

# Geant4 Hadron Validation Framework Requirements

Spring 2009

Written by:

S. Banerjee	FNAL/CD/
S.Gysin (Editor)	FNAL/CD/CET
J.Yarba	FNAL/CD/

Checked by:

J.Apostolakis	CERN/PH
G.Folger	CERN/PH
D.Elviria	FNAL/CD
C.Green	FNAL/CD
V.Ivantchenko	CERN/PH
A.Ribon	CERN/PH
D.Wright	SLAC

Copied to :

J.Kowalkowski	FNAL/CD
P.Spentzouris	FNAL/CD

Version	Date	Comments
2.1	7/1/09	Revision for approval, with updates from review 6/19
2.0	6/10/09	Incorporated comments
1.0	3/5/09	For external review
0.1	3/2/09	First draft for authors

## Table of Contents

<a href="#">1 Introduction.....</a>	<a href="#">3</a>
<a href="#">2 Terminology.....</a>	<a href="#">4</a>
<a href="#">3 Workflow for Testing a Release.....</a>	<a href="#">4</a>
<a href="#">4 Selecting Tests and Test Groups.....</a>	<a href="#">5</a>
<a href="#">5 Executing Tests.....</a>	<a href="#">5</a>
<a href="#">6 Comparison .....</a>	<a href="#">7</a>
<a href="#">7 Storing Results.....</a>	<a href="#">8</a>
<a href="#">8 Publishing Comparisons.....</a>	<a href="#">8</a>
<a href="#">9 Authentication Requirements.....</a>	<a href="#">9</a>
<a href="#">10 Deployment and Operations .....</a>	<a href="#">10</a>

# GEANT4 Hadron Validation Framework Requirements

---

<a href="#">11 Issues.....</a>	<a href="#">11</a>
<a href="#">12 References.....</a>	<a href="#">11</a>

## 1 Introduction

Geant4 [1] is a software toolkit for the simulation of the passage of particles through matter. It is used by a large number of experiments over many application domains, high energy and nuclear physics, astrophysics and space science, medical physics and radiation protection. Geant4 describes physics of all different particles over a wide energy region: from optical photon and thermal neutron to particles with energy of several TeV.

Geant4 divides the physics process into several categories: electromagnetic, hadronic, weak interactions, and optical. The electromagnetic physics processes include ionization, bremsstrahlung, multiple scattering, Compton and Raleigh scattering, pair production, photo-electric effect among others. All these processes are theoretically well understood and the descriptions match well with the existing measurements. The physics of hadrons is not as well-understood from first principles and thus must be described in terms of one or more models. The main challenge is to provide a good description of the physics over a range of energy spanning 15 orders of magnitude. Several models are required to do this, even though a typical user may use only one or two. These models are either driven by data, motivated by theory, or based on parameterizations and extrapolation of cross sections. Each of these models is valid for a restricted number of incident particles and over a restricted energy region.

It is essential to find out the range of application of these models by examining them against available data. Also the models are periodically improved by injecting new ideas. Validation of a physics model is an integral part of commissioning the model in the Geant4 application and has been a part of Geant4 activity from the very early days. This work has been done within the Geant4 collaboration using published data and also by users with complete detector setup.

These requirements cover two types of tests for hadronic model validation:

- Regression Tests, which are comparison of physics output from release to release.
- Validation Tests, which are comparison of physics output with published data.

Tests such as unit tests and system tests are expected to be performed outside of the frame work.

The metric of the comparisons will be developed over time. To the first approximation MC/data for validation and Chi-squared for regression will be used. A number of comparison plots with superposition of symbols specifying data point and histograms representing model predictions will also be available.

This document contains the requirements for procedures and applications to enable a Geant4 developer to validate the hadronic models for a Geant4 release. This validation framework has been requested to ensure a higher degree of consistency in executing the tests, publish the results in a central place, and guarantee some quality assurance of the test procedures.

The requirements document describes 'what' the framework does, but leaves 'how' this is done to the design stage. The requirements have one of three priorities.

- Critical: the framework is useless without it
- Expected: expected in an early release
- Desired: not needed in the first release, but nice to have

## 2 Terminology

For the purpose of this document, we define some terms to describe the specific functionality.

**Model:** a set of classes implementing or to be used in a physics process for a certain primary particle(s).

**Test Executable:** a set of classes and libraries that define a hadronic physics model, compiled into an executable.

**Starting Condition:** these are the input parameters for the tests, like primary particle type, energy, and materials. The starting conditions are kept in an ASCII file under version control.

**Test:** a test is a combination of one test executable and one starting condition.

**Group of Tests:** a group of tests is a named collection of tests, it is usually one starting condition combined with different test executables.

**Test Results:** the result of executing a test is a Root file containing histograms. There is usually one Root file per test.

**Reference:** a reference is a test result from a previous release.

**Experimental Data:** these are published data points used for comparison. They are also stored in ASCII files under version control.

**Comparison:** there are three types of comparisons:

- a) a reference histogram super imposed on a test result
- b) experimental data superimposed on a test result
- c) all histograms for a group of tests superimposed on one axis.

**Regression Tests:** comparison of physics output from release to release.

**Validation Tests:** comparison of physics output with published data.

## 3 Workflow for Testing a Release

Geant4 hadron models are released to the public once or twice a year. In addition, there are about 10 internal releases per year. The internal releases are limited to the members and institutions of the collaboration. Internal and public releases are validated, but only public releases are published.

About 10000 tests are executed to validate the hadron models of a release.

Geant4 has on the order of 500 users who read the comparisons, and about 10 developers who execute the hadronic model tests. The majority of users (500) are potentially interested in the public release comparison (published results).

The testing is done by the developers and the internal release test results and comparisons are only of interest to the (<10) developers.

The steps in validating a release are:

1. Select and build tests and group of tests.
2. Execute tests
3. Compare the test results
4. Store the test results and comparisons
5. Publishing the test results (for a public release only)

## 4 Selecting Tests and Test Groups

The first step in validating a release is to select tests and groups of tests. This may include creating new test and new grouping.

No.	Requirement	Source	Priority
G-100	<b>Creating a Test:</b> A Geant4 developer shall be able to select a test executable and a starting condition to create a test.	J.Yarba 2-2009	Critical
G-110	<b>Test Contents:</b> Each test shall have a name, creation date, the name of the test executable, a copy of the start condition, and owners associated with it.	J.Yarba 2-2009	Critical
G-120	<b>Creating a Test Groups:</b> A Geant4 developer shall be able to combine tests into test groups.	J.Yarba 2-2009	Critical
G-130	<b>Test Group Contents:</b> Each test group shall have a name, creation date, list of tests, and owners associated with it.	J.Yarba 2-2009	Critical
G-140	<b>Test Group Change History:</b> A test group shall have a persistent change history.	J.Yarba, S.Banerjee 2-2009	Desired
G-180	<b>Test and Group Persistence:</b> Tests and test groups shall be persistent, meaning their definition shall be available from one release test to the next.	J.Yarba 2-2009	Critical
G-200	<b>Selecting Tests:</b> The Geant4 developer shall be able select tests and test groups.	J.Yarba 2-2009	Critical
G-205	<b>Automated Test and Group Selection:</b> Since the number of tests is often very large (40,000). The developer shall be able to select tests or a group of tests automatically. The intent of this requirement is to avoid the developer to select 40,000 tests manually.	G.Folger 6-2009	Critical
G-210	<b>Central Services:</b> selecting tests and building test groups shall be centralized on a server.	S. Banerjee S.Gysin 5-2009	Desired
G-190	<b>Web Interface for Selecting and creating Test:</b> The creating, selecting, and grouping of tests shall be available through a web interface.	J.Yarba, S. Banerjee 2-2009	Desired

## 5 Executing Tests

We can assume the tests for a release are created and selected. The next step is the execution of the tests. There is a diversity in the number of test executions and the average execution time of a test depending on the developer. A quick review of present hadron validation usage is summarized below:

## GEANT4 Hadron Validation Framework Requirements

Developer	Number of test runs	Approximate CPU Usage
1	20000	80000 hr
2	35	1500 hr
3	40	1000 hr
4	90	600 hr
5	42	200 hr
6	24	100 hr

This diversity in number of tests and CPU usage will be reflected in the requirement document.

The variance in CPU hours is due to the extent of the simulation. The 80,000 hours include a full high energy hadron shower simulation in several detector configurations.

