Review of the Geant4 Project
January 19 – 20, 2009

Report submitted on March 16, 2009

Members of the Review Team:

Irène Buvat, INSERM
Fabio Cossutti, INFN Trieste
Frank Gaede, DESY
Andreas Morsch, CERN
Robert Reed, Vanderbilt University
Sayed Rokni, SLAC
Hirohi Sakamoto, University of Tokyo
Lembit Sihver, Chalmers University of Technology
Charles Young, SLAC (chair)

1. Executive Summary

The committee is impressed by the progress made by the Geant4 collaboration since the last review in April 2007. Many issues identified in the previous review have been addressed and resolved, and some of the others are being actively pursued. We take particular note of the reorganization of the two electromagnetic (EM) groups into a more coherent effort. There has also been important progress in the hadronic group and in the area of documentation. We encourage Geant4 to continue some recent efforts so the few outstanding items will be dealt with expeditiously.

2. Introduction

This report contains the result of the delta review of the Geant4 project held at CERN on January 19-20, 2009 to assess the progress on the recommendations from the previous review in April 2007. The Geant4 Oversight Board appointed members of the review committee and provided its mandate. In performing this task, the committee may also address auxiliary issues that it considers relevant.

Geant4 provided before the review a written summary of responses to the previous set of recommendations. The review started with a brief introduction by John Apostolakis, spokesperson of the Geant4 collaboration. This was followed by eight reports summarizing the progress in various domains, in particular the recommendations from the last review. The last report was on manpower and resource planning. The committee discussed its observations and findings, and then presented its preliminary recommendations on the last day. The complete agenda is included in this report as Appendix A.

The committee would like to thank the Geant4 collaboration and CERN for their organization and support of the meeting, and for their hospitality.

3. Follow-ups to Previous Recommendations.

There were thirty one recommendations from the last review in April 2007. A summary of the responses was provided before the meeting, and the issues were addressed in detail during the presentations. This report is divided into sub-sections each covering one area in the previous review report. We start by recalling the recommendations from the April 2007 review, then report the observations at this review, and finally report any new recommendations.

3.1. Electromagnetic Physics

Previous Recommendation 1: We recommend setting up an easily accessible central repository with detailed references to data comparisons and validation papers for the models, e.g. a web page linked from the EM home page.

Previous Recommendation 2: We recommend rapid integration of the ICRU 73 heavy ion stopping power model.

Previous Recommendation 3: We recommend providing guidance on the tradeoff between physics accuracy and computing speed. For example, collect and publish via the web results from users in various domains regarding computing speed vs range cut, and provide a Wiki-like forum where users can document their experience; etc.

Previous Recommendation 4: We recommend providing guidance on the choice between the two EM models for specific particle species, energies, etc.

Previous Recommendation 5: We recommend integrating the two EM models into a single package, similar to what exists in hadronic models. This will allow a user to choose one model in one energy range and the other model in a different energy range in order to optimize physics and computing performance for his application.

The EM working group has improved considerably since the last review. We observe significant progress in many areas:

- A common interface and validation plan for the standard and low energy electromagnetic processes is being put in place, and the work has been re-organized towards a tighter link between the different components. When this work is finished, it should open the possibility to combine low-energy and standard EM models when designing a simulation.

- A number of models have been updated and new ones have been added, e.g.

    - Density effect in the relativistic Bremsstrahlung model for E > 1 GeV.
    - ICRU 73 data for ion ionization.
    - Multiple scattering for specific particles.

        - Cerenkov effect.
        - Inclusion of Birks Law in scintillation process.

    - New functionalities have been introduced.

        - Spline in addition to linear interpolation for physics tables.
        - Ability to configure EM models per energy range and G4Region.

    - There is an effort to identify the most computationally expensive tasks, to be followed by optimization of code.

    - The EM documentation has been completely reorganized for improved clarity and accessibility. An extensive review of the different processes has been started, with emphasis on the validation on available real data. This improvement and validation effort has been accompanied by a better organization of references and a set of dedicated web Twiki pages.

    - There is a coherent work plan for 2009 and later. It addresses code infrastructure, the evolution of existing models, the inclusion of new models, their validation, as well as documentation.

The standard EM and low-energy EM groups are now working more coherently, coordinating the development and validation of different models and on documentation. They seem to be structured and organized, with relatively clear plans for the future. They could even serve as an example for other working groups. This is a major achievement of Geant4 in the period elapsed since last review.

We encourage Geant4 to build on this progress, and expeditiously complete the many tasks that have been recently initiated.

The growing number of models and options available for EM processes is a welcome evolution of the code. There is also a large body of validation results for specific applications. However, this wealth of choices and information can also reduce user friendