CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

ACGRID School

and Symposium

Nov. 5-16, 2007, IoIT (VAST)

Advanced Computing and GRID Technologies for Research

Sponsors

IN2P3: Institute of Nuclear and Particule Physics, a CNRS institute

ICT-ASIA network: French sponsored IT programmme:
Foreign Affairs Ministry, CNRS, INRIA, GET, ...

Computing in research

- No scientific research without computing expertise
  - Experimental design, simulation, experiment construction, data taking, data analysis, model interpretation and theory development

    → **All these activities need computing support**

- New computing tools are complex and need training

- ACGRID stands for:
  - Advanced Computing
    - Software engineering: Languages, CASE, Databases,
    - Artificial Intelligence: Symbolic manipulation, Genetic algorithm,
    - Distributed computing: parallelism, cluster, GRID, BOINC
    - General purpose Packages: ROOT, GEANT4, TAVERNA
  - Grid
    - Hidden computing from distributed resources
    - Analog to the electrical power grid

CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE

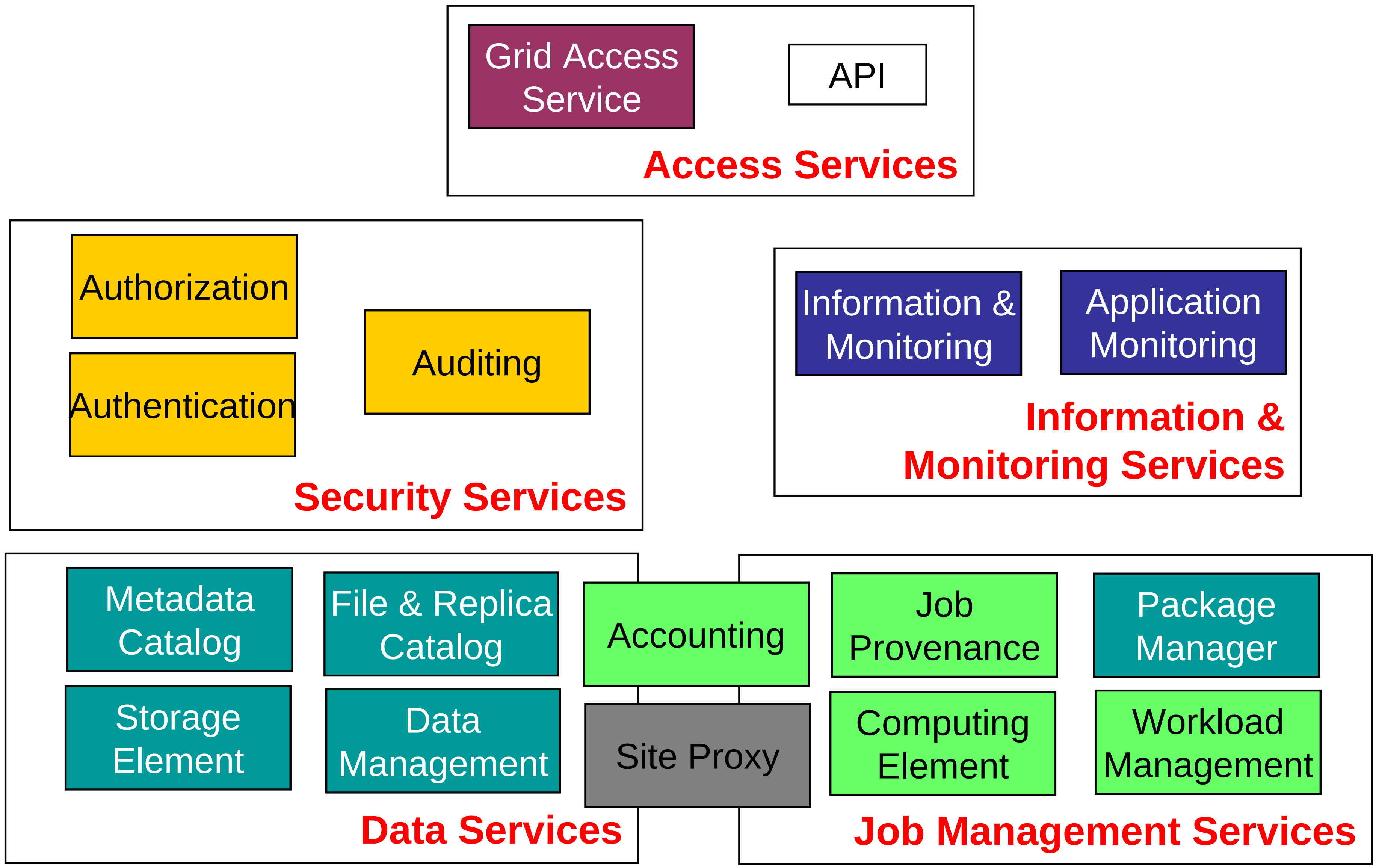
School programme Week I

- All you need to use the GRID
- All you need to GRIDify your application

gLite Middleware

- Vincent Breton: Grids: a new paradigm for science
- Jean Salzemann: Embrace: Integrated system for Bioinformatics
- Matthieu Reichstadt: Bioinformatics portal, AuverGrid
- Vincent Bloch: WISDOM: Wide In Silico Docking On Malaria
- Hung-Chun Lee: AMGA: Access Metadata, GANGA: user interface to the GRID

gLite Services

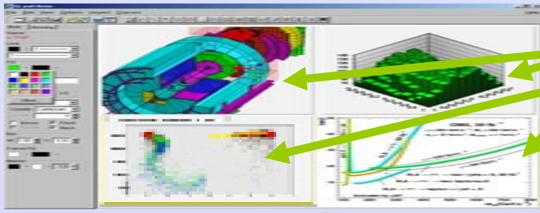


School programme Week II

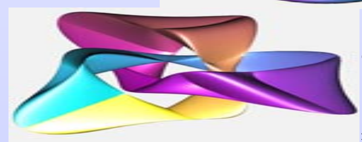
- René BRUN
  - ROOT: Object Oriented Data Analysis Framework
  - physics, astronomy, biology, genetics, finance, insurance, pharmaceuticals, etc.
  - PROOF: Parallel ROOT Facility
- Sébastien Incerti: GEANT4
  - simulation of the passage of particles through matter.
  - high energy, nuclear and accelerator physics, as well as studies in medical and space science
- Georgina Moulton: TAVERNA
  - language and software tools to facilitate easy use of workflow and distributed compute technology within the eScience community
- Nicolas Maire: BOINC **B**erkeley **O**pen **I**nfrastructure for **N**etwork **C**omputing
  - Volunteer computing and desktop grid computing

ROOT Workshop 2007: Graphics News (2D and 3D)

This poster shows some of the new features recently introduced in ROOT 2D and 3D graphics.

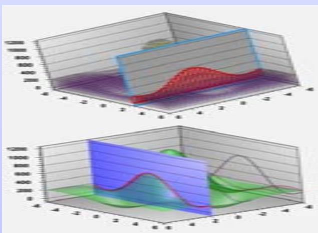


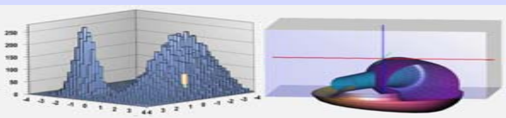
OpenGL 3D graphics can be mixed in a TPad with **standard 2D graphics**. Output can be generated in various formats (postscript, gif, jpeg etc ..).



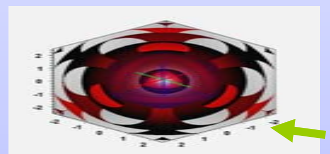
Parametric functions:

```
TGLParametricEquation p1("Conchoid",
                         "1.2 ^ u * (1 + cos(v)) * cos(u)",
                         "1.2 ^ u * (1 + cos(v)) * sin(u)",
                         "1.2 ^ u * sin(v) - 1.5 * 1.2 ^ u",
                         0., 6 * TMath::Pi(),
                         0., TMath::TwoPi());
p1.Draw();
```





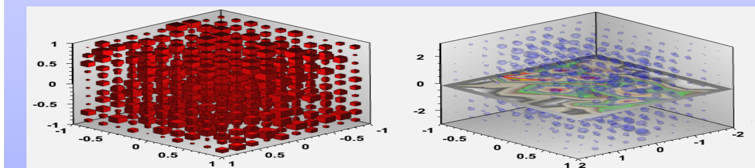
Several interactions are possible on 3D OpenGL plots:
- zooming, panning, rotation,
- bins highlight,
- moving profile planes,
- projection on walls,
- cutting boxes …

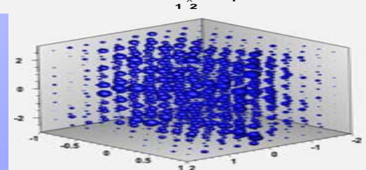


TF3 representations:

```
TF3 fun3("fun3",
         "sin(x*x+y*y+z*z-36)",
         -2,2,-2,2,-2,2);
fun3.Draw();
```

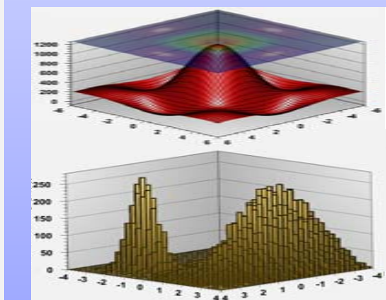
One can interact with the function representation using, for instance, the cutting box.

