Journal of Physics: Conference Series **396** (2012) 032033 doi:10.1088/1742-6596/396/3/032033

Validation of Geant4 Releases with distributed resources

Andrea Dotti

CERN , PH-SFT ; J09610 Genève 23 Genève CH-1211 Switzerland

E-mail: andrea.dotti@cern.ch

Abstract. In this paper we present the strategy and tools used to accomplish physics validation of the GEANT4 simulation toolkit.
GEANT4 is a toolkit for the simulation of passage of radiation through matter. It is successfully used by all LHC experiments as well as by medical applications, space physics and industrial applications. A publicly available version of GEANT4 is released once per year and a beta preview release is also made available.
The developed testing suite is extended with a job submission system that allows for the generation of high statistics data-sets on distributed resources. The analysis of the data is also integrated and tools to store and visualize the results are provided.

1. Introduction

GEANT4 is a toolkit for the simulation of the passage of particles through matter [1, 2]. It is used in many applications domains: High Energy Physics, Space Science, Medical Simulations and Technology Transfer [3].

It is the detector simulation engine used by all LHC experiments and these are the largest communities that employ the toolkit. While in the past it has been used to design and optimize the detectors under construction at LHC; GEANT4 is now used to tune the reconstruction and trigger algorithms used in the collision data analysis. It plays an important role in many physics studies both for precise measurement of standard model parameters and new-physics searches enabling the production large Monte Carlo data-samples.

The GEANT4 collaboration releases two new versions of the code each year: a beta preview release just before the summer and a public release at the end of the year[1]. In addition to the public releases monthly development ones (*reference releases*) are produced to allow GEANT4 collaboration members to test new developments and improvements before these are released for public use.

A rigorous approach to validate new developments is used in the GEANT4 collaboration. Each code change is automatically tested every night with incremental builds adding recent developments on top of the most recent development release. This system is focused to test the technical aspects of the code.

In the past the physics validation (regression testing against data or other stable versions of the code) was carried out mainly by the developers of the physics algorithms themselves.

[1] At the time of writing this note the last public release is December 2011 GEANT4 version 9.5

The main focus of this validations is on studying the physics performance of specific aspects of physics modeling. For example: predicted cross-sections are compared to measured ones, final state generators output are validated against alternative models.

Physics models may have a restricted range of validity. This is particularly true for models describing the hadronic interactions with nuclei: their results are optimal on a limited range of energy or primary type. A set of models is assembled in a so-called *physics list* that covers all the simulation needs of a particular use-case.

Thus an additional level of validation is required in which an entire physics list is tested both to verify the stability of the code and the physics results in realistic setups. These second type of validations are time-consuming and in the past years they were performed only for the public releases on limited data sets. An increased interest to strengthen this type of validation has been expressed with the goal to verify the absence of rare bugs and to increase the variety of use-cases being tested. The validation system has thus been extended with the main goal to provide new resources to test GEANT4 applications.

In this paper we will describe the progresses in this area of the last two-three years: a simplified version of a particularly challenging application has been extended to include more detailed analysis of the simulated data, to significantly increase the produced data samples sizes, and to automatically perform the most repetitive aspects of the regression testing procedures.

In Section 2 the general structure of the testing suite is presented together with details on the chosen use-case to test GEANT4 physics. The analysis strategy of the results, their collection and visualization of them are discussed in Section 3. Finally in Section 4 future extensions of the system are discussed.

2. Testing Strategy

One of the most challenging use-cases is the simulation of the LHC experiments. The simulation of calorimeters is particularly demanding: it challenges all aspects of physics simulation (tracking in magnetic field, electromagnetic and hadronic interactions). Many billions of events are simulated each year by LHC experiments and the stability of simulation code has to be guaranteed. In addition, with LHC in a stable data-taking phase, it is particularly important that deviations of the physics results from the expected ones are constantly monitored and justified.

Table 1. Calorimeter types currently implemented in the simulation program.

