Apple at the CERN LHC

Bleeding Edge Physics and Bleeding Edge Computing

Dr. Fons Rademakers
Senior Scientist, ALICE Software Architect
CERN

The CERN Large Hadron Collider – LHC

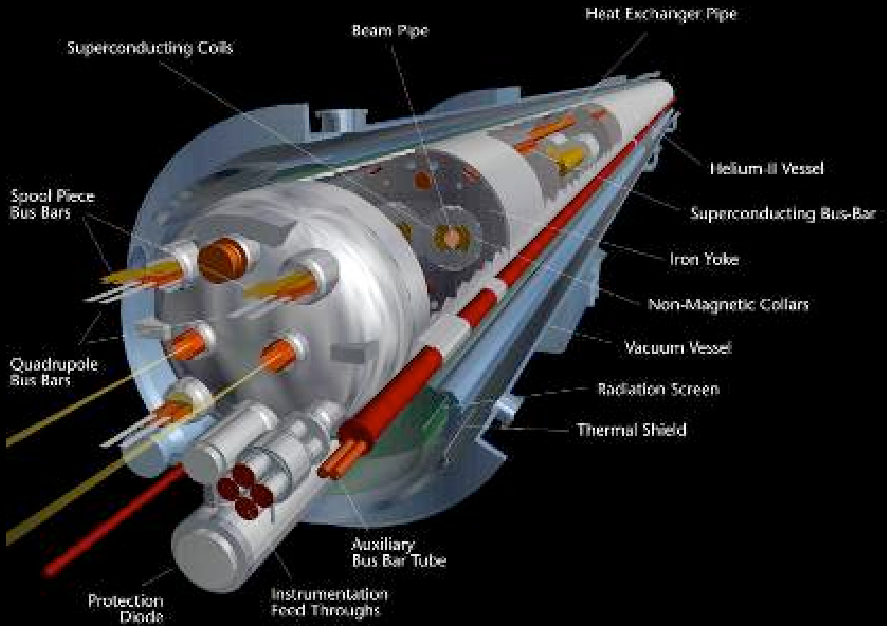
- LHC is a proton-proton and heavy ion collider
- Proton-proton center-of-mass energy $\sqrt{s_{pp}} = 14$ TeV
- Scheduled start: Sept 2007

The CERN Large Hadron Collider – LHC

- LHC is a proton-proton and heavy ion collider
- Proton-proton center-of-mass energy $\sqrt{s_{pp}} = 14$ TeV
- Scheduled start: Sept 2007

The 15-m Long LHC Cryodipole

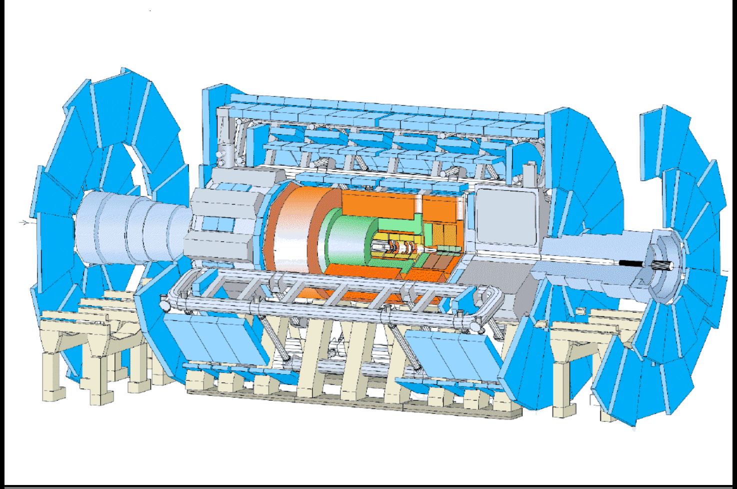
- High energy protons are bent in the ring by 1232 super-conducting magnets that provide a field never reached before of 8.3T at 1.9K (-271^{0}C)
- The largest cryogenic system in the world



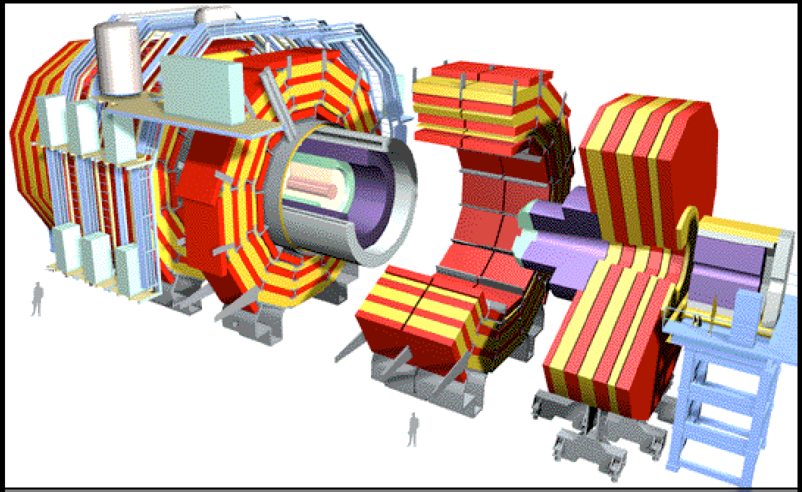
The LHC Detectors

The LHC Detectors

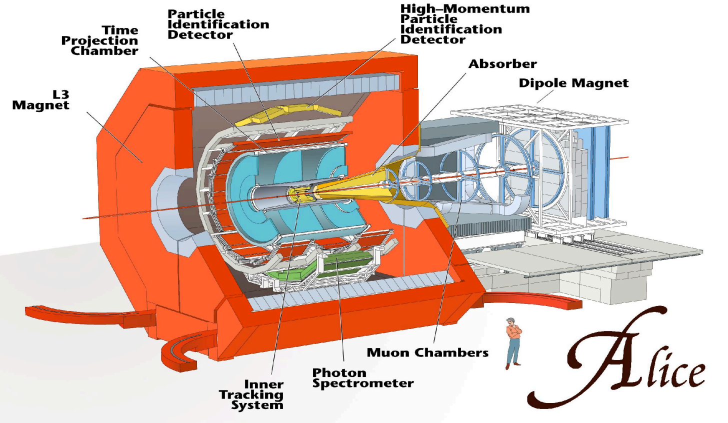
ATLAS



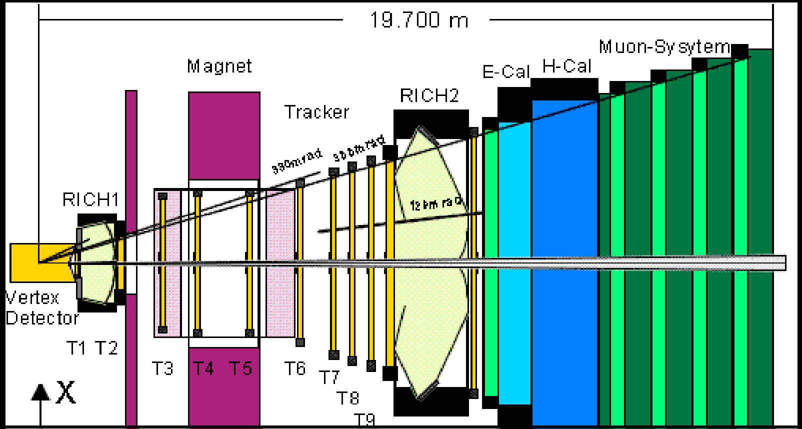
CMS



Alice

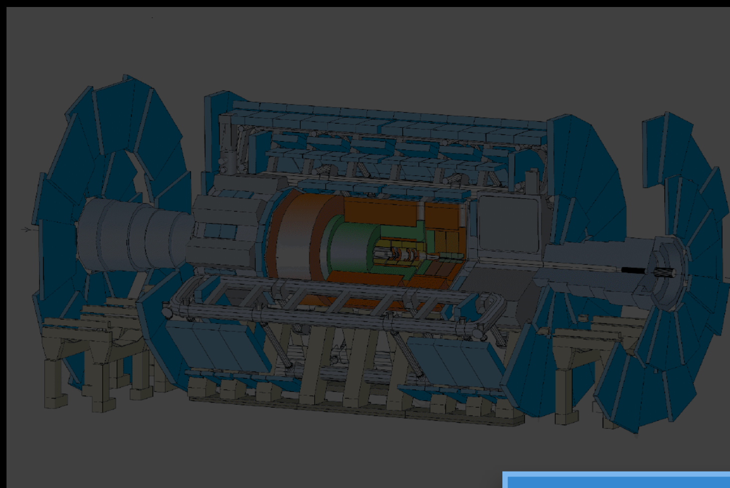


LHCb



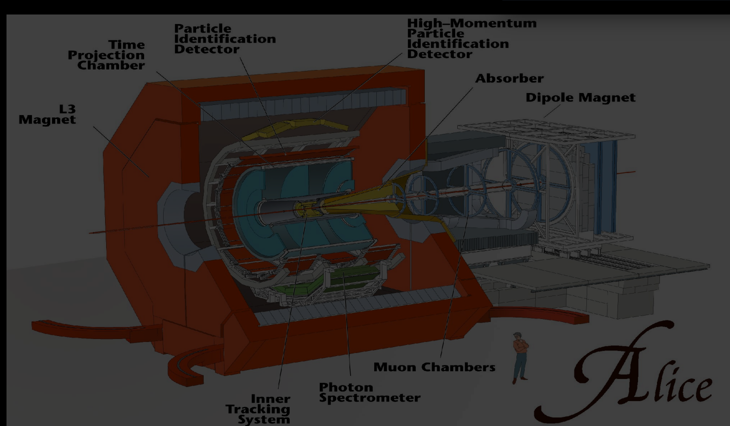
The LHC Detectors

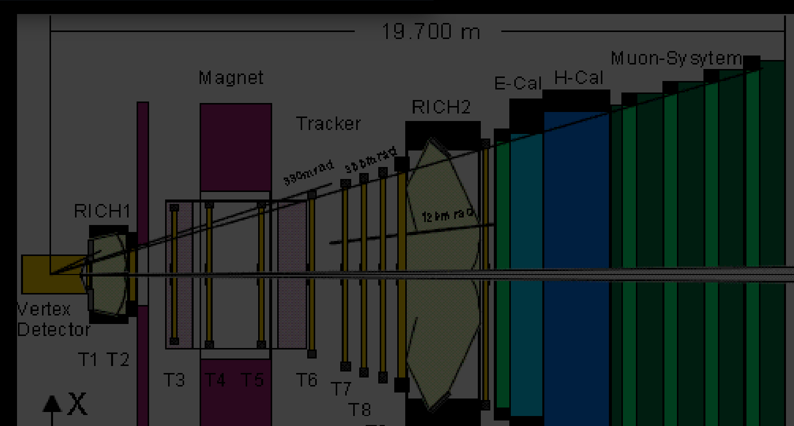
ATLAS



CMS

Alice





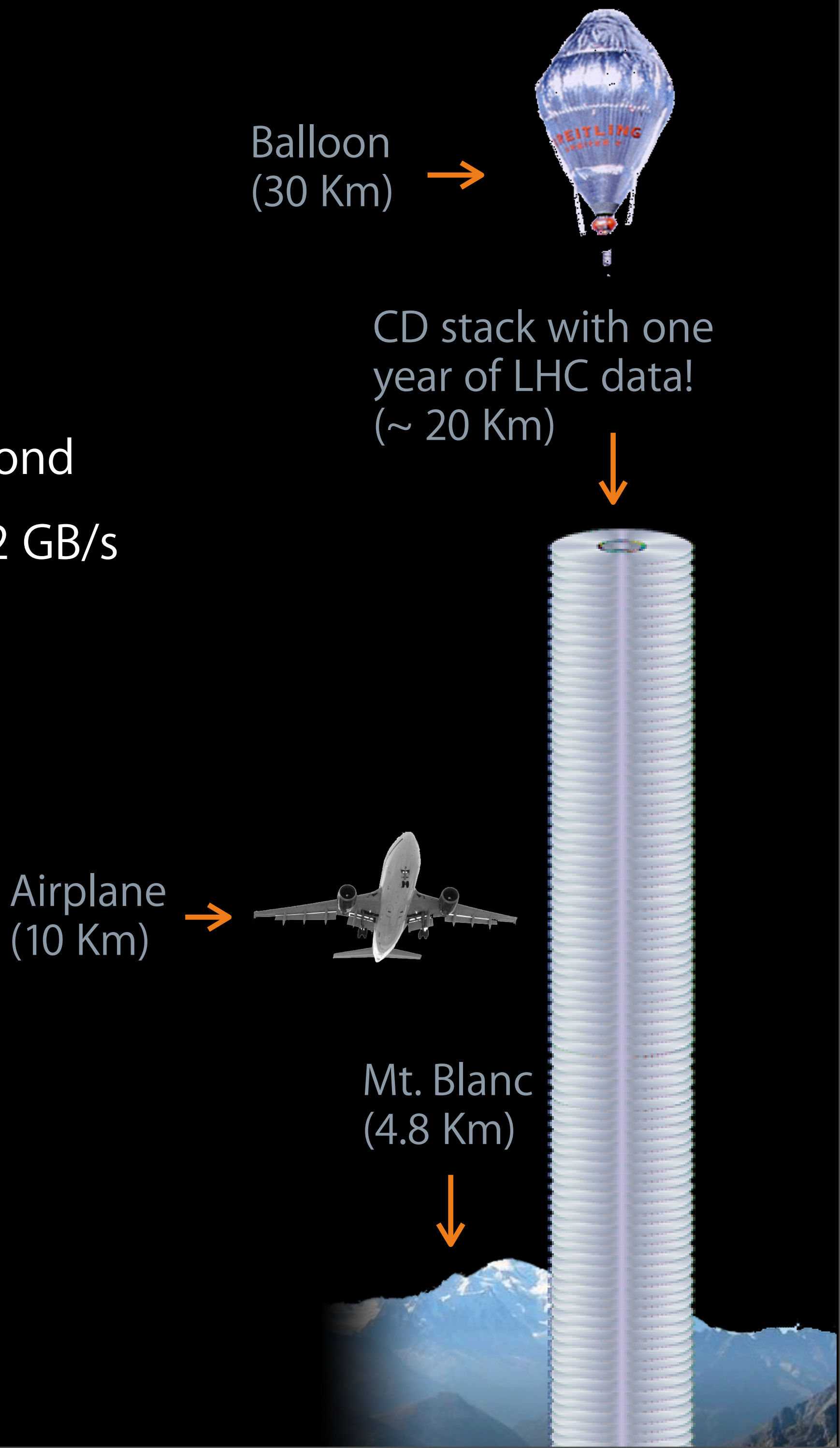
~8-10 PetaBytes/ year
~10^{10} events/year
~10^3 batch and interactive users

LHC Data

- The LHC generates:
  - 40 million collisions per second
- Combined the 4 experiments record:
  - After filtering, 100 interesting collision per second
  - From 1 to 12 MB per collision ⇒ from 0.1 to 1.2 GB/s
  - 10^{10} collisions registered every year
  - ~ 10 PetaBytes (10^{15} B) per year
  - LHC data correspond to
    20 millions CD per year!
  - Computing power equivalent to
    100.000 of today's PC
  - Space equivalent to 400.000 large PC disks

LHC Data

- The LHC generates:
  - 40 million collisions per second
- Combined the 4 experiments record:
  - After filtering, 100 interesting collision per second
  - From 1 to 12 MB per collision ⇒ from 0.1 to 1.2 GB/s
  - 10^{10} collisions registered every year
  - ~ 10 PetaBytes (10^{15} B) per year
  - LHC data correspond to 20 millions CD per year!
  - Computing power equivalent to 100.000 of today's PC
  - Space equivalent to 400.000 large PC disks



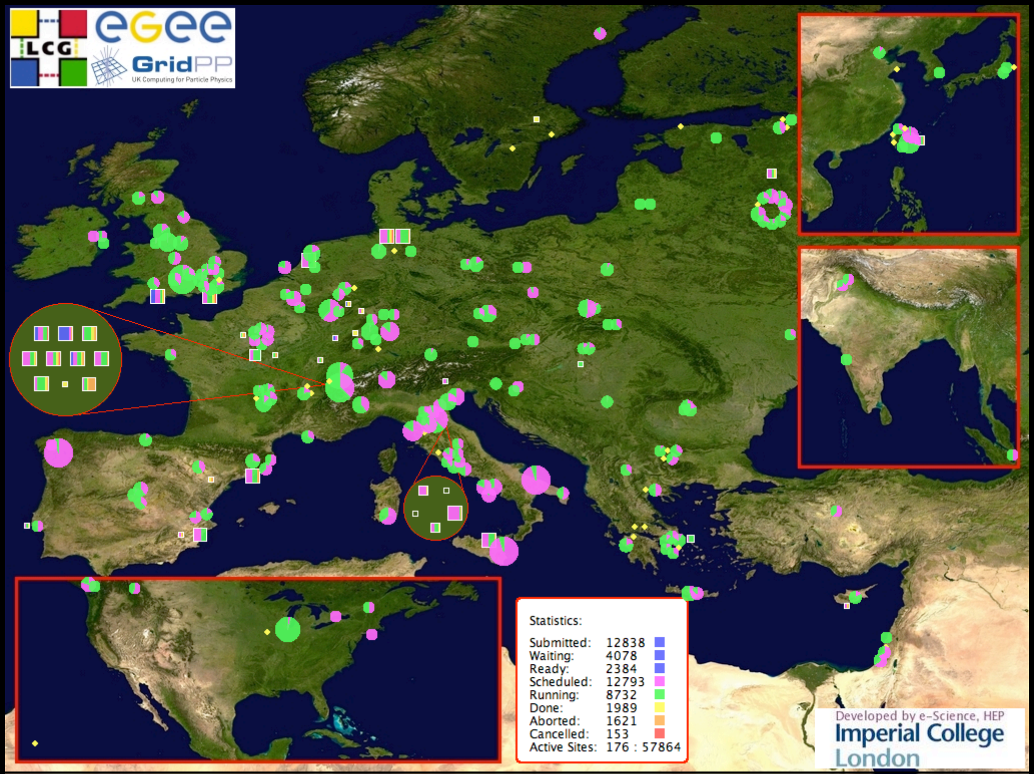
LHC Data Processing Needs the Grid

- The LHC computing needs are enormous
- Computing resources cannot be concentrated in a single center
  - Nations prefers local investments
  - Competences are naturally distributed
  - Not enough real-estate to house all the computers at CERN
- Resources will be collocated in centers of different dimensions, running different hardware and OS's
- The Grid middleware integrates all these distributed centers into one large virtual center
- Actually we will need Grids of Grids

The EGEE Grid Project

"Enabling Grids for E-sciencE" started as an EU funded project, now world-wide deployed

- >180 sites
- >15 000 CPUs
- ~14 000 jobs completed per day

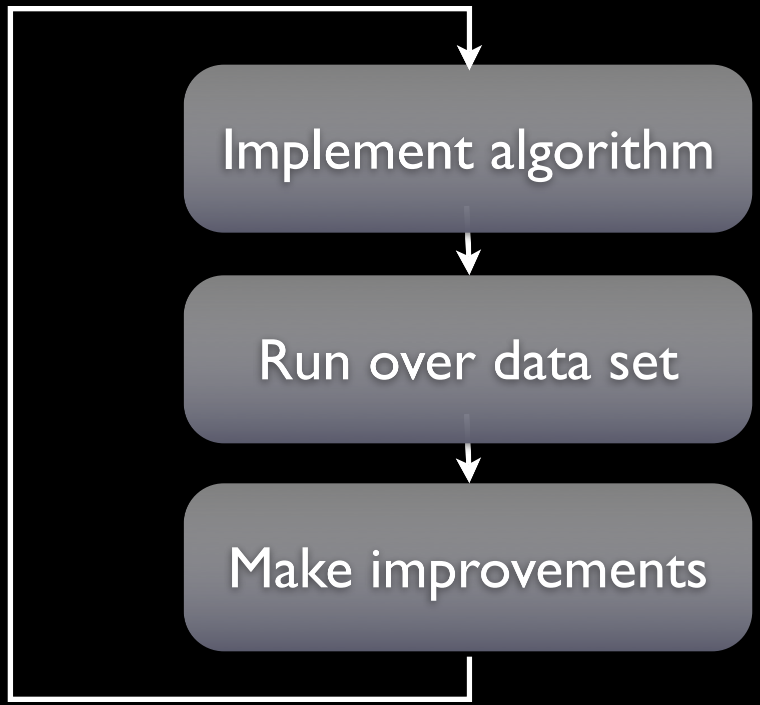


Apple in the Grid

Apple in the Grid

HEP Data Analysis

- Typical HEP analysis needs a continuous algorithm refinement cycle



HEP Data Analysis

- Ranging from I/O bound to CPU bound
- Need many disks to get the needed I/O rate
- Need many CPUs for processing
- Need memory to cache as much as possible

Some ALICE Numbers

- 1.5 PB of raw data per year
- 360 TB of ESD+AOD per year (20% of raw)
- One pass using 400 disks at 15 MB/s will take 16 hours

Some ALICE Numbers

- 1.5 PB of raw data per year
- 360 TB of ESD+AOD per year (20% of raw)
- One pass using 400 disks at 15 MB/s will take 16 hours

Using parallelism is the only way to analyze this amount of data in a reasonable amount of time

PROOF Design Goals

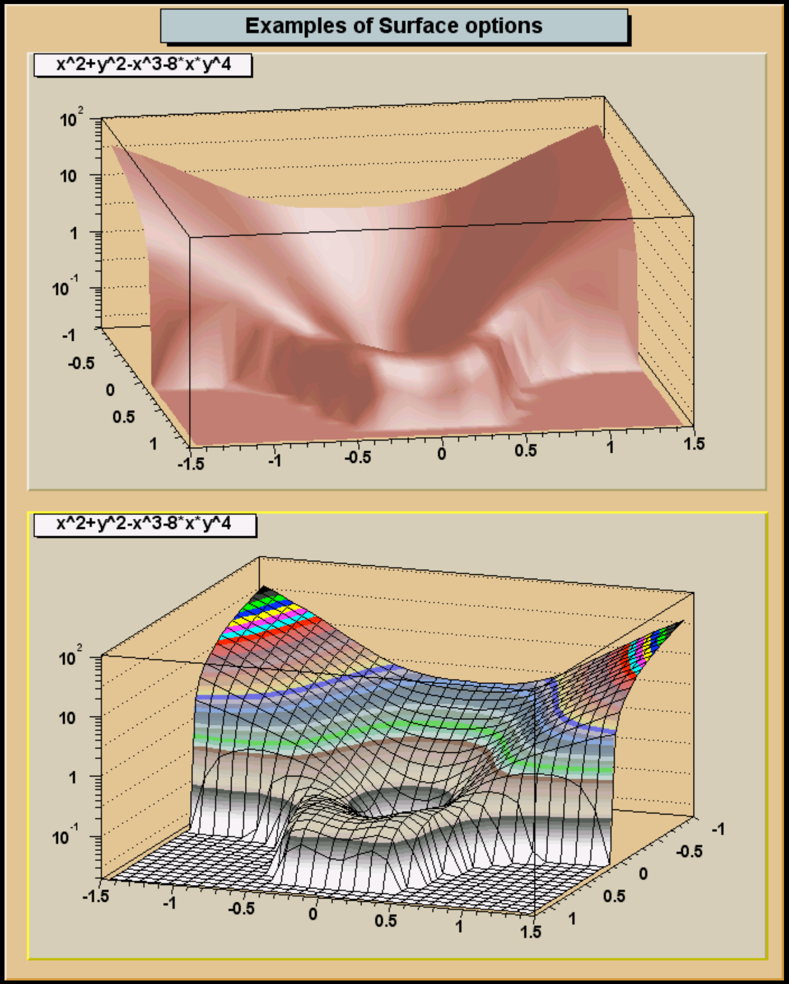
- System for running ROOT queries in parallel on a large number of distributed computers or multi-core machines
- PROOF is designed to be a transparent, scalable and adaptable extension of the local interactive ROOT analysis session
- Extends the interactive model to long running “interactive batch” queries

The ROOT Data Analysis Framework

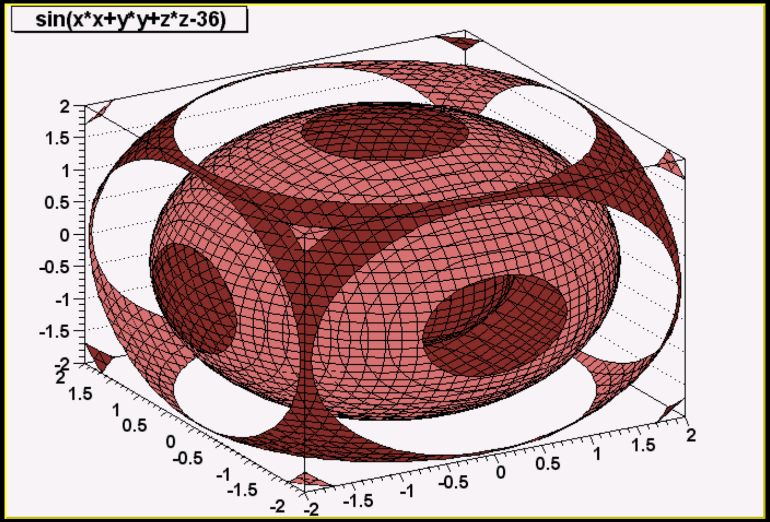
- ROOT is a large Object-Oriented data handling and analysis framework
  - Efficient object data store scaling from KB's to PB's
  - C++ interpreter
  - Extensive 2D+3D scientific data visualization capabilities
  - Extensive set of data fitting and modeling methods
  - Complete set of GUI widgets
  - Classes for threading, shared memory, networking, etc.
  - Fully cross platform, Windows, Unix/Linux, Mac OS X
  - 1.5 million lines of C++
  - Licensed under the LGPL
- Used by all HEP experiments in the world
- Used in many other scientific fields and in commercial world

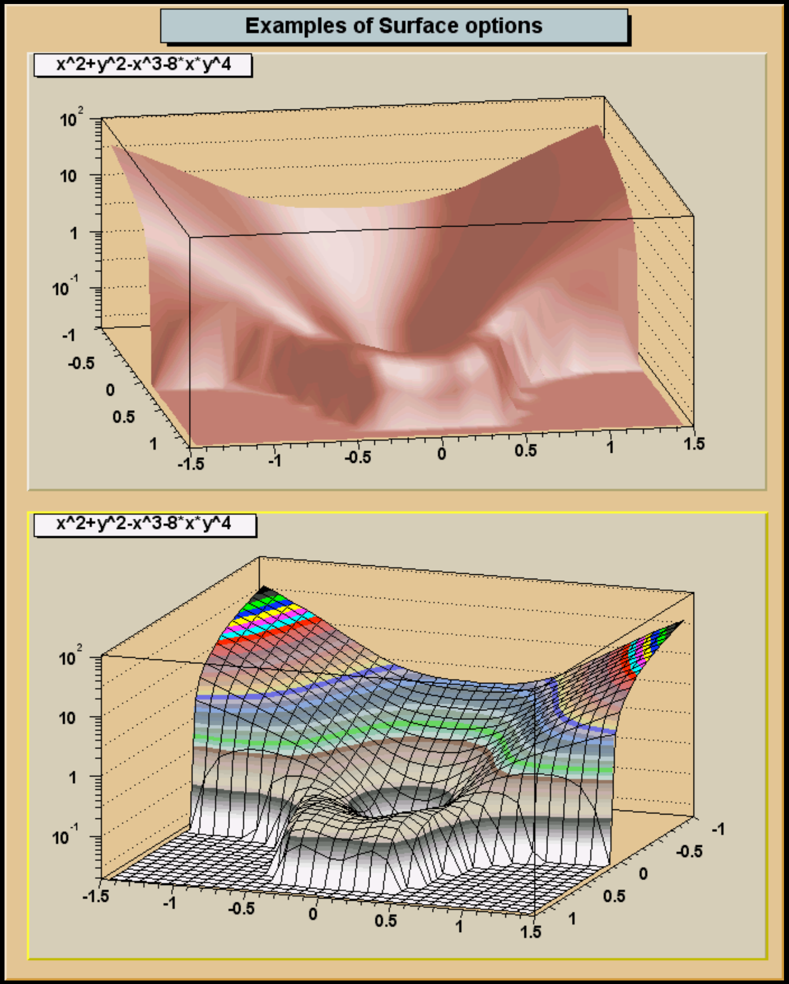
ROOT Image Gallery

ROOT Image Gallery

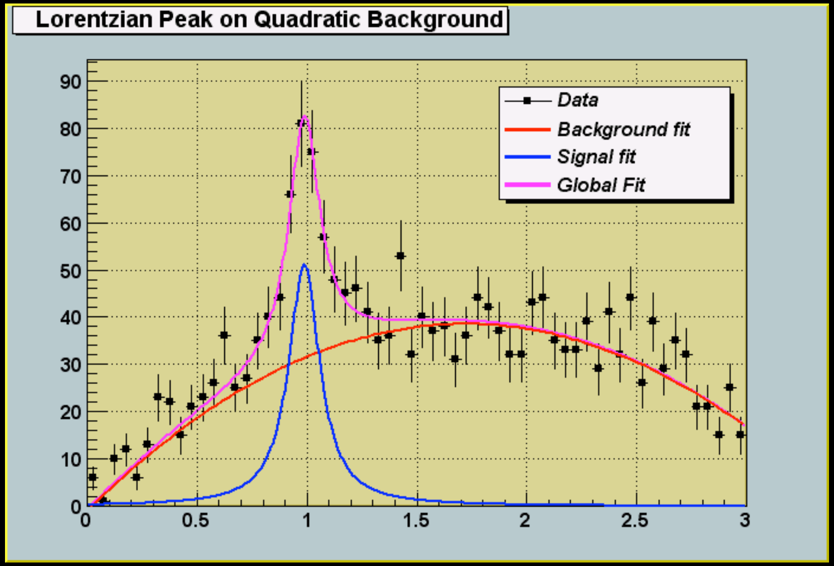


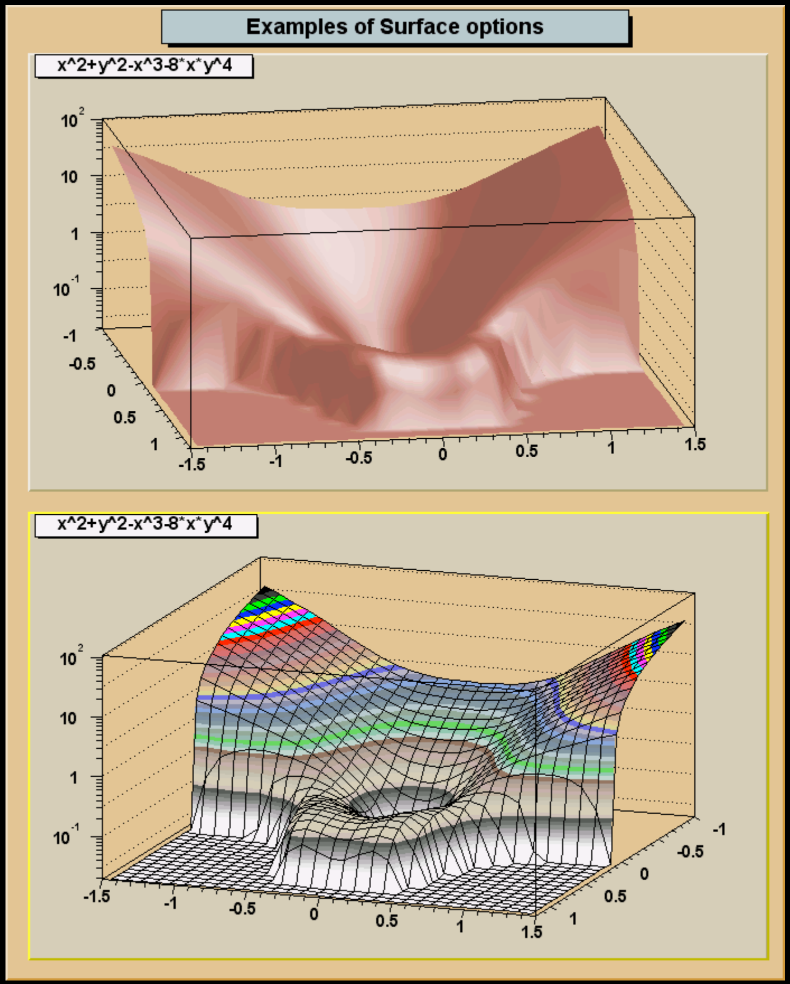
ROOT Image Gallery

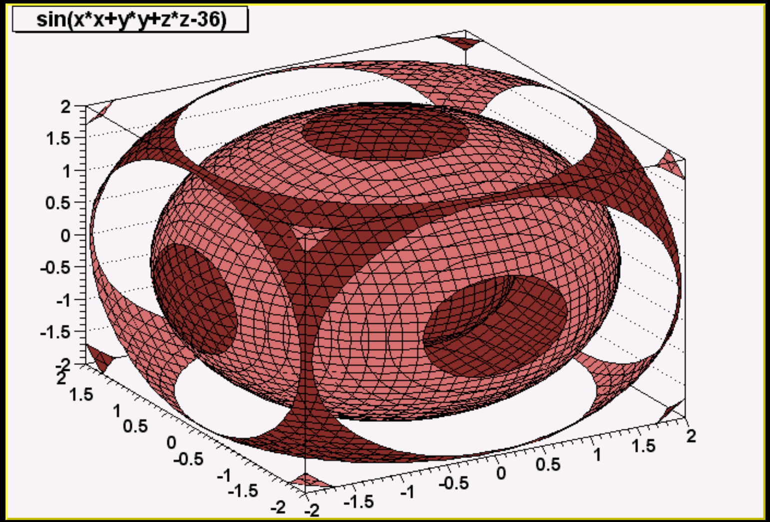




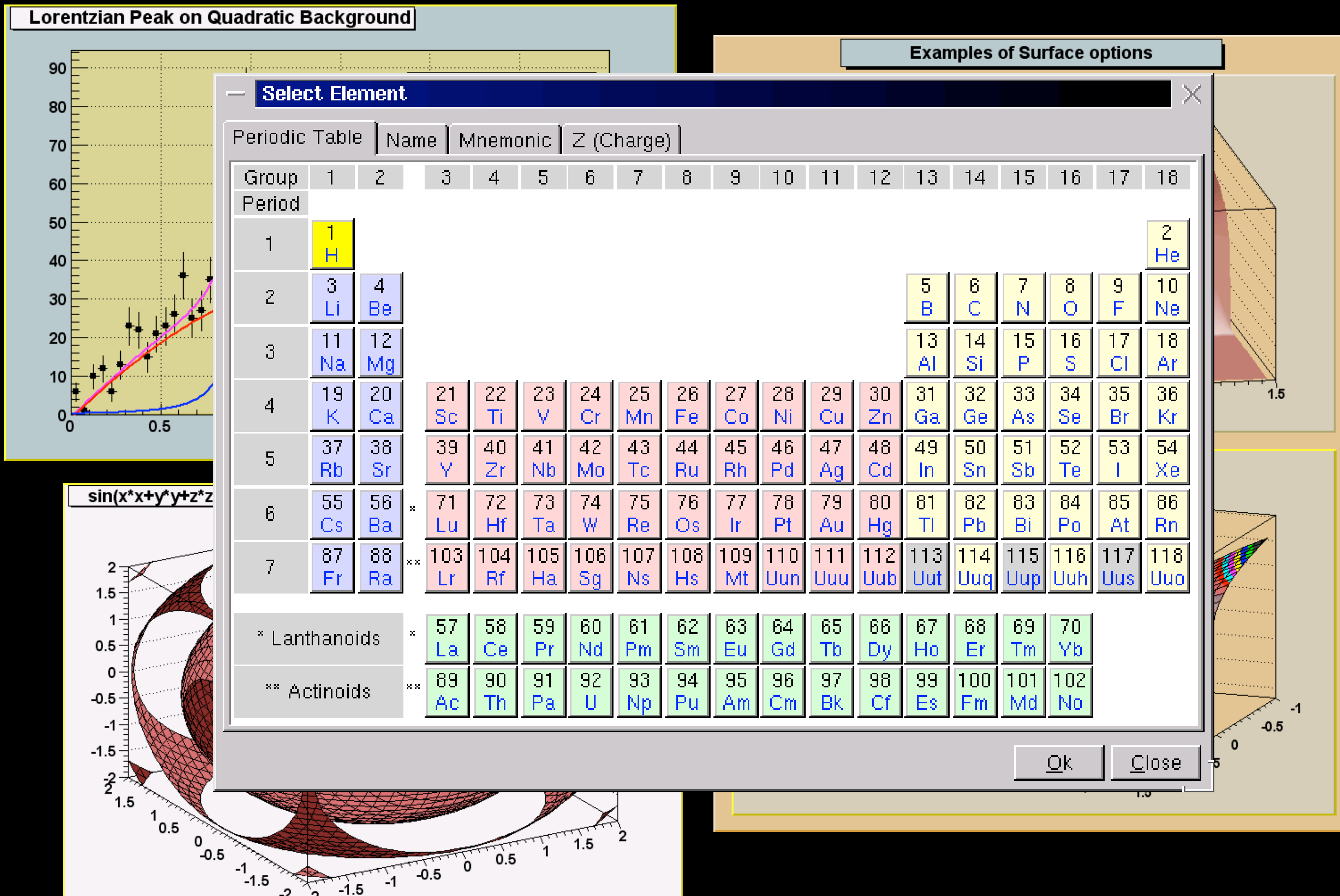
ROOT Image Gallery



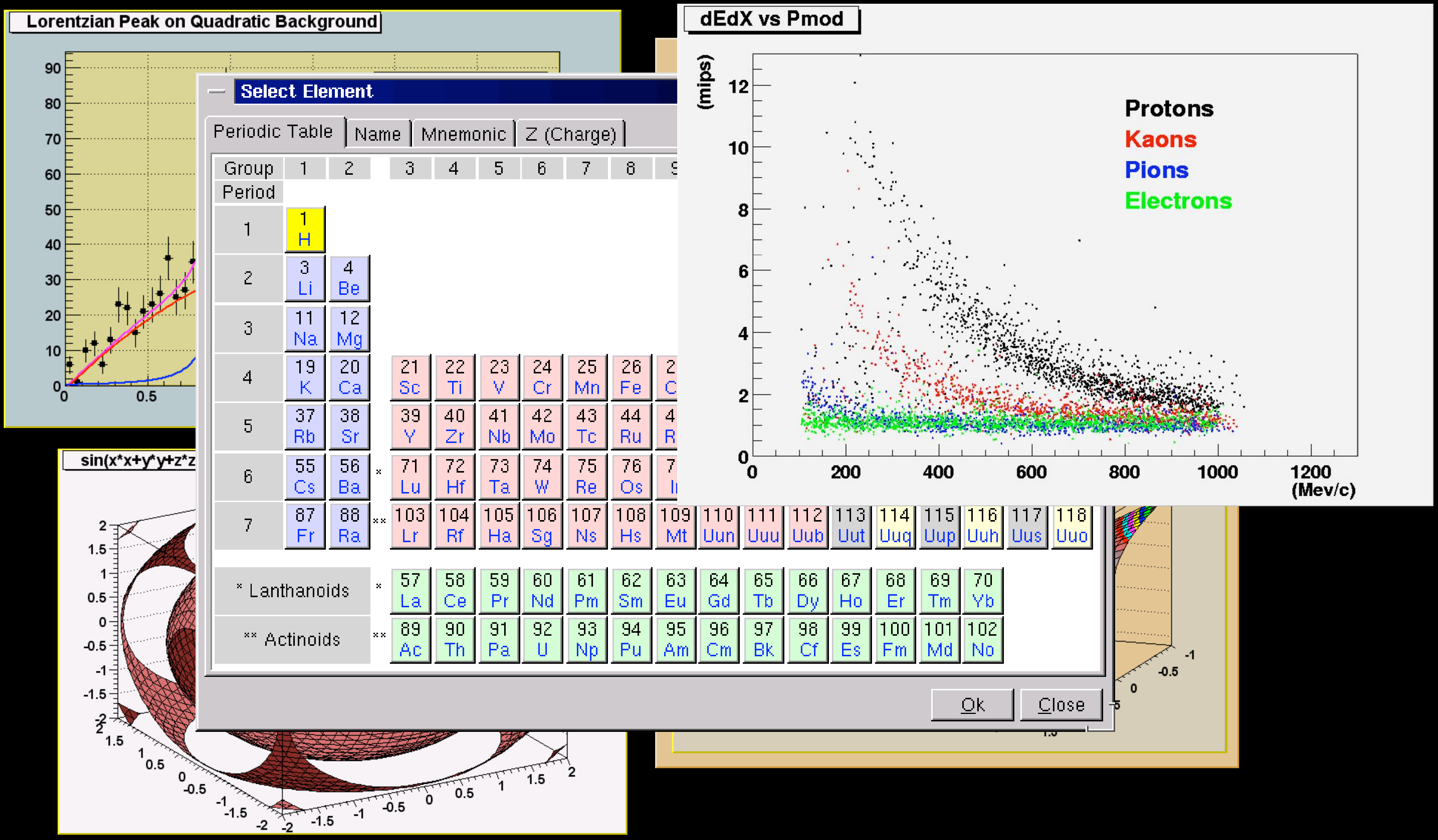




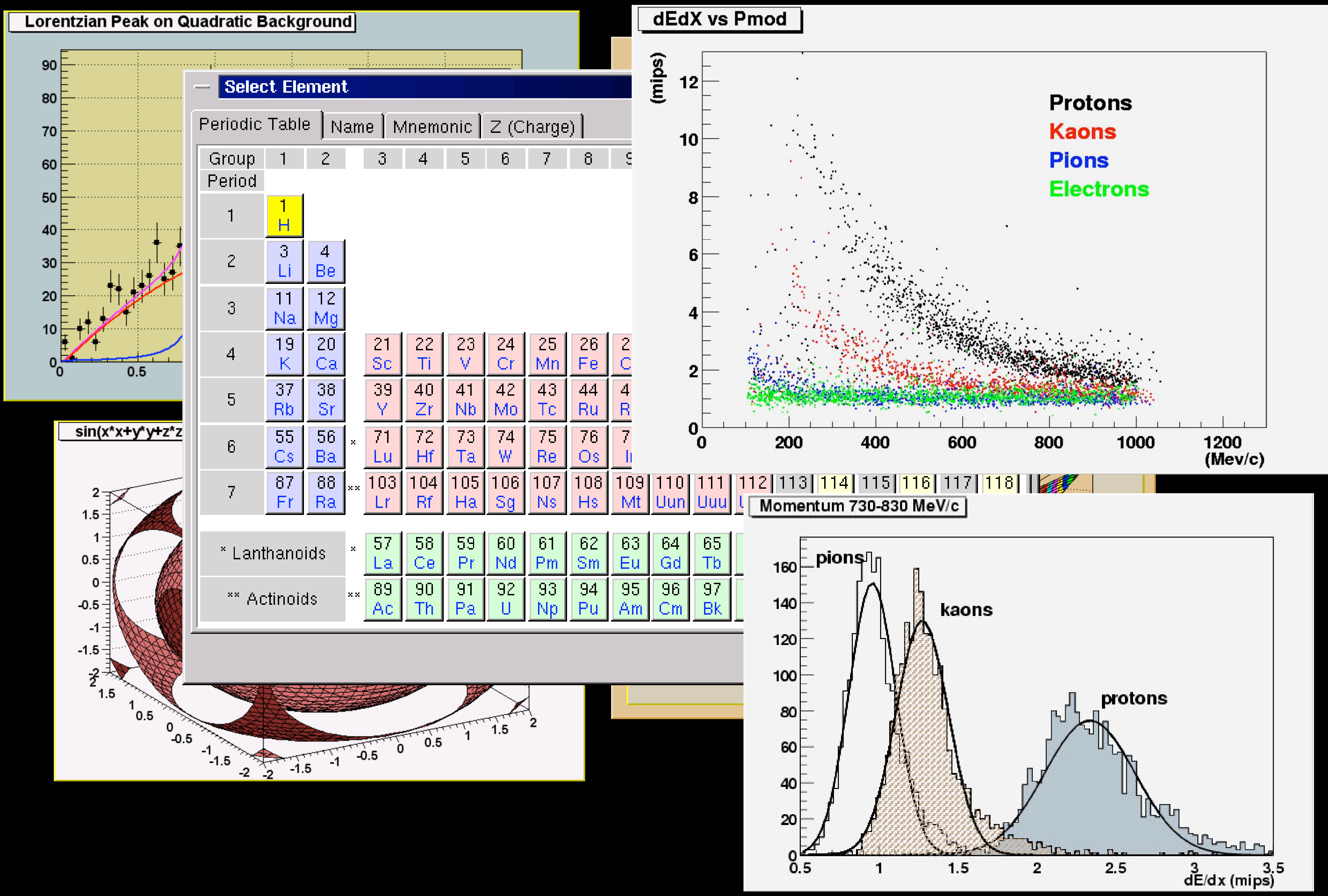
ROOT Image Gallery



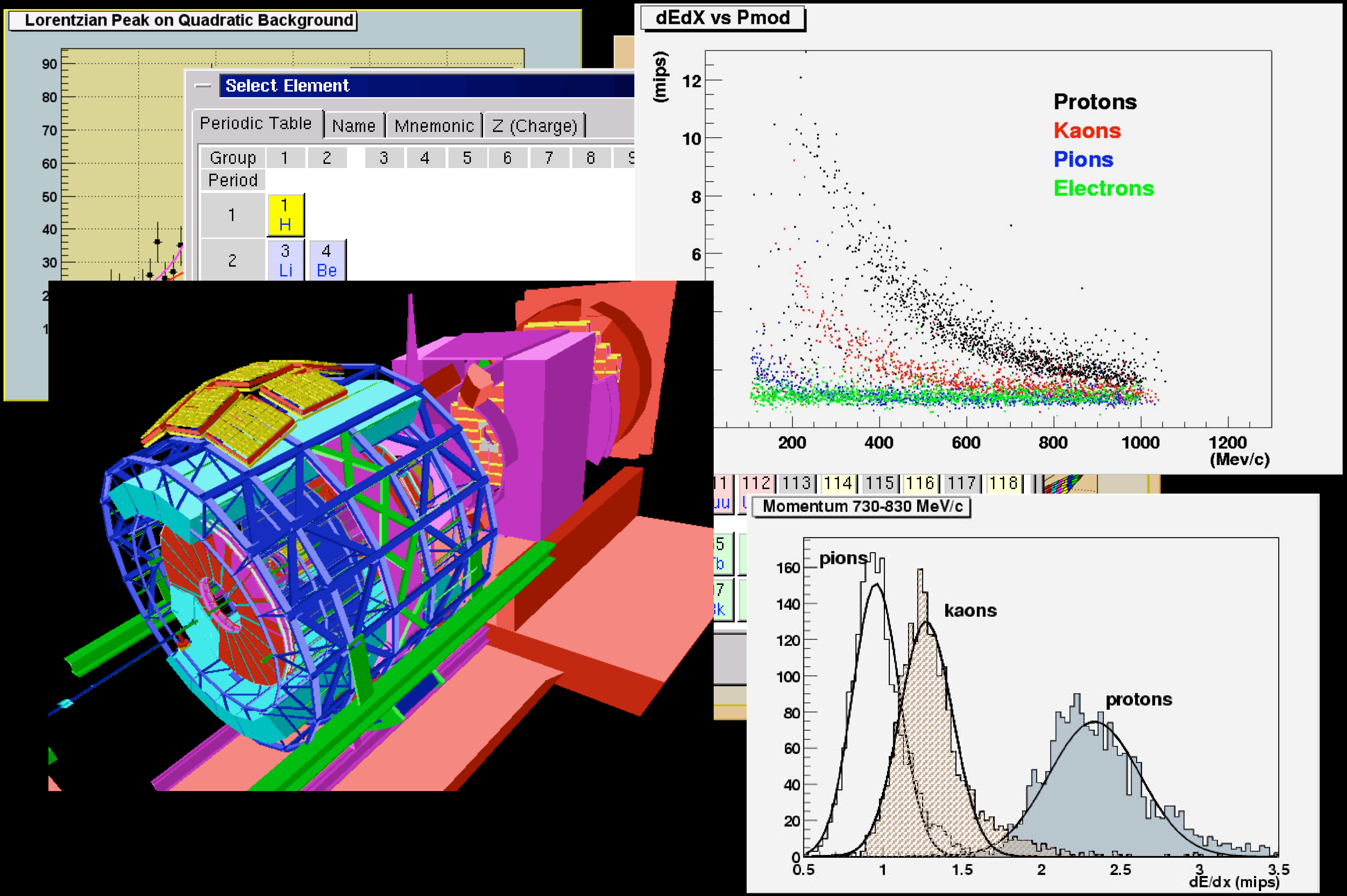
ROOT Image Gallery



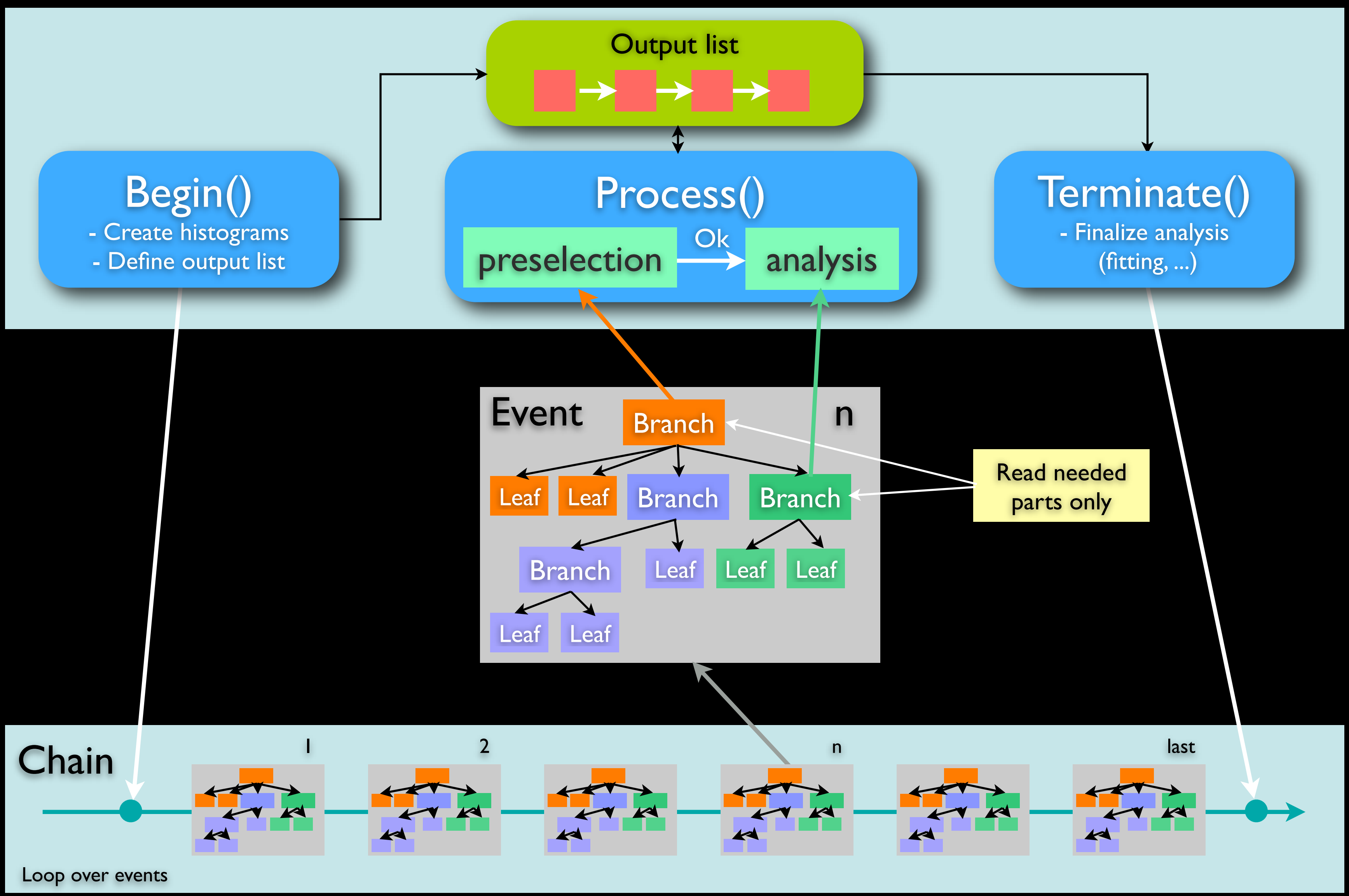
ROOT Image Gallery



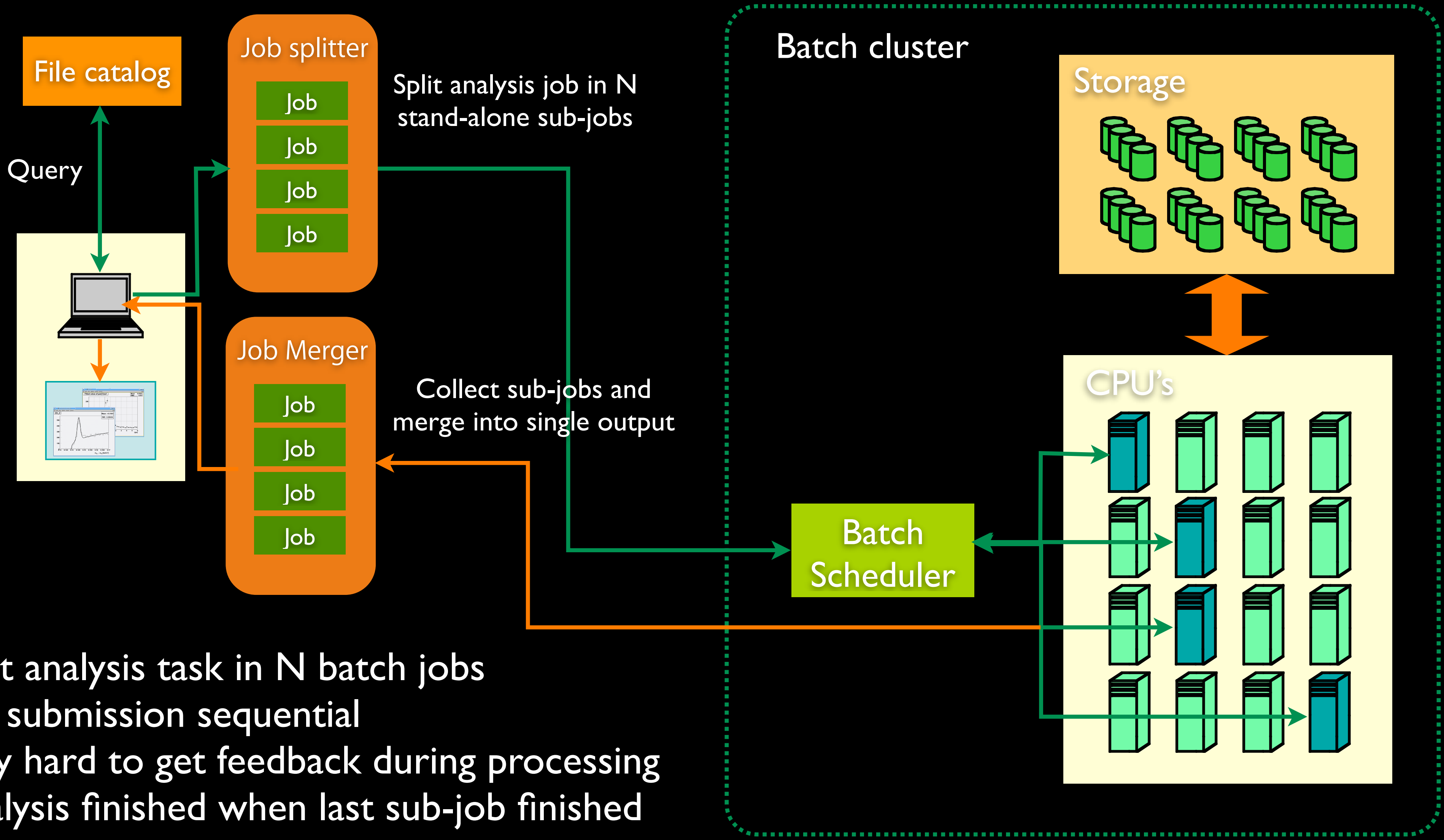
ROOT Image Gallery



The ROOT Data Model
Trees & Selectors

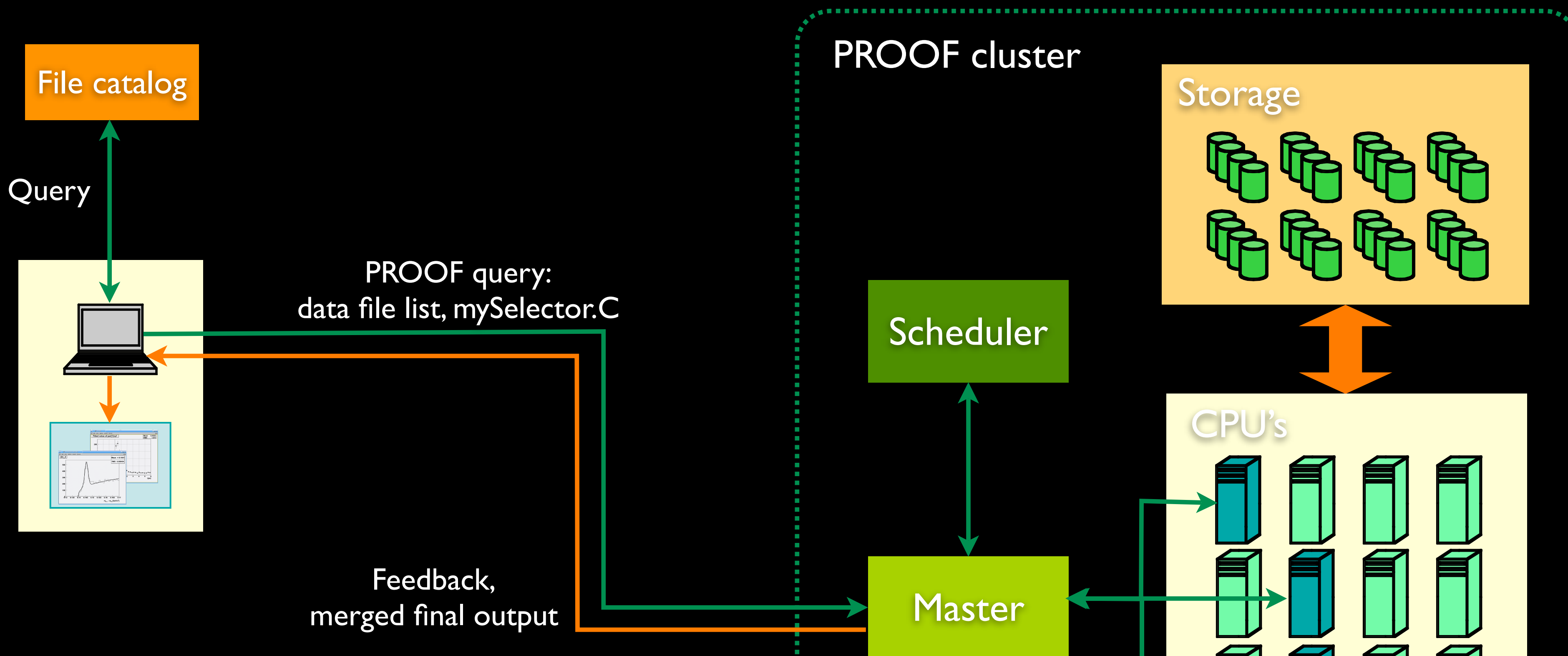


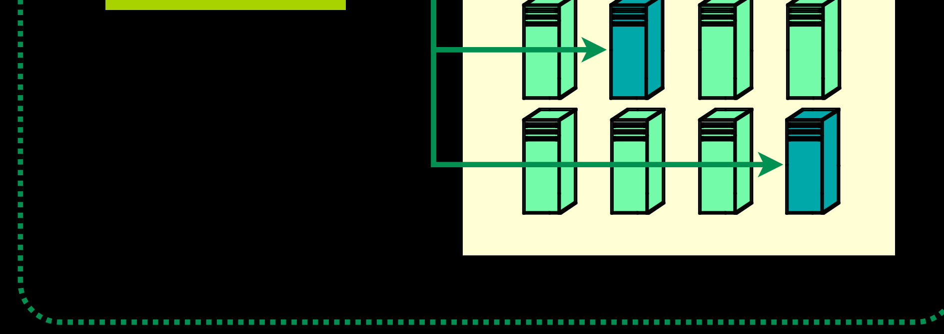
The Traditional Batch Approach



- Split analysis task in N batch jobs
- Job submission sequential
- Very hard to get feedback during processing