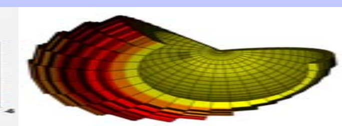


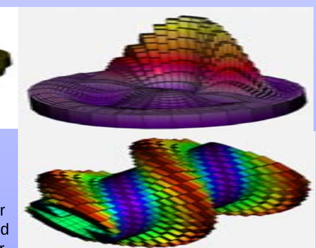


TH3 representations:

3D histograms (`TH3`) can be drawn using different kinds of box plots (cubes or spheres). A 2D contour plot corresponding to a cutting plane can be drawn in real time.

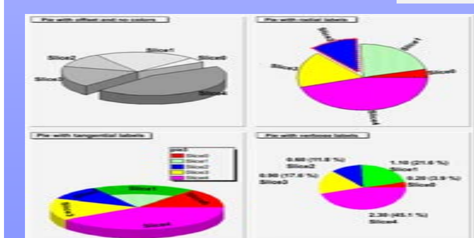






TH2 representations:

Lego and surface plots can be drawn with several options, projection modes and color palettes. Many extensions have been added compared to the “non-GL” way, in particular in the GUI area.

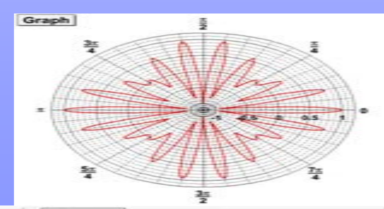


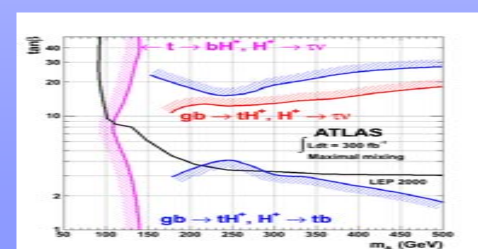
TPie:

The `TPie` class allows to define and draw pie charts. This class offers various options to draw a pie chart (flat, 3D effect, label format etc …) and provides a very flexible and intuitive way to manipulate the drawing interactively. This class can also be used to draw `TH1` histograms.

TGraphPolar:

`TGraphPolar` creates a polar graph (including error bars). A `TGraphPolar` is a `TGraphErrors` represented in polar coordinates. It uses the class `TGraphPolargram` to draw the polar axis.





Exclusion graphs:

A `TGraph` extension allows to draw exclusion graphs. One can choose on which side of the graph the hatches are drawn, the width of the hatched zone, the type of hatches (or patterns) used.

User Interface Classes

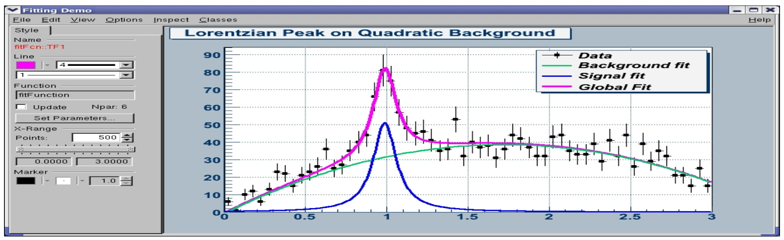
The ROOT GUI classes provide a rich and complete set of widgets allowing the construction of modern looking graphical user interfaces.

Like everything else in ROOT the GUI classes are fully cross platform and provide the same look and feel on either X11, Win32 or Mac OS X.

Complex GUI's can easily be constructed using a GUI builder, which allows widgets to be dragged and dropped into frames.

The GUI and the ROOT graphics classes are fully integrated and it is simple to embed a scientific data display into a GUI .

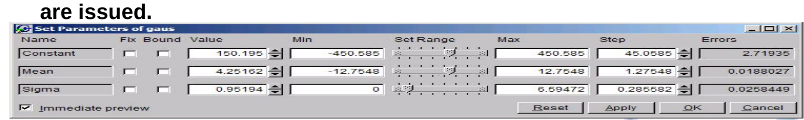
ROOT comes with many examples of high level GUI's like the browser, tree viewer, fit panel, etc.



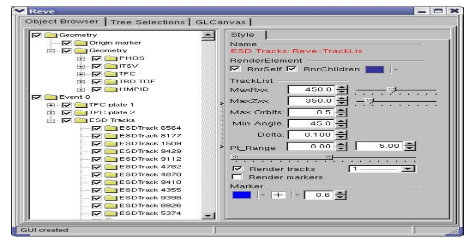
Signals / Slots

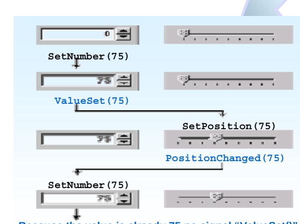
Using the signal/slot communication mechanism GUI elements can be easily connected to any number of action (slot) methods.

Signal/slots are integrated into the ROOT core and heavily use CINT to connect the signals to the slots and to call the slot methods when signals are issued.



On interaction, widgets send out various signals. Any public object method can be connected to these signals.





Fast Prototyping

Like all classes in ROOT the GUI classes are fully scriptable allowing for fast prototyping via the embedded CINT C++ interpreter. In addition any GUI can be saved as C++ macro by typing ctrl-s when the mouse is over a GUI window. As macros can be stored in ROOT files one can envisage to store the GUI with the data:

```
root[] TMacro m("myApplication")
root[] m.ReadFile("myApplication.C")
root[] m.Exec()
root[] TFile f("myFile.root","recreate")
root[] m.Write()
root[] hpxpy.Write()
```

Executing the saved macro restores the complete applicati... Users can ...tor... e

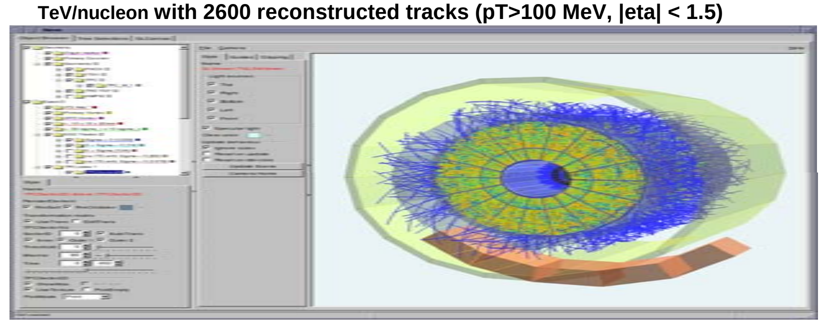
```
root[] TFile f("myFile.root")
root[] f.ls()
TFile**         myFile.root
  TFile*         myFile.root
   KEY: TMacro   myApplication;1
   KEY: TH2F     hpxpy;1 py vs px
root[] TMacro *d = f.Get("myApplication")
root[] d.Exec()
```

Examples

The ALICE Event Visualization Environment (AliEVE) is based on ROOT and its GUI, 2D/3D graphics classes. A small application kernel provides for registration and management of visualization objects. CINT scripts are used as an extensible mechanism for data extraction, selection and processing as well as for steering of event-related tasks.

AliEVE is used for event visualization in offline and high-level trigger frameworks.