No.	Requirement	Source	Priority
G-300	<b>Number of tests:</b> An average release test has about 20000 tests (assumes hadronics only). The framework shall be able to accommodate 40000 tests.	S. Banerjee 5-2009	Critical
G-305	<b>CPU hours:</b> The current CPU hours consumed by the hadronics tests is on the order of 100,000 hours. The validation framework shall be able to accommodate twice as much.	C.Green 5-2009	Critical
G-307	<b>Completion Time:</b> The tests for a release shall be complete within 3 weeks.	C.Green 5-2009	Critical
G-315	<b>Interrupt Option:</b> The Geant4 developer shall have the option to cancel and restart a test.	C.Green 5-2009	Critical
G-320	<b>Storing Results:</b> The test results (Root histograms) shall be available for 2 months after execution.	S. Banerjee 2-2009	Critical
G-325	<b>Promoting Results to Reference Store:</b> The developer shall be able to promote a test result to the official reference store. (see G-600)	M.Fishler 6-2009	Critical
G-330	<b>Access Results:</b> The test results (Root histograms) shall be accessible to all authenticated users.	S. Banerjee 2-2009	Critical
G-335	<b>Central Services:</b> executing tests shall be centralized on a server.	S. Banerjee S.Gysin 5-2009	Desired
G-340	<b>Web Interface for Execution:</b> There shall be a web interface to execute the selected tests.	S. Banerjee, J.Yarba 2-2009	Desired
G-350	<b>Diversity:</b> The system shall accommodate a user with high CPU consumption without penalizing a user with a small CPU consumption.	S. Banerjee, S. Gysin 6-2009	Critical

## 6 Comparison

This section describes the types of comparisons used for testing. The histograms obtained from the executed tests will be compared with static data from ASCII files under version control (experimental data). They will also be compared with references from previous release(s). There will be an option to save the results from these comparisons to the reference store. A release validation will typically have 500 comparisons to save.

No.	Requirement	Source	Priority
G-400	<b>Compare with Experimental data:</b> The Geant4 developer shall be able to compare a test result with a static experimental data file.	S. Banerjee 2-2009	Critical
G-410	<b>Static Experimental data:</b> The experimental data file for comparisons shall be under version control.	S. Banerjee 2-2009	Critical
G-420	<b>Compare with Reference:</b> The Geant4 developer shall be able to compare a test result with a reference test result.	S. Banerjee 2- 2009	Critical
G-430	<b>Compare within Groups:</b> The Geant4 developer shall be able to view a superimposed view of all test results for a group of tests.	S. Banerjee 2-2009	Critical
G-440	<b>Comparison Output:</b> The output of each comparison shall be Root files (histograms) and meta information.	S. Banerjee 2- 2009	Critical
G-450	<b>Comparison Meta Information:</b> The meta information shall contain the following: - time of execution - test name - reference name - tester's name - comment	S. Banerjee 2- 2009	Critical
G-470	<b>Quality Measure:</b> MC/data for validation, and Chi-squared for regression shall be available for quality measures.	S. Banerjee 5- 2009	Expected
G-472	<b>List of Measure:</b> We expect to develop more quality measure. As the list grows, the user shall be able to view all available measure.	S.Gysin 6-2009	Expected
G-474	<b>Choice of Quality Measure:</b> The developer shall be able to choose the quality measure for a comparison.	S.Gysin 6-2009	Expected
G-475	<b>Central Services:</b> comparison of tests shall be centralized on a server.	S. Banerjee S.Gysin 5-2009	Desired
G-460	<b>Web Interface for Comparisons:</b> There shall be a web based interface to make the comparisons.	S. Banerjee, J.Yarba 2-2009	Desired

## 7 Storing Results

Storing the results and comparisons is the next step after making the comparison. The output of the comparison is a Root file and meta data as described in section requirement G-430 and G-440.

This section describes how new references will added to the reference store.

No.	Requirement	Source	Priority
G-500	<b>Saving as Reference:</b> The Geant4 developer shall be able to submit the comparisons and/or test results to a long term reference store.	S.Gysin 2-2009	Critical
G-510	<b>Reference Store capacity for comparisons:</b> Comparisons stored in the reference store shall be available for at least 10 years. (see G-300)	S. Banerjee 2-2009	Critical
G-520	<b>Reference Store capacity for test results:</b> Test results store in the reference store shall be available for at least one year.	S. Banerjee 2-2009	Critical
G-530	<b>Saving to Short Term Storage:</b> The Geant4 developer shall be able to store comparisons to short term storage.	S. Banerjee 2-2009	Critical
G-540	<b>Short Term Store:</b> The comparisons shall be stored for no longer than 2 months in short term storage.	S. Banerjee 2-2009	Critical
G-570	<b>Setting Permissions:</b> The test owner shall be able to assign read and write permissions for each stored item.	J. Yarba 2-2009	Desired
G-575	<b>Central Services:</b> storing results shall be centralized on a server.	S. Banerjee S.Gysin 5-2009	Expected
G-550	<b>Web Interface for saving:</b> There shall be a web based interface to enable the user to choose a storage option.	S. Gysin, J. Yarba 2-2009	Desired
G-560	<b>Web Interface for browsing stores:</b> There shall be a web interface for the user to browse, add, change, or remove reference information from the reference store and from the short term store.	S.Gysin, J. Yarba 2-2009	Desired

## 8 Publishing Comparisons

Publishing the comparisons is the final step in the testing procedure.