- Analysis finished when last sub-job finished

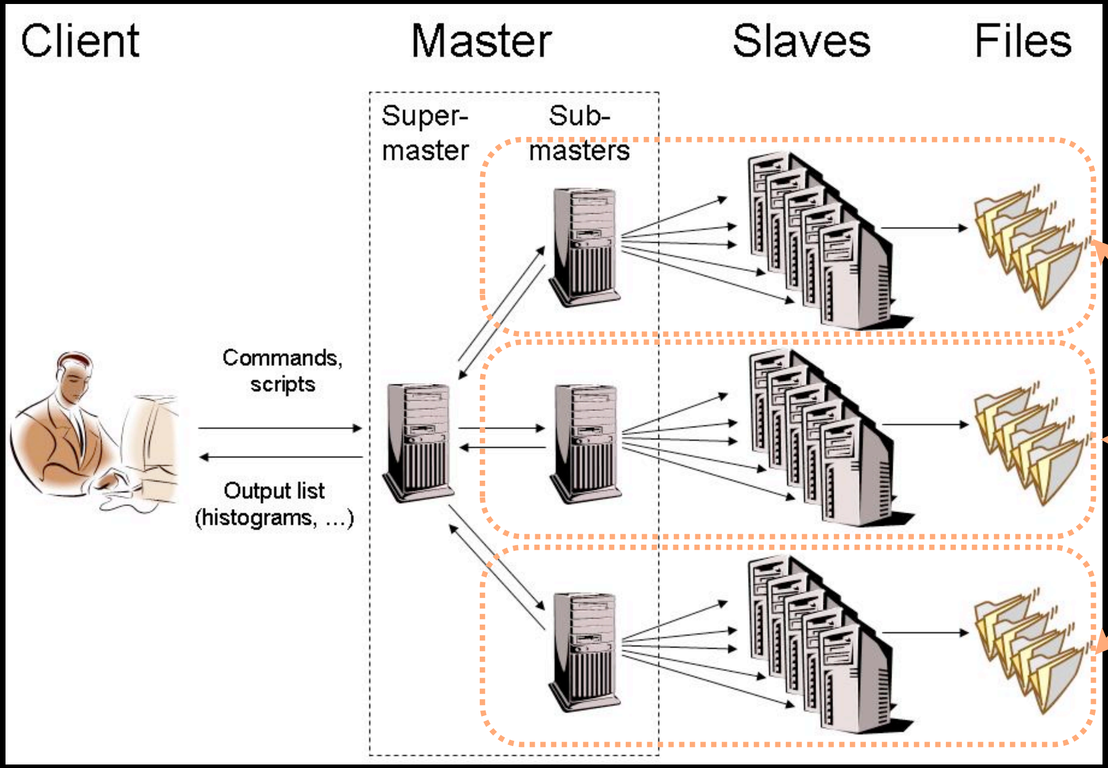
The PROOF Approach





- Cluster perceived as extension of local PC
  - *Same macro and syntax* as in local session
- More *dynamic* use of resources
- Real-time feedback
- Automatic *splitting* and *merging*