The event below is a simulated peripheral lead-lead collision at 5.5 TeV/nucleon with 2600 reconstructed tracks (pT>100 MeV, |eta| < 1.5)



For more information see: http://root.cern.ch
For any questions please use the address: rootdev@pcroot.cern.ch

School programme Week II

- René BRUN
  - ROOT: Object Oriented Data Analysis Framework
  - physics, astronomy, biology, genetics, finance, insurance, pharmaceuticals, etc.
  - PROOF: Parallel ROOT Facility
- Sébastien Incerti: GEANT4
  - simulation of the passage of particles through matter.
  - high energy, nuclear and accelerator physics, as well as studies in medical and space science
- Georgina Moulton: TAVERNA
  - language and software tools to facilitate easy use of workflow and distributed compute technology within the eScience community
- Nicolas Maire: BOINC **B**erkeley **O**pen **I**nfrastructure for **N**etwork **C**omputing
  - Volunteer computing and desktop grid computing

B. Bellenot, R. Brun, G. Ganis, J. Iwaszkiewicz, F. Rademakers, *CERN, Geneva, Switzerland*,
M. Ballintijn, *MIT, Cambridge, MA, USA*

PROOF

PROOF enables interactive analysis with ROOT [1] on distributed computing resources. It realizes basic parallelism by exploiting the independence of uncorrelated events.

PROOF is designed for use at
- Central Analysis Facilities
- Departmental workgroup computing facilities (Tier-2's)
- Multi-core, multi-disks desktops

Design goals

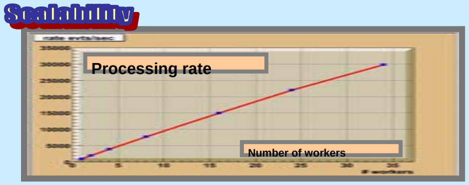
- Transparency: distributed system perceived as an extension of the local ROOT session (same syntax, scripts, …)
- Scalability: efficient use of the available resources: performance scales with the number of CPUs and disks.
- Adaptability: adapt to heterogeneous resources

Features

- Package manager
  - Handling of additional libraries, classes, data files, etc.
- Support for dynamic environment setting
  - On-the-fly definition of variables and/or sourcing of relevant scripts
- Query manager for easy handling of results
  - Results can be saved on any mass storage
- Support for "interactive batch"
  - Smooth interactive -> batch transition
  - Client disconnection / reconnection
    - Users can reconnect later from a different place to e.g. check a long-query status and retrieve the results
  - Background, non-blocking running model
  - Multiple-session control from single ROOT shell
    - Concurrent execution of queries on different sessions
- Dataset manager and uploader

Scalability

- ALICE and PROOF
- Central Analysis Facility (CAF)
- 34 Xeon 2.8 GHz
- 4 GB RAM, GB Ethernet
- Data read from mass storage via XROOTD [2]
- Processing rate up to

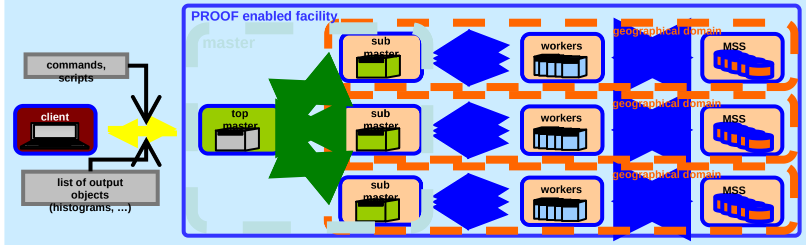


Multicore Scalability

- ALICE analysis repeated *ad infinitum* on dual-socket machines equipped with quad-, dual- and single-core processors
- Speed-o-meters show the instantaneous event and MB processing rates: the advantage of having more CPU is clear
- The rate normalized by the clock speed and # of CPU sockets scales nearly with the # of cores, indicating that the available computing power is fully exploited

[1] http://root.cern.ch
[2] http://www.slac.stanford.edu/xrootd

Multi-tier architecture



Connection lay-out set up via dedicated daemons in charge of authenticating the clients and spawning the server applications. XROOTD [2] has been instrumented for this purpose.

Work distribution and Data Access Strategies

Low-latency access to data is crucial
- Optimizing to process data in it's current location
- Using caching and pre-fetching techniques for data from mass storages

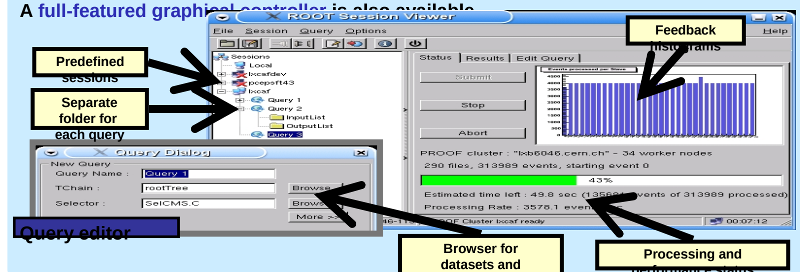
Dynamic load balancing: pull architecture
- Workers ask for more work when they are ready
- Packet generation (*packetizer*):
  - Adaptive mechanism to avoid node overloading

Resource Scheduling

To face the needs of large, multi-user analysis environments expected in the LHC era, optimized sharing of resources among users is required.
The resource scheduling improves the system utilization, insures efficient operation with any number of users and realizes the experiment's scheduling policy. To achieve the goal, two levels of resource scheduling are introduced:
- At worker level, a dedicated mechanism controls the fraction of resources used by each query, according to the user priority and current load.
- At master level, a new central component, the *scheduler,* decides which resources can be used by a given query, based on the overall status of the cluster, the query requirements (data location, estimated time for completion, …), the client history, etc.

Running PROOF: ROOT shell or GUI

A PROOF session is controlled by a dedicated class which can be instantiated on the ROOT shells or on ROOT-enabled applications (see the *grid interface* box for an example).
A full-featured graphical controller is also available.

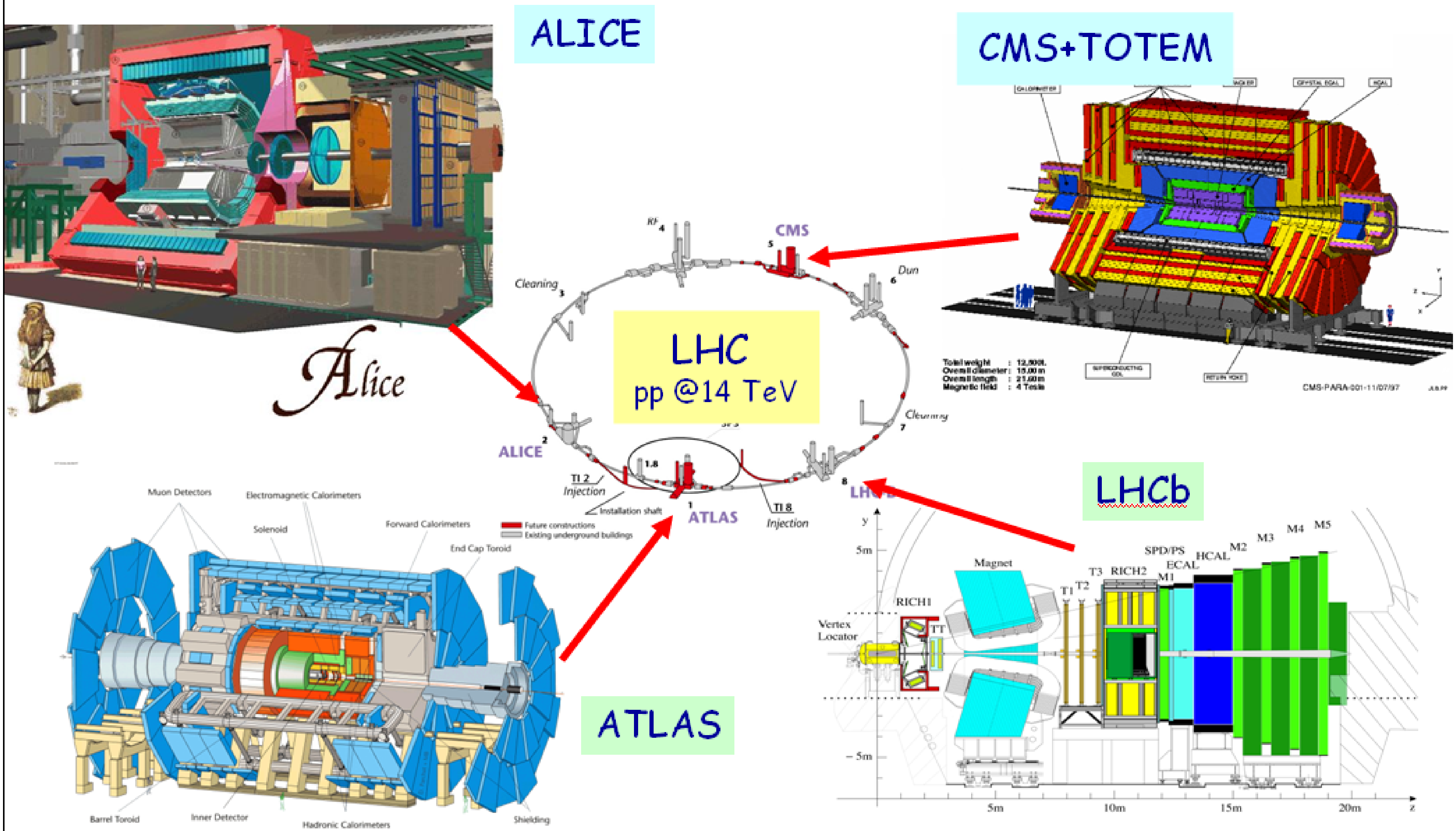


School programme Week II

- René BRUN
  - ROOT: Object Oriented Data Analysis Framework
  - physics, astronomy, biology, genetics, finance, insurance, pharmaceuticals, etc.
  - PROOF: Parallel ROOT Facility
- Sébastien Incerti: GEANT4
  - simulation of the passage of particles through matter.
  - high energy, nuclear and accelerator physics, as well as studies in medical and space science
- Georgina Moulton: TAVERNA
  - language and software tools to facilitate easy use of workflow and distributed compute technology within the eScience community
- Nicolas Maire: BOINC **B**erkeley **O**pen **I**nfrastructure for **N**etwork **C**omputing
  - Volunteer computing and desktop grid computing

Geant 4

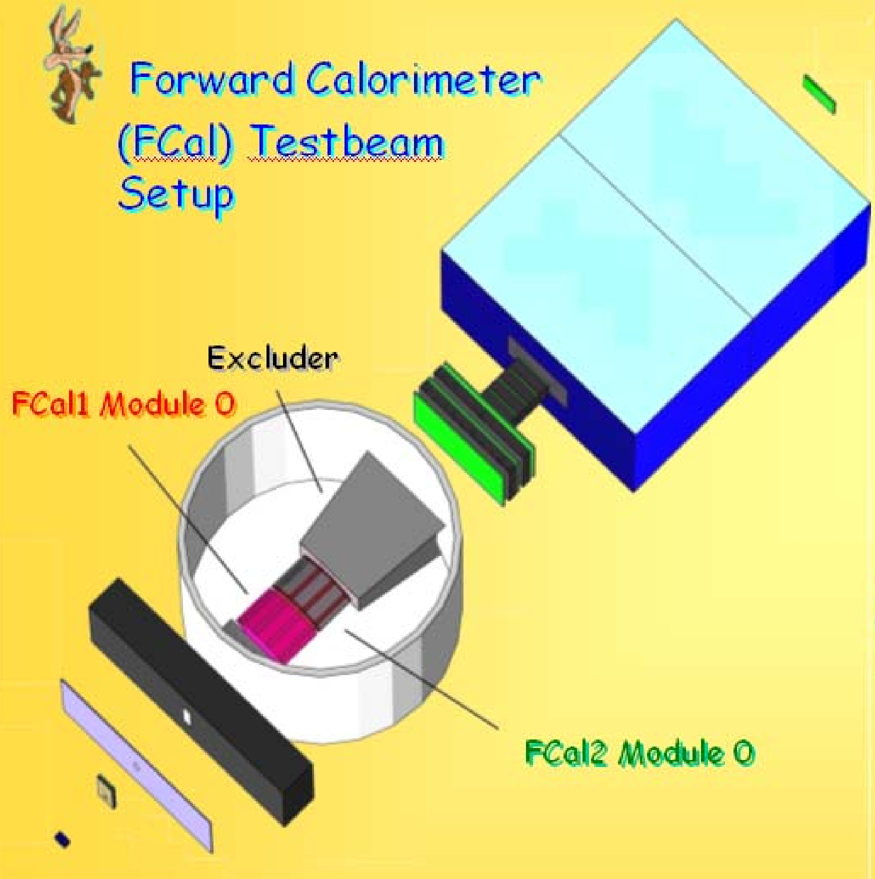
The Large Hadron Collider Experiments



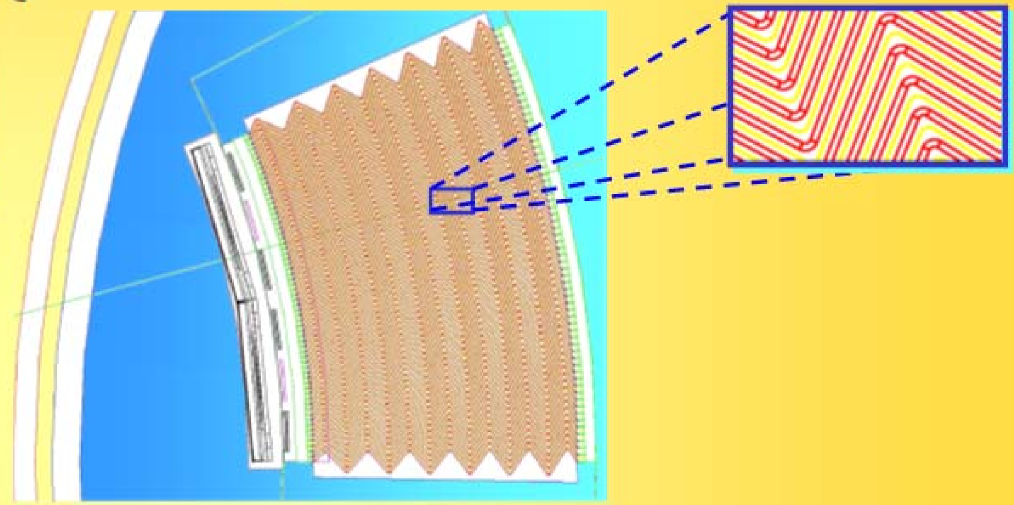
Peter Loch
University of Arizona
Tucson, Arizona 85721

Geant4 Setups (2)

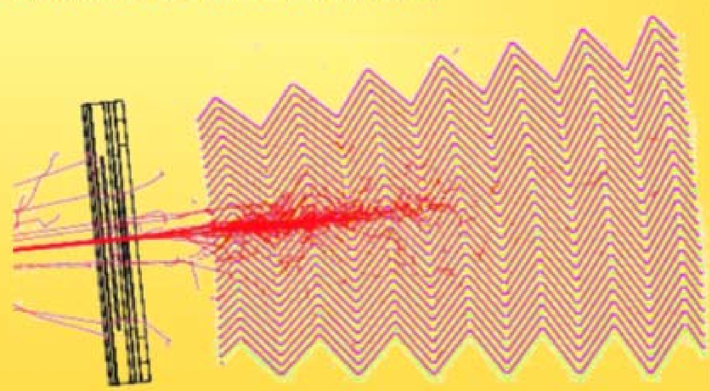
- Forward Calorimeter (FCal) Testbeam Setup



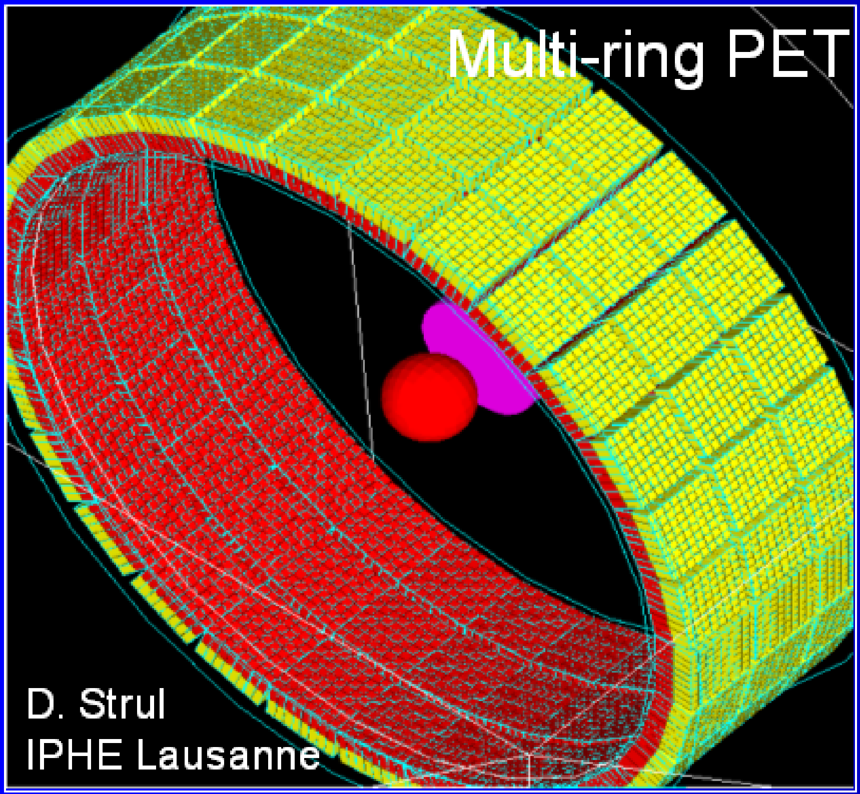
- Electromagnetic Barrel Accordion Calorimeter