Experiment	Calorimeter	Absorber	Active Material	Notes
ATLAS/LHCb	Had Calorimeter	Iron	Scintillator	
ATLAS	Barrel E-Cal	Lead	Liquid Argon	Simplified geometry
ATLAS	Had End-Cap	Copper	Liquid Argon	Simplified geometry
ATLAS	Forward	Tungsten	Liquid Argon	Simplified geometry
CMS	E-Cal	-	$PbWO_4$	"Infinite" crystal length
CMS	Had Calorimeter	Brass	Scintillator	
LHCb	E-Cal	Pb	Scintillator	
ZEUS	T60A1	Pb	Scintillator	NIM**A**274 (1989) 134-144
ZEUS	Had Proto	Pb	Scintillator	NIM**A**262 (1987) 229-242
CALICE	A-HCAL	W	Scintillator	

A simplified version of HEP calorimeter has been implemented with Geant4: all LHC calorimeters materials are described. In addition some calorimeters from experiments of the

Journal of Physics: Conference Series **396** (2012) 032033 doi:10.1088/1742-6596/396/3/032033

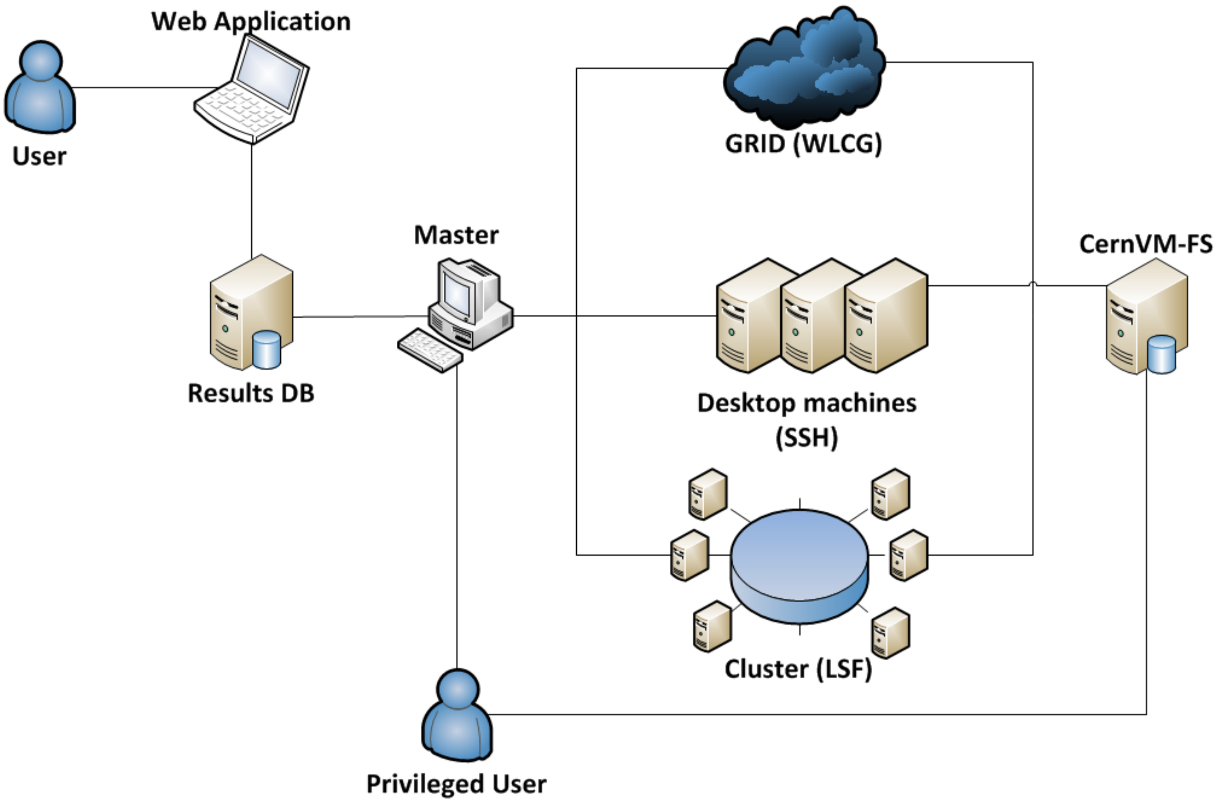

Figure 1. Simplified architecture view of the validation testing suite. The software is distributed on remote sites via CernVM-FS. The validation team publishes pre-compiled binaries for each version of GEANT4 on the CernVM-FS primary server. The client machines will automatically pull the data when needed. The validation team starts the DIANE master that defines the tasks to be executed. Jobs are submitted, via GANGA, to GRID sites, LSF batch queues and linux desktop machines (SSH enabled). Workers connect to the master and download a task configuration and perform the simulation. After completion output files are pushed back to the master node. Results are published in a database and made available to produce plots via a web-application.

past are also implemented for their particular interest, implementation of calorimeters currently in the R&D phase (from CALICE collaboration) are also being included.

2.1. The Simplified Calorimeter application

The simulation program is implemented as a simplified geometry of sampling calorimeters. The calorimeter geometry consists of a cylinder with a radius of $5\times\lambda_I$ (interaction lengths of the absorber material) and $10\times\lambda_I$ length. These large dimensions ensure that the leakage is kept as small as possible and the full hadronic shower is absorbed. The axis of the cylinder is parallel to the beam one.

For sampling calorimeters the layers of active materials are perpendicular to beam axis, the geometrical sampling fraction is as close as possible to the real one. The read-out of the deposited energy is obtained measuring the energy released in the active material. The simplified version of the CMS electromagnetic calorimeter (a lead tungstate crystal calorimeter) is composed of active material only.

The read-out allows to measure the energy deposited in the sensitive volumes and, grouping together read-out planes, to measure the longitudinal shower profile. Concentric cylinders are also defined to read-out the transverse shower profile.

To describe the showers profile in more details an additional three-dimensional read-out mesh is defined. The granularity of this mesh is similar to the one of R&D imaging calorimeters: the deposited energy is collected in "cells" of volume $5 \times 5 \times 5$ cm^3. This fine granularity allows for measuring detailed properties of the showers. To efficiently summarize the physics information (shower shapes) appropriate statistics have been developed (details on the *shower moments* analysis can be found in [4]).

A real calorimeter has to deal with electronic noise, dead cells, cracks and other experimental effect that may play an important role on the measurements. We have not implemented any of these instrumental effects with the exception of the Birks attenuation effect and the read-out timing window (each electronic read-out is sensitive only for a limited time after the data acquisition has been triggered). Slow charged particles highly ionize the active medium and the signal measured, particularly for scintillators, may "saturate" introducing a small non-linearity: this process is described by Birks law [5]. A similar effect has also been observed for liquid-argon based calorimeters [6]. When available the coefficients measured by the experiment collaborations are used, in the other cases the original Birks coefficients are used.

Table 1 show the different calorimeter types currently available in the simulation program: the parameters (dimensions, read-out segmentation, Birks coefficients) for these cases are pre-configured, but any similar setup with cylindrical geometry can be defined at run time.

The total energy deposit, the longitudinal and transverse profiles are the most simple observables measured in the simulation program. In addition energy spectra of all secondary species created during the shower development are also registered. This information is not available in real-setups but it is very useful to study the details of the simulation process [10, 9].

An additional component is optionally available to further increase the level of detail of the analysis. The hadronic component of showers is analyzed and the *hadronic shower tree* is reconstructed. Each new particle is analyzed in terms of the parent track generating it. For each event summary information are then created: number of hadronic interactions of a given type with information on the generating primaries, number, species and average energy of secondaries. This type of analysis is time consuming and produces very large outputs thus it is disabled by default.

The described physics quantities are recorded in ROOT [7] files in the form of histograms and *trees*. In addition all warning and error messages produced during the simulation are recorded together with physics summaries in form of text in log files.

The Simplified Calorimeter application is also used by the Geant4 collaboration to perform CPU and memory performance measurements. For this particular use the entire physics analysis code is disabled.

2.2. Testing suite architecture

The Simplified Calorimeter simulation program is run via a driver scripts that is responsible of setting up the environment with the correct version of the needed external dependencies (ROOT and Geant4) and defining the correct configuration.

The simulation of full showers and their analysis is a time consuming task, especially for high-energy primaries. CPU time scales roughly linearly with primary energy and is of the order of few seconds per event at energies of the order of 10 GeV, but each single event of few hundreds GeV of initial energy can take several minutes to be simulated. To obtain high enough statistics 5000 events are needed for each run. A run is a single simulation job with a defined calorimeter type and a defined primary type and energy and a defined physics list.

It is important to test each version of Geant4 at least on the most common materials to validate new developments. Useful physics results are obtained with a scan of the energy of the primary (from 1 GeV up to 500 GeV) to study the dependence of observables on the beam energy. Even if LHC experiments are converging on 3 main hadronic models to cover their simulation needs (namely the Fritiof model at high energy, the Bertini intra-nuclear cascade at intermediate energies and the Pre-compound model to de-excite the wounded nucleus) we test 10 different physics lists to study the physics performance also of less used models [2]. In total about 10 millions events are produced to cover all necessary combinations.

This set of simulation configurations is adequate to validate the monthly reference releases, while for the public releases data samples of at least a factor of 10 larger are produced. Distributed GRID resources are used to produce the needed statistics in a reasonable time.

The Geant4 virtual organization (*VO*) makes use of WLCG resources and the Simplified Calorimeter application has been extended to a fully fledged GRID-enabled application. Compared to other use-cases of other VOs, the validation of Geant4 has some peculiarities:

- It does not need large input data sets. Actually no input is needed at all, if we exclude very small configuration files
- Large productions are essentially concentrated in few weeks during the public releases preparation phases. During these periods several versions of the Geant4 code are tested: new updates of the code have to be verified on a almost a daily timescale
- The produced output is relatively small, of the order of 10 GB, and needs to be analyzed only once, summaries should be centrally collected and kept for reference. The summaries are very small, of the order of 10 MB

The general architecture of the testing suite is shown in Figure 1. In Figure 2 additional details on the testing suite deployment are shown. The simulation executable is run on remote worker nodes.

To steer, communicate, monitor the remote application and transfer the data between the master and the worker nodes a DIANE application [12] has been prepared. The DIANE Geant4 application receives as input the list of configuration parameters: Geant4 version, calorimeter types, primary species and energies. A queue of *tasks* is created (one task corresponds to one simulation run: a calorimeter, a beam type and energy and a physics list). A DIANE *master* is started and waits for workers to connect.

DIANE *workers* communicate with the master via CORBA [13] and pull the configuration for the simulation corresponding to a given task. To efficiently use remote resources the workers can execute one or more simulation runs until all the available CPU time is consumed. The diane worker is thus a pilot job that *pulls* new tasks. The simulation output (ROOT and log files) is *pushed* back to the master node for permanent storage and further analysis.

The DIANE master has responsibility for accomplishing three other important tasks:

(i) In a preparatory stage it splits the most time consuming tasks (the one corresponding to higher energy) in smaller sub-tasks
(ii) When output is received from workers it merges the ROOT files from sub-tasks
(iii) It then analyzes the merged output to produce and store summaries

Summaries are labelled depending on the conditions on which they were produced, the most important one being the Geant4 version. Summaries are stored in a MySQL database. In this way it is possible to query the database and compare results corresponding to different conditions.