The comparisons for the official release are published to a webpage, to make them available to the wide community of Geant4 users. The comparisons should illustrate how a Geant4 hadronic physics model performs, and highlight the improvements.

## GEANT4 Hadron Validation Framework Requirements

---

The published results page will be the most exposed part of this web application. It is expected to be viewed by all users (500), maybe 5 users at a time

No.	Requirement	Source	Priority
G-600	<b>Explicit Publishing:</b> Publishing shall be the choice of the Geant4 developer, it shall not be automated	J.Yarba 2-2009	Critical
G-610	<b>Selective Publishing:</b> The developer shall be able to choose a subset of results to publish.	G.Folger 6-2009	Critical
G-630	<b>Administer the published results:</b> A user with administrative authority shall be able to create hierarchies to organize the comparisons.	J.Yarba 2-2009	Critical
G-640	<b>Published information content:</b> The information published shall include: <ul style="list-style-type: none"> <li>• Comparisons (3)</li> <li>• Version control tag of the software to run the test</li> <li>• References</li> </ul> Note: historically the Tests are not stored.	J.Yarba 2-2009	Critical
G-640	<b>Published information display:</b> The published information shall display the charts from the comparison.	J.Yarba 2-2009	Critical
G-650	<b>Web interface for browsing published results:</b> There shall be a web interface for browsing published results.	J.Yarba 2-2009	Critical
G-660	<b>Web interface load:</b> The web page displaying the published information shall be able to accommodate 20 simultaneous users.	J.Yarba 2-2009	Critical
G-670	<b>Central Services:</b> publishing results shall be centralized on a server.	S. Banerjee S.Gysin 5-2009	Critical

## 9 Authentication Requirements

These are the authentication requirements that apply to each step. Authentication identifies the user as who he says he is. This is done via a user name and password, certificates, etc. Authorization is an additional step and checks if a user is authorized to take a specific action.

When implementing these requirements, the permission on the Geant4 repository should be considered.

No.	Requirement	Source	Priority
G-10	Authentication shall be required on the central server where the tests are to be executed.	S.Banerjee, S.Gysin, J.Yarba, 3-2009	Critical
G-30	<b>Authorization:</b> the following tasks shall require the user to be authorized: <ul style="list-style-type: none"> <li>• Creating tests or test groups</li> </ul>	S.Banerjee, S.Gysin,	Critical

## GEANT4 Hadron Validation Framework Requirements

	<ul style="list-style-type: none"> <li>• Storing comparisons and test results</li> <li>• Publishing results</li> <li>• Organizing the published display</li> </ul>	J.Yarba, 3-2009	
G-20	<p><b>Authentication only:</b> the following tasks shall require the user to be only authenticated (vs. authorized):</p> <ul style="list-style-type: none"> <li>• Selecting tests</li> <li>• Executing tests</li> <li>• Comparing tests</li> </ul> <p>The following tasks do not require authentication:</p> <ul style="list-style-type: none"> <li>• Viewing published results</li> </ul>	S.Banerjee, S.Gysin, J.Yarba, 3-2009	Critical
G-40	<p><b>Ownership of tests:</b> the following tasks shall be limited to the authors of the tests and the administrator:</p> <ul style="list-style-type: none"> <li>• Editing tests or test groups</li> </ul>	S.Banerjee, S.Gysin, J.Yarba, 3-2009	Critical
G-50	<p><b>Ownership of comparisons:</b> the following tasks shall be limited to the person who executed the tests and the administrator.</p> <ul style="list-style-type: none"> <li>• Comparing tests</li> <li>• Publishing results</li> </ul>	S.Banerjee, S.Gysin, J.Yarba, 3-2009	Critical

## 10 Deployment and Operations

G-700	<p><b>Deployment:</b> the centralized service shall be deployed on a server matching the hardware and software capacity in the previous requirements.</p>	S.Banerjee, S.Gysin 5-2009	Critical
G-710	<p><b>Operation:</b> the web application shall be operated and maintained by an administrator designated by the Geant4 collaboration.</p>	S.Banerjee, J.Yarba	Critical

## 11 Issues

1. Where do we introduce such concepts as “failure”, “deviation”, “tolerance” and the likes? It may be part of the comparison, and affect the decision whether or not results are to be published.

## 12 References

- [1] S. Agostinelli et al. (Geant4 collaboration), Nuclear Instruments and Methods **A506**, (2003) 250.  
J. Allison et al. (Geant4 collaboration), IEEE Trans, Nuclear Sciences **53**, (2006) 270.