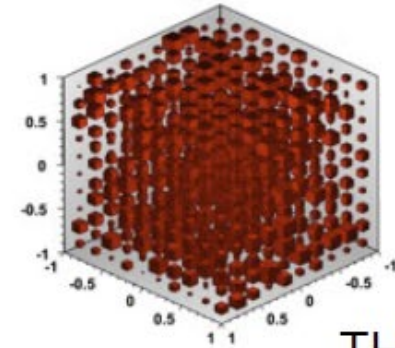
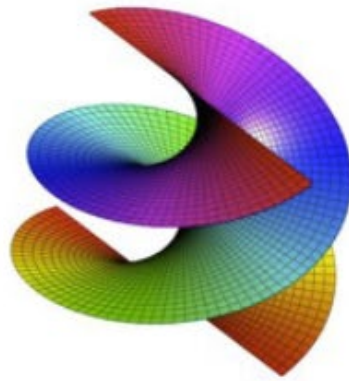
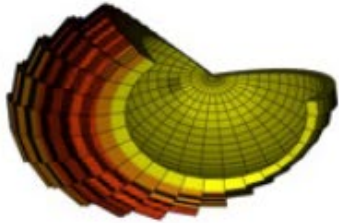


2D Graphics

- New functionality added at every new release
- Always new requests for new style of plots
- Can save graphics in many formats:
ps, pdf, svg, jpeg,
png, c, root

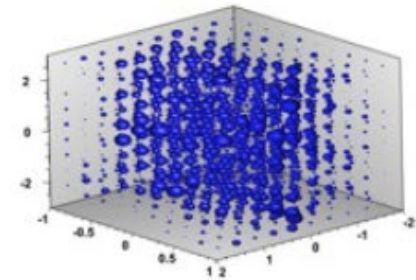
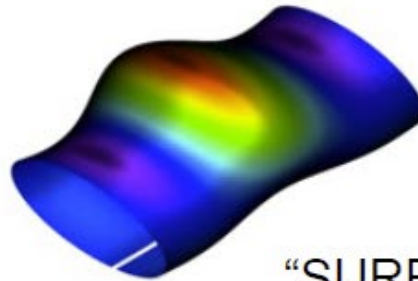
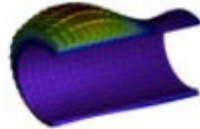
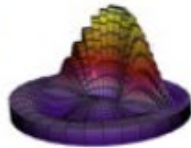
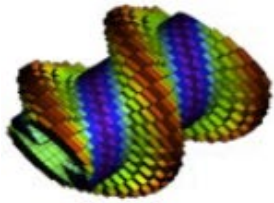


3D Graphics

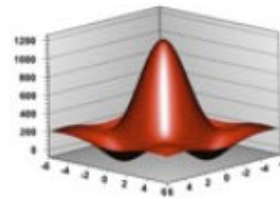
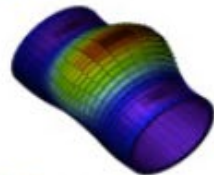
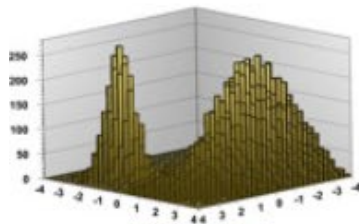


TH3

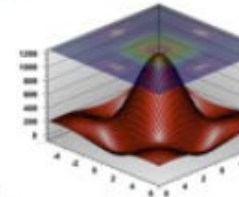
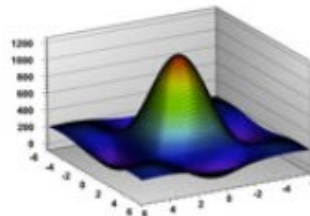
TGLParametric



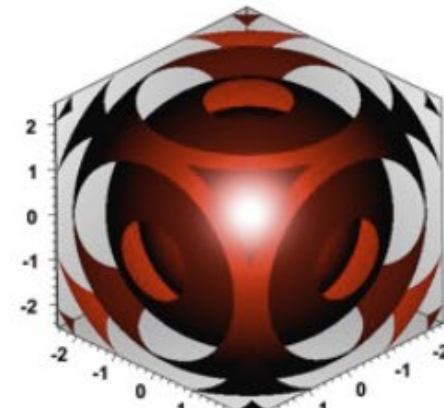
“LEGO”