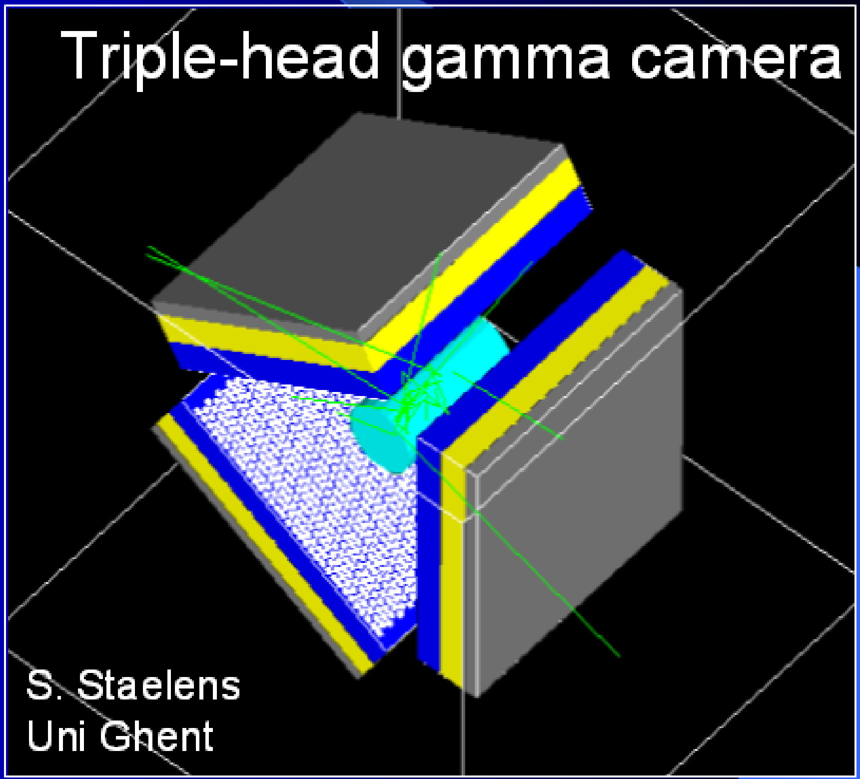


- 10 GeV Electron Shower



Geometry examples of GATE applications





Comparison with commercial treatment planning systems

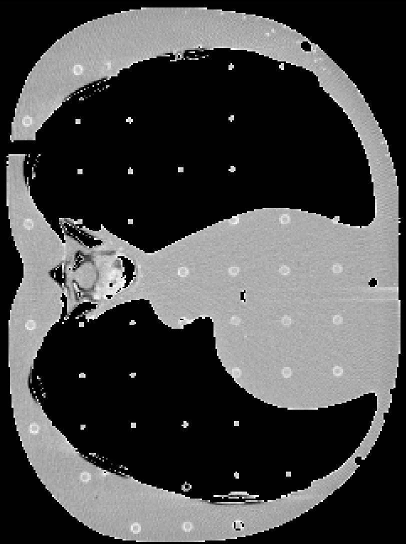
M. C. Lopes [1], L. Peralta [2], P. Rodrigues [2], A. Trindade [2]

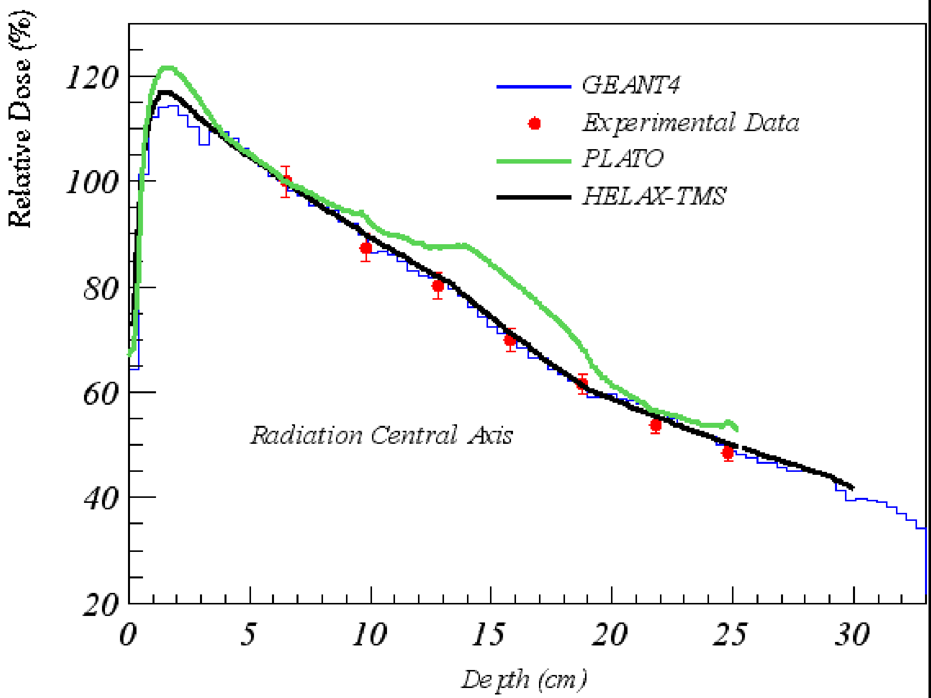
[1] IPOFG-CROC Coimbra Oncological Regional Center - [2] LIP - Lisbon

CT-simulation with a Rando phantom
Experimental data obtained with TLD LiF dosimeter

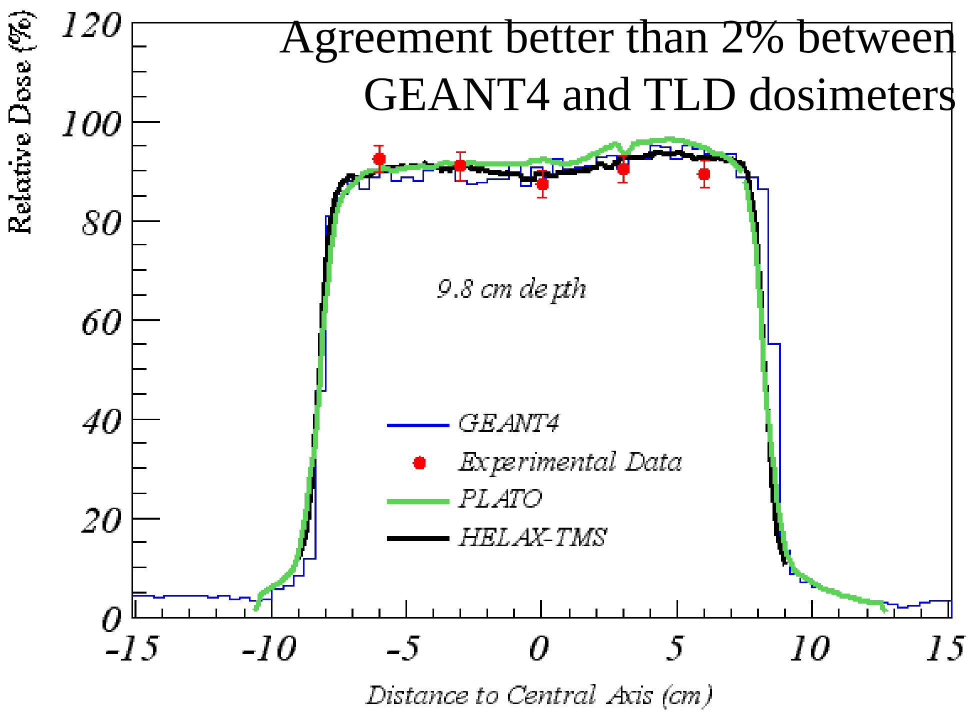
CT images used to define the geometry:

a thorax slice from a Rando anthropomorphic phantom



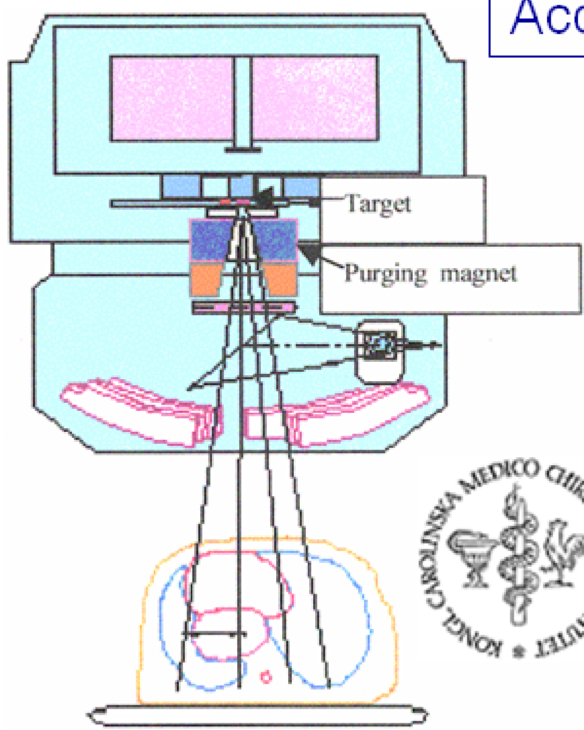


Agreement better than 2% between GEANT4 and TLD dosimeters

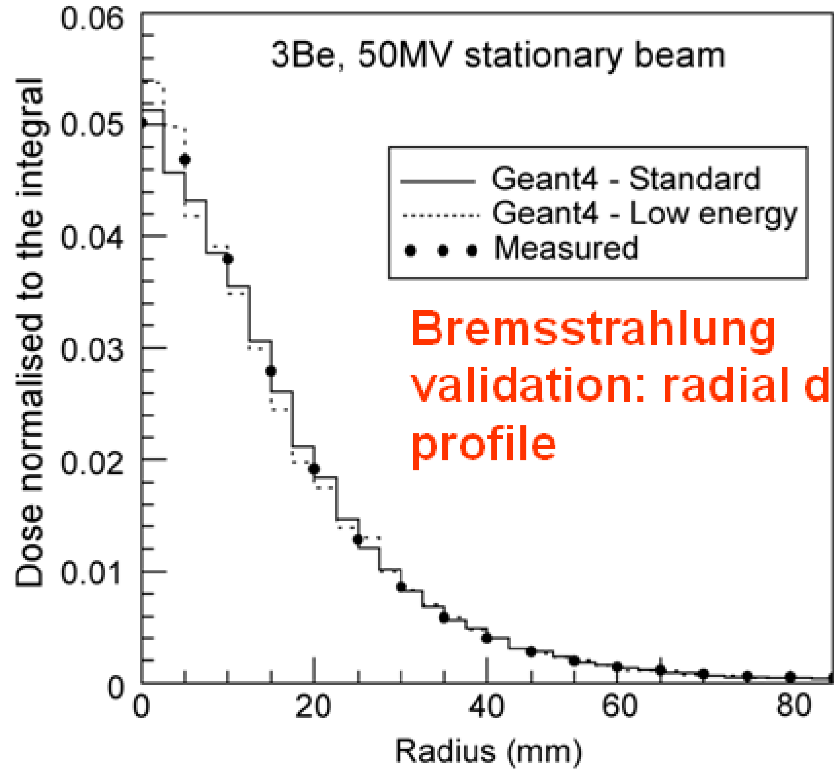


Simulation of a treatment head

Accuracy in the geometry and magnetic field modeling



- High energy electron beam, 50 MeV
- Target 3 mm Be

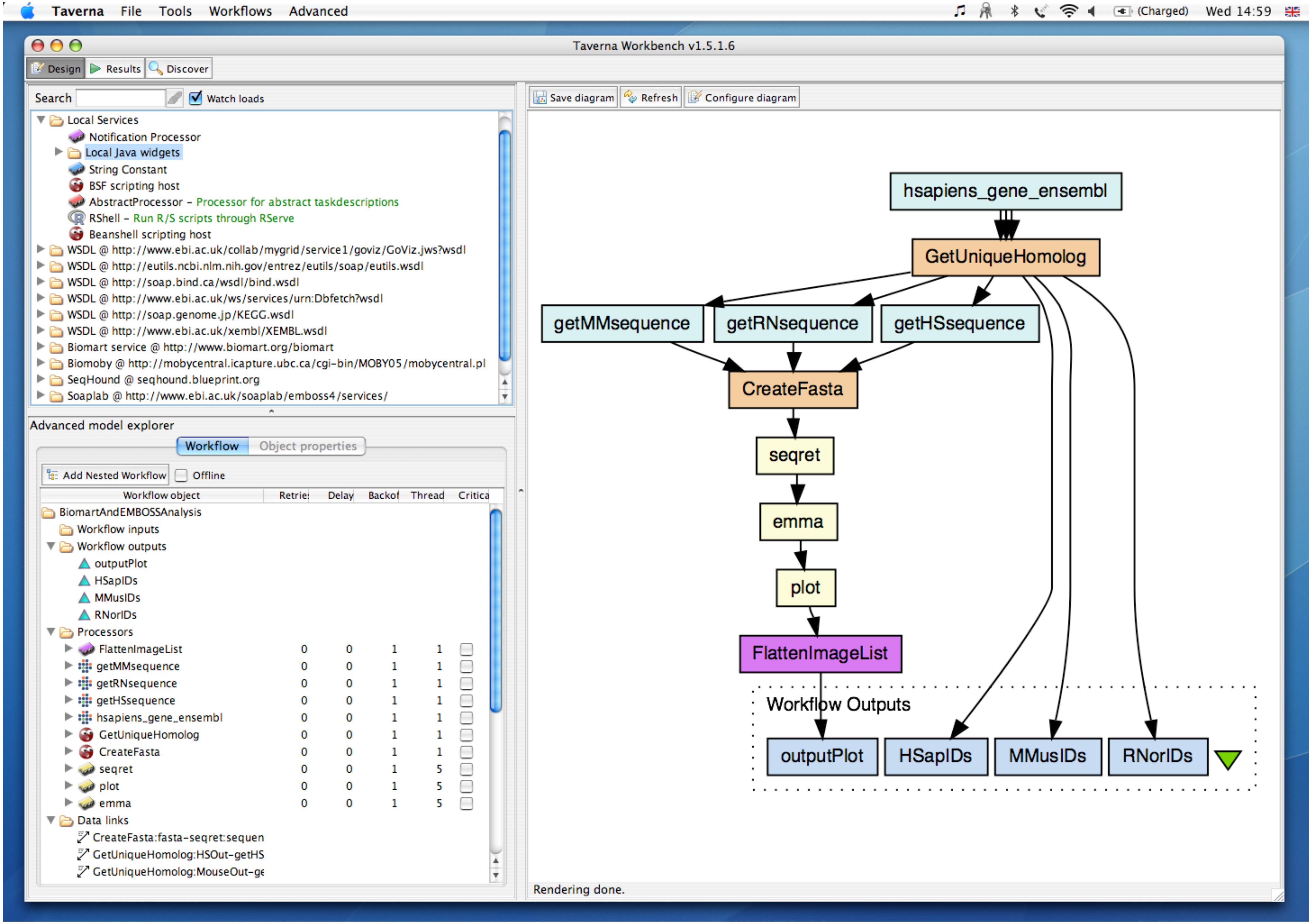


Bremsstrahlung validation: radial dose profile

Susanne Larsson, Roger Svensson
Irena Gudowska, Björn Andreasen (Karolinska Institutet, Stockholm),
Vladimir Ivanchenko (CERN)

School programme Week II

- René BRUN
  - ROOT: Object Oriented Data Analysis Framework
  - physics, astronomy, biology, genetics, finance, insurance, pharmaceuticals, etc.
  - PROOF: Parallel ROOT Facility
- Sébastien Incerti: GEANT4
  - simulation of the passage of particles through matter.
  - high energy, nuclear and accelerator physics, as well as studies in medical and space science
- Georgina Moulton: TAVERNA
  - language and software tools to facilitate easy use of workflow and distributed compute technology within the eScience community
- Nicolas Maire: BOINC **B**erkeley **O**pen **I**nfrastructure for **N**etwork **C**omputing
  - Volunteer computing and desktop grid computing