Multi-Tier Architecture



Adapts to wide area *virtual* clusters

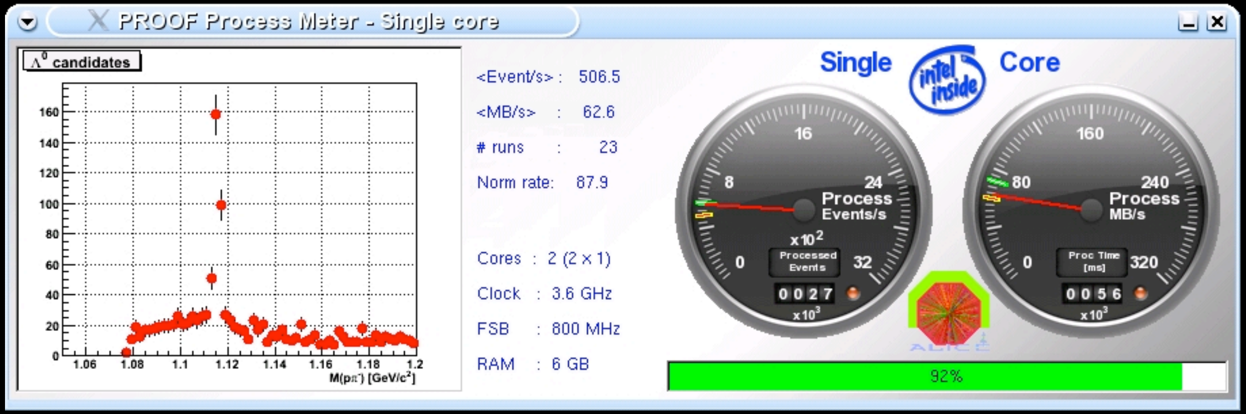
Geographically separated domains, heterogeneous machines

Less important — Network performance → VERY important

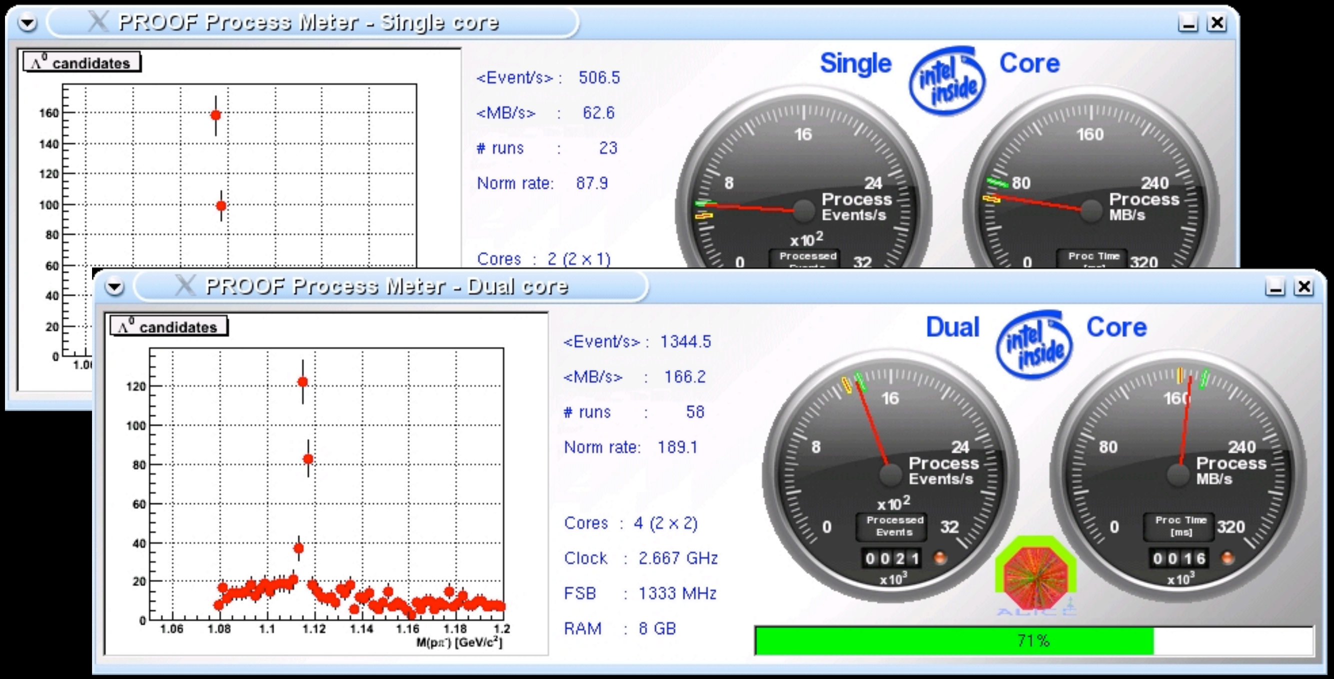
Optimize for data locality or high bandwidth data server access

PROOF Scalability on Multi-Core Machines

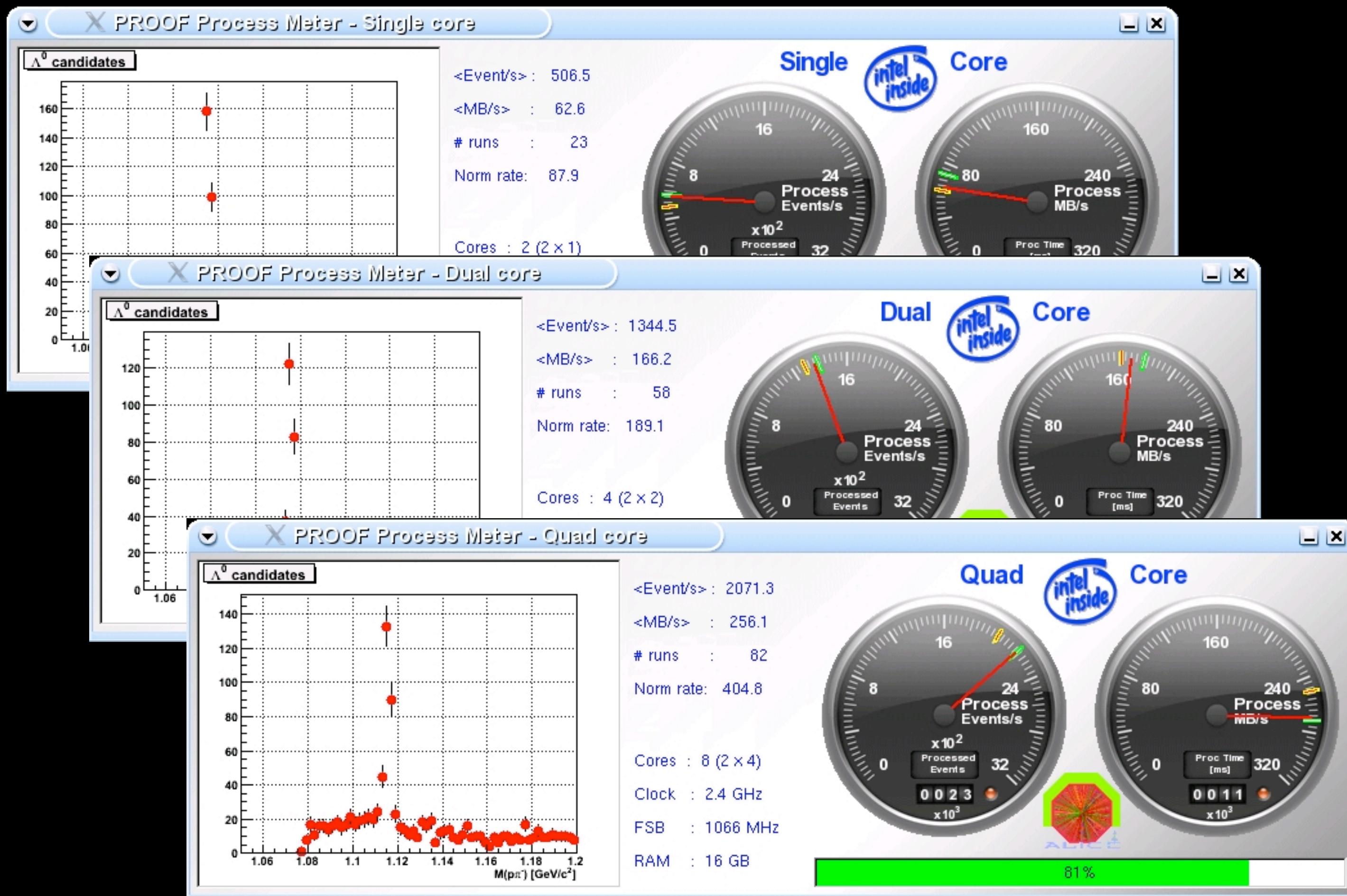
PROOF Scalability on Multi-Core Machines



PROOF Scalability on Multi-Core Machines



PROOF Scalability on Multi-Core Machines



Current version of Mac OS X fully 8 core capable. Running my MacPro since 4 months with dual Quad Core CPU's.

Apple Software and LHC Computing

- All scientific computing is Unix based, Mac OS X a natural and good fit
- Possibility to mix scientific and office software on the same platform is a huge advantage
- Excellent software development tools
  - Xcode
  - gcc, gfortran, Intel's icc and ifort
  - Shark, mallocdebug, etc.
- Intel compilers deliver up to 25% faster executables
- Excellent cluster monitoring software
- Leopard will be even better with Xray and fully 64 bit

Apple Hardware and LHC Computing

- The move to Intel was an enormous step forward in terms of performance
- MacBook Pro’s are now by far the most popular laptops at CERN
- Powerful OpenGL based graphics used for event displays
- ALICE uses an 8 node, 16 CPU, Xserve and Xserve RAID cluster as one of its Grid computing elements
- We are currently upgrading to an Intel based Xserve cluster

Meeting of the LCG Software Architects

Conclusions

- The LHC will generate data on a scale not seen anywhere before
- LHC experiments will critically depend on the Grid to process their enormous amounts of data
- Apple’s move to Intel opens up tremendous possibilities in this market
- Wish us good luck!