[2] A description of Geant4 physics is out of the scope of this paper, additional information can be found in [11] and the Geant4 physics manual [8]

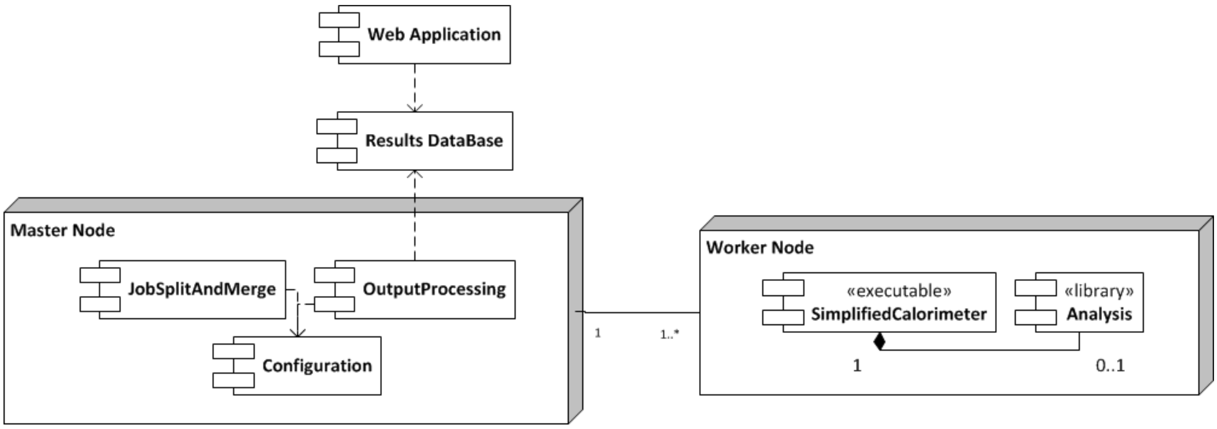

Figure 2. UML deployment diagram of the testing suite. The simulation task is run on worker nodes. The workers are controlled by the DIANE master node, responsible to steer the workers, collect produced output, summarize results and fill the results database. The relevant information is accessed with dedicated scripts or via a DRUPAL web-application.

Several helper scripts have been created to extract relevant information from the results database and to perform the most important actions: parse log files in search of unexpected warning or error messages, create physics summaries with the most relevant plots, compare different versions of Geant4 for the most important physics observables. In particular this last task can be achieved also via a DRUPAL web-application [14].

Since the code changes relatively often, especially during release phases, we have searched an efficient way to distribute the Geant4 libraries to remote nodes. The Cern Virtual Machine File System (*CernVM-FS*) project fully satisfies our needs [15]. CernVM-FS is a FUSE distributed file-system based on the HTTP protocol. The Geant4 VO manager installs the needed software in a central server. Remote sites that support CernVM-FS get the update after few hours and thanks to local caches the network requirements are significantly reduced. Since CernVM-FS uses the well known FUSE technology it can be installed on a variety of systems. In the case of Geant4 validation it allows to use transparently GRID, local clusters or batch systems without the need to adapt the application layer (see Figure 1).

To submit workers to remote sites GANGA [16] is being used. Once the GANGA job is scheduled for execution on the remote site it starts the DIANE worker and simulation tasks are executed. The use of GANGA allows for a transparent submission to different backends. Submitter scripts for GRID, LSF based batch systems and SSH-enabled local machines are available.

3. Analysis of Results

The validation of any new Geant4 release is a two-steps process. At first the more technical aspects of the execution are verified: in particular possible crashes have to be investigated and reproduced. Each job is identified by a unique random seed that is used to re-run a failing job and perform a full debug. In addition the output and error streams of the simulation application are searched for unexpected warning or error messages.