Taverna File Tools Workflows Advanced
(Charged) Wed 14:59
Taverna Workbench v1.5.1.6
Design Results Discover
Search Watch loads
Local Services
Notification Processor
Local Java widgets
String Constant
BSF scripting host
AbstractProcessor - Processor for abstract taskdescriptions
RShell - Run R/S scripts through RServe
Beanshell scripting host
WSDL @ http://www.ebi.ac.uk/collab/mygrid/service1/goviz/GoViz.jws?wsdl
WSDL @ http://eutils.ncbi.nlm.nih.gov/entrez/eutils/soap/eutils.wsdl
WSDL @ http://soap.bind.ca/wsdl/bind.wsdl
WSDL @ http://www.ebi.ac.uk/ws/services/urn:Dbfetch?wsdl
WSDL @ http://soap.genome.jp/KEGG.wsdl
WSDL @ http://www.ebi.ac.uk/xembl/XEMBL.wsdl
Biomart service @ http://www.biomart.org/biomart
Biomoby @ http://mobycentral.icapture.ubc.ca/cgi-bin/MOBY05/mobycentral.pl
SeqHound @ seqhound.blueprint.org
Soaplab @ http://www.ebi.ac.uk/soaplab/emboss4/services/
Advanced model explorer
Workflow Object properties
Add Nested Workflow Offline
Workflow object Retrie Delay Backof Thread Critica
BiomartAndEMBOSSAnalysis
Workflow inputs
Workflow outputs
outputPlot
HSapIDs
MMusIDs
RNorIDs
Processors
FlattenImageList 0 0 1 1
getMMsequence 0 0 1 1
getRNsequence 0 0 1 1
getHSsequence 0 0 1 1
hsapiens_gene_ensembl 0 0 1 1
GetUniqueHomolog 0 0 1 1
CreateFasta 0 0 1 1
seqret 0 0 1 5
plot 0 0 1 5
emma 0 0 1 5
Data links
CreateFasta:fasta-seqret:sequen
GetUniqueHomolog:HSOut-getHS
GetUniqueHomolog:MouseOut-ge
Save diagram Refresh Configure diagram
hsapiens_gene_ensembl
GetUniqueHomolog
getMMsequence
getRNsequence
getHSsequence
CreateFasta
seqret
emma
plot
FlattenImageList
Workflow Outputs
outputPlot
HSapIDs
MMusIDs
RNorIDs
Rendering done.

School programme Week II

- René BRUN
  - ROOT: Object Oriented Data Analysis Framework
  - physics, astronomy, biology, genetics, finance, insurance, pharmaceuticals, etc.
  - PROOF: Parallel ROOT Facility
- Sébastien Incerti: GEANT4
  - simulation of the passage of particles through matter.
  - high energy, nuclear and accelerator physics, as well as studies in medical and space science
- Georgina Moulton: TAVERNA
  - language and software tools to facilitate easy use of workflow and distributed compute technology within the eScience community
- Nicolas Maire: BOINC **B**erkeley **O**pen **I**nfrastructure for **N**etwork **C**omputing
  - Volunteer computing and desktop grid computing

Jan. 30 2007

1 credit=1/100 cpu PC hour

Statistics
For the WORLD!

Heat Maps of your Pages
...sitors do? Find
...e Trial

39 projects

478,000
CPU Hours/day
~ 20000 CPUs full time

222 M CPU Hours

BOINC Statistics for the WORLD!

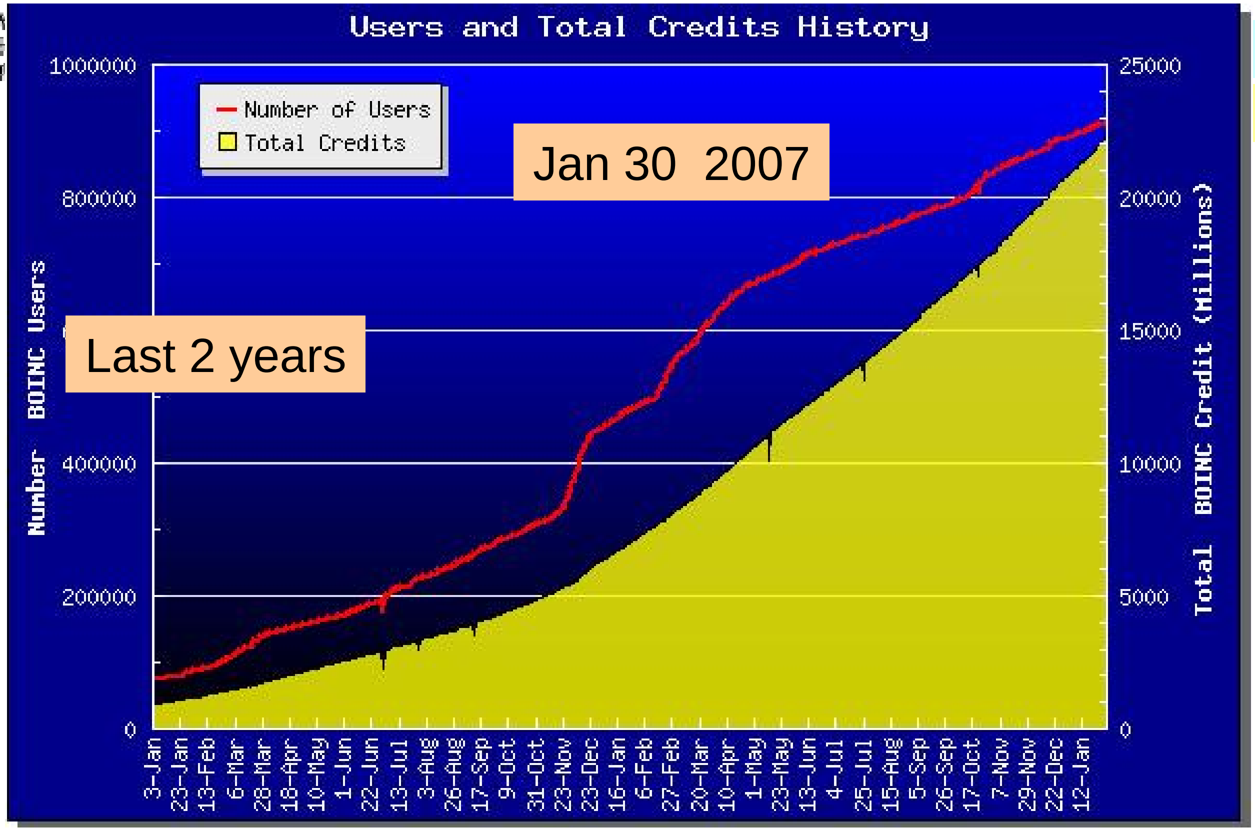
...y BOINC Team - "BOINC Synergy" ("100%" BOINC is ...

...t to leave a comment? Goto our FORUM to post ... the friendly crew of BOINC Synergy.

Project	Development Stage		Total Credit	Recent Credit	Users	Teams	Hosts (Computers)	Countries	Data last upd
Combined BOINC			**22,237,035,221**	**47,807,355**	**914,601**	**57,690**	**N/A**	**250**	-
SETI@Home	-		12,577,720,601	26,244,800	595,177	46,460	1,310,456	245	23.99 hours a
Einstein@Home	-		3,002,622,440	6,634,314	155,311	6,507	329,194	199	4.41 hours ag
ClimatePrediction.Net	-		2,518,092,858	3,615,558	107,421	4,333	199,679	186	14.04 hours a
Rosetta@Home	-		1,298,041,671	3,827,521	108,648	4,486	257,858	193	7.10 hours ag
BBC Climate Change Experiment	-		1,259,204,840	2,113,118	120,243	1,169	136,250	103	14.23 hours a
World Community Grid	-		482,513,459	2,570,984	29,618	3,806	80,347	153	6.71 hours ag
Predictor@Home	-		420,699,910	0	54,792	2,960	131,867	170	5.13 hours ag
QMC@Home	Beta		123,901,047	600,501	12,911	913	24,541	125	7.65 hours ag
LHC@Home	-		111,322,684	0	33,244	1,992	72,360	141	146.94 days a
SIMAP	-		103,670,942	271,985	14,771	1,084	35,947	132	4.64 hours ag
MalariaControl.net	Beta		40,258,960	224,416	3,906	453	11,363	95	12.57 hours a
TANPAKU	Alpha		39,813,670	233,277	5,496	491	12,306	101	17.98 hours a
Seasonal Attribution Project	-		31,569,611	75,052	4,214	360	5,880	82	13.28 hours a
Spinhenge@home	Beta		27,546,204	277,784	10,407	717	16,931	110	8.23 hours ag
SETI@Home Beta	Perm Testing		27,280,277	85,502	3,206	475	7,449	83	4.22 hours ag
SZTAKI Desktop Grid	-		26,820,789	47,482	10,001	829	39,307	111	11.07 hours a
PrimeGrid	Alpha		26,820,355	120,745	3,365	503	13,142	77	6.31 hours ag
uFluids	Alpha		23,791,951	81,736	4,913	599	14,902	88	9.48 hours ag
XtremLab	Alpha		16,963,176	77,431	2,153	361	6,810	75	5.48 hours ag
RieselSieve	Beta		16,788,081	72,719	3,265	342	6,812	84	6.31 hours ag
Leiden Classical	Alpha		13,271,586	59,280	2,468	340	7,135	82	5.22 hours ag
Proteins@home	Beta		9,577,010	208,233	3,414	316	6,068	81	6.57 hours ag
Rectilinear Crossing Number	Beta		9,257,174	78,851	3,517	388	7,860	89	5.15 hours ag