“SURF”

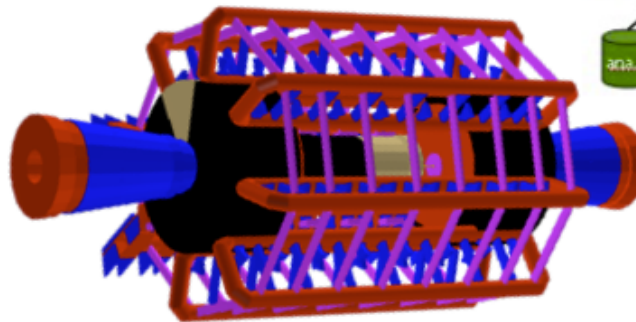
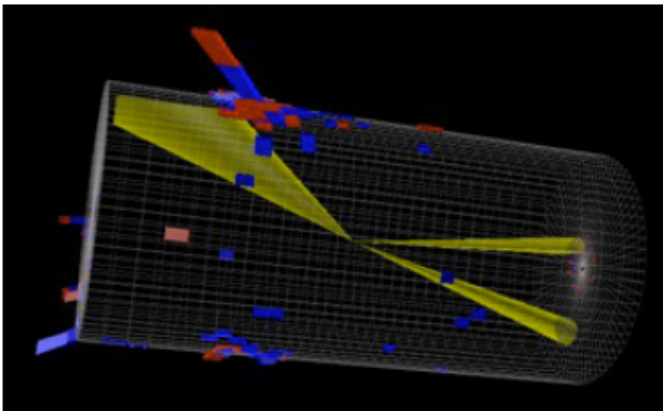


TF3

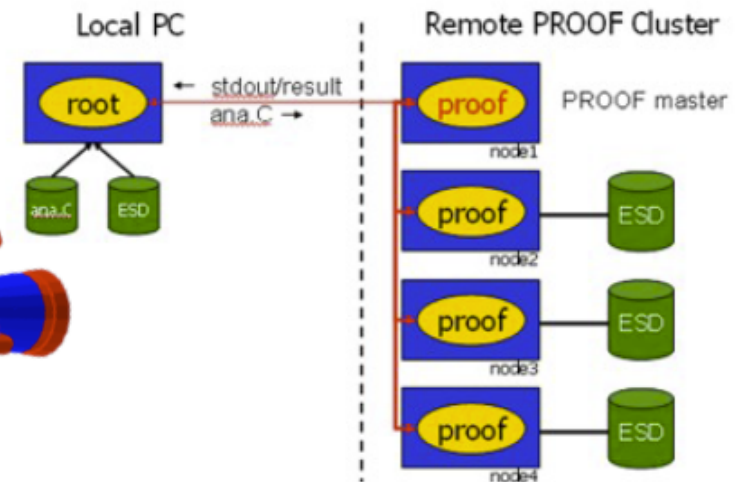


Other ROOT Features

- Geometry Toolkit
 - Represent geometries as complex as LHC detectors.
- Event Display (EVE)
 - Visualise particles collisions within detectors.
- PROOF: Parallel ROOT Facility
 - Multi-process approach to parallelism
 - A system to run ROOT queries in parallel on a large number of distributed computers
 - Proof-lite: does not need a farm, uses all the cores on a desktop machine



PROOF Schema



Getting Information

- **ROOT website:** <http://root.cern.ch>

You can find ROOT user references and guides there.

- **the ROOT mailing list:** roottalk@root.cern.ch

- **subscribe to the mailing list:**

<http://root.cern.ch/root/Registration.phtml>

- **forums and archives:**

<http://root.cern.ch/root/roottalk/AboutRootTalk.html>

<http://root.cern.ch/phpBB2/>

Summary

- CERN represents excellence in the field of Particle Physics research. To cope with the needs, it always pushed innovation in IT technologies.
- ROOT: complete analysis framework offering a C++ interpreter, a powerful persistency mechanism, advanced mathematical tools and even a parallel facility
 - it is designed for data analysis of very large shared data sets
 - it is the result of 18 years of cooperation between the development team and thousands of heterogeneous users
 - it is used extensively and profitably by physicists
 - all Higgs discovery plots made with ROOT
- We will see now how to start using ROOT !