3.1. Summaries production and results database

The physics observables are stored in ROOT files in the form of histograms or TTrees. To further reduce their size and summarize them in an efficient way statistics are stored in a MySQL

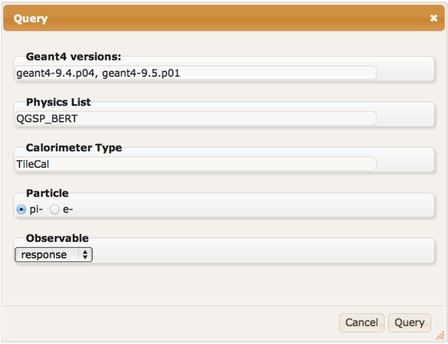

Figure 3. Query panel of the web-application to produce the comparison plots. After selecting the calorimeter type, the beam particle and the observable to be shown two types of plots can be created: for a given physics list it is possible to select many GEANT4 versions or the comparison is made between physics lists for a given GEANT4 version.

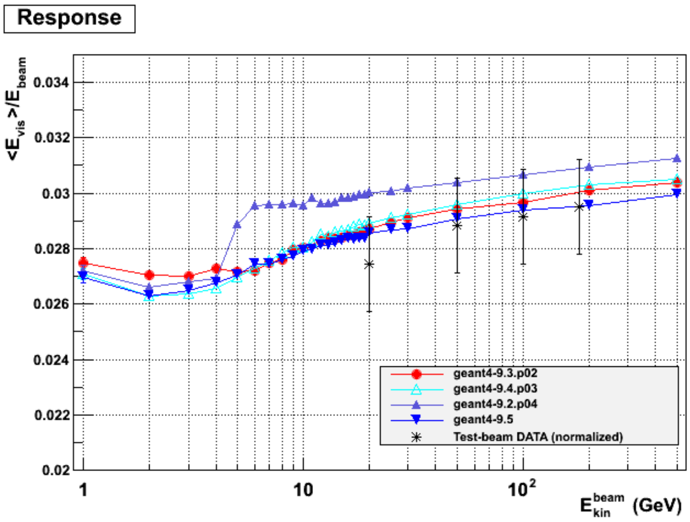

Figure 4. Example of a comparison plot obtained with the analysis macros and web-application. The colored lines represents the simulation results obtained for one physics observable: in this case the visible energy (i.e. the energy in the active material) for a particular calorimeter type (simplified TileCal) for a pion beam as a function of beam energy. The simulations are obtained with the FTFP_BERT physics list. For reference the test-beam data results, published by the TileCal collaboration are also shown (markers).

database. For each distribution the mean, variance, kurtosis, skewness and percentiles values are calculated and stored with references to the simulation setup (GEANT4 version, primary type, energy, etc).

The analysis of the data is done by pre-defined python macros that perform the most common studies. A set of python modules is also available to ease the creation of additional analysis. In particular the modules provide the interfaces needed to upload and get data from the results database and they provide helpers to produce histograms and graphs with a common look-and-feel (for example using always the same colors and markers for a given physics lists). The analysis modules also contain references to the experimental test-beam data if these are available.

Figure 4 shows an example of the output of one of the analysis macro: the visible energy of a pion beam as a function of the beam momentum for different public releases of GEANT4 simulations (FTFP_BERT physics list) are compared against the test-beam data of the TileCal collaboration.

3.2. Web application

To further streamline the production of relevant results a simple web-application has been developed: it runs automatically the most relevant analysis macros[3]. A DRUPAL module has been developed that allows to setup the query to the results database and runs the analysis macro (see Figure 3). The plots are produced and presented to the user (see Figure 4). A sub-set

[3] For the time being only the validation team has access to the web-application, further improvements are needed (in particular simplifying the query panel) before the application can be open to public use.

of the most relevant plots are also copied to the central GEANT4 validation database that collects the most important physics results being advertised to the GEANT4 users community [17].

4. Conclusions and future directions

In the last years a testing suite to validate GEANT4 has been developed; the main focus being the validation of hadronic interactions. Simplified but realistic versions of calorimeters from high energy physics experiments have been implemented in GEANT4 . Analysis of simulated data has been integrated and the system has been embedded in a testing suite that enables running on distributed resources. Results are collected, further analyzed and summaries are stored in a database. Regression testing plots can be produced via macros or a simple web-application.