914 K users

222 M hours



Complementary to the GRID

- Large CPU power: 20,000 CPU full time and growing
- BUT
  - •Low reliability: redundant computations
  - •Not for time critical application

Project name Mouse over for details; click to visit web site	Project U Copy and paste into
Biology and Medicine	
SIMAP	http://boinc.bio.wzw.tum.de/boincsimap/
Predictor@home	http://predictor.scripps.edu/
Rosetta@home	http://boinc.bakerlab.org/rosetta/
Malariacontrol.net	http://www.malariacontrol.net
Tanpaku	http://issofty17.is.noda.tus.ac.jp/
World Community Grid	http://www.worldcommunitygrid.org/
Astronomy/Physics/Chemistry	
Spinhenge@home	http://spin.fh-bielefeld.de/
SETI@home	http://setiathome.berkeley.edu/
LHC@home	http://lhcathome.cern.ch/lhcathome/
Leiden Classical	http://boinc.gorlaeus.net/
uFluids@home	http://www.ufluids.net/
Einstein@home	http://einstein.phys.uwm.edu/
Quantum Monte Carlo at Home	http://qah.uni-muenster.de/
Earth Sciences	
Climateprediction.net	http://climateprediction.net
Mathematics and strategy games	
PrimeGrid	http://www.primegrid.com

Symposium

Friday 16 November 2007

Researches using grid

Symposium: Intoduction and Scientific Applications - IoIT (VAST Campus) (16 November 09:00-13:00)

time	[id] title	presenter
09:00	[4] Opening and welcome (00h10')	Prof. MINH, Chau Van
09:10	[67] Institute of Information Technology and IT in Vietnam (00h30')	Prof. THI, Vu Duc Prof. CHI MAI, Luong
09:40	[10] Research and development cooperation (00h20')	FRENCH EMBASSY AND CNRS
10:00	[7] High Energy Physics physics and the GRID (00h30')	Prof. LE DIBERDER, Francois
10:30	break (00h30')	
11:00	[6] The International Linear Collider project (00h30')	Dr. MIYAMOTO, Akiya
11:30	[15] Astroparticles, Space detectors: JEM-EUSO and the GRID (00h30')	Prof. EBISUZAKI, Toshikazu
12:00	[14] Hot issues in the field of emerging diseases (00h30')	Prof. DUNG, Nguyen Tien
12:30	[66] Bioinformatics Grid-based Applications and IOIT-HCM Grid (00h30')	Prof. LANG, Tran Van Mr. LONG, Do Van

Lunch - IoIT (VAST Campus) (13:00-14:00)

Grids in ASIA and Franc

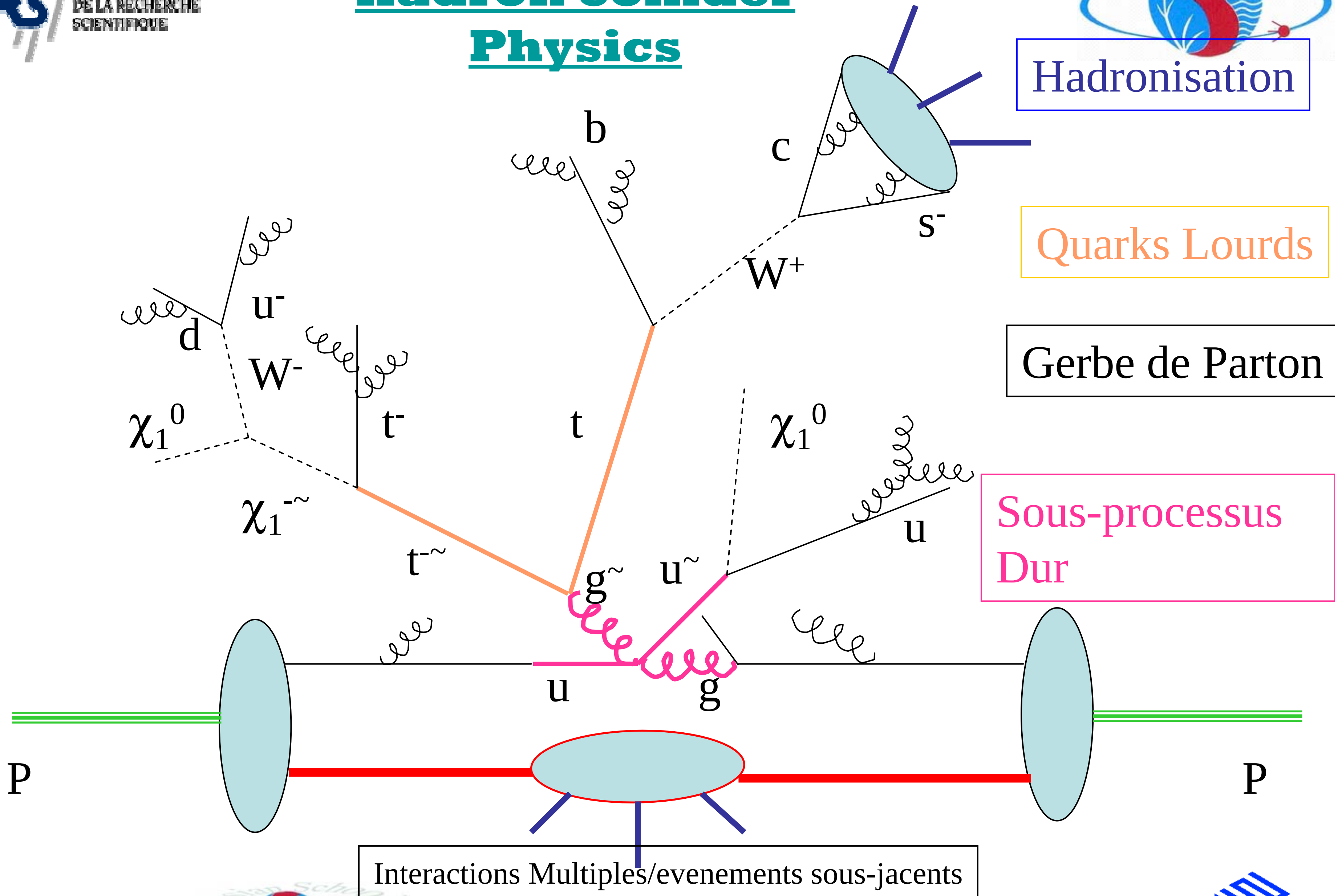
Symposium: the GRID and Conclusions - IoIT (VAST Campus) (16 November 14:00-18:30)

time	[id] title	presenter
14:00	[5] The EGEE GRID in Asia (00h30')	Prof. LIN, Simon
14:30	[9] Grid as a tool for e-science (00h30')	BOUTIGNY, Dominique
15:00	[65] Networking in Vietnam (00h30')	
15:30	[8] VNGRID (00h30')	Prof. LANG, Tran Van Prof. THUY, Nguyen Thanh
16:00	break (00h30')	
16:30	[13] The GRID in Japan (00h30')	Dr. SASAKI, Takashi
17:00	[11] The GRID in China (00h30')	Prof. CHEN, Gang
17:30	[64] Conclusions and perspectives (00h30')	Dr. AURENCHE, Patrick

If you want to know more on:

- Particle Physics
- Automatic Feynman diagram Computations
- BOINC: Feynman@Home Project
- Advanced Computing and Analysis Technologies (ACAT workshop series)
- FJPPL: France-Japan Particle Physics Lab.
- Cooperation Asia-Pacific in Nuclear and Particle Physics

Hadron collider Physics



Bonne chance…

Do not forget about the Nov. 15 Banquet