The described GEANT4 validation system has been developed and commissioned during the period 2009-2011. GEANT4 version 9.5.beta (June 2011 release) was the first one to be validated with this system. During the 3 weeks of the validation phase more than 90 million events are produced on about 1000 GRID nodes at CERN, France, England, Japan and the Netherlands. More than a thousand simultaneous GANGA jobs have been used with success to produce the data. In critical phases of the release phase, additional workers can be added to the workers pool using local CERN batch queues (up to a maximum of 300 workers) demonstrating the flexibility of the system in allocating resources.

The system in its current state has met the requirements and allows for a detailed validation of GEANT4 . However some components will be improved. Simplifications in the submission mechanism are foreseen to ease the job steering and run-configuration definition. DRUPAL has shown very good flexibility to develop the web-application to present the results, this benefit comes at a price of some complexity in developing even relatively simple custom modules. For this reason the possibility to use the MCPLOT [18] system (developed for the theory group at CERN) to present the results in web-pages will be investigated.

With an increased streamlining of the production and monitoring system it will be possible to open the validation procedures to an increased number of users, establishing a *shift rotation* mechanism among collaboration members and share the duties of physics validation. While the system was developed with LHC experiments in mind, it can be used to validate GEANT4 with other use-cases. The system could be adapted to different use cases that need to run large GEANT4 simulations: other users may benefit of components of the system or even integrate their application in the same system.

Acknowledgments

The author would like to acknowledge the help of J.T. Mościcki for his support with DIANE, V.D. Gonzales for his support with the MySQL databases and J. Blomer and P. Buncic for their help in setting up CernVM-FS for GEANT4 .

References

[1] S. Agostinelli *et al*, Nuclear Instruments and Methods in Physics Research **A** 506 (2003), 250-303
[2] J. Allison *et al*, IEEE Transactions on Nuclear Science 53 No. 1 (2006), 270-278
[3] GEANT4 User domains webpage http://geant4.cern.ch/applications/index.shtml
[4] A. Dotti, Simplified Calorimeter: shower moments, CERN-LCGAPP-2012-01
[5] R.L. Craun, D.L. Smith, Nuclear Instruments and Methods 80 (1970), 239-244
[6] S. Amorusoa *et al*, Nuclear Instruments and Methods in Physics Research **A** 523 (2004) 275-286
[7] I. Antcheva *et al*, ROOT A C++ framework for petabyte data storage, statistical analysis and visualization, Computer Physics Communications, Volume 180, Issue 12, December 2009, Pages 2499-2512
[8] GEANT4 Collaboration, *Physics Reference Manual*, available online from: http://www.geant4.org
[9] A. Dotti *et al*, Nuclear Science Symposium Conference Record (NSS/MIC), 2011 IEEE (2011), 2128-2134
[10] A. Dotti *et al*, *Recent improvements on the description of hadronic interactions in* GEANT4 , Journal of Physics: Conference Series Vol. 293 (1011), 012022
[11] A. Ribon *et al*, Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE (2009), 526-529

[12] J.T. Mościcki, Understanding and Mastering Dynamics in Computing Grids Processing Moldable Tasks with User-Level Overlay, Available at: http://dare.uva.nl/en/record/373814
[13] CORBA webpage, http://www.corba.org/
[14] DRUPAL support at CERN, http://www.cern.ch/drupal
[15] J. Blomer *et al*, Proceedings of the first international workshop on Network-aware data management NDM '11, 2011, 49-56
[16] J.T. Mościcki *et al*, Ganga: a tool for computational-task management and easy access to Grid resources; Computer Physics Communications, Volume 180, Issue 11, (2009)
[17] J. Jarba *et al*, Recent Developments in Geant4 Hadronics, Published in this conference proceeding.
[18] MCPLOTS webpage, http://mcplots.cern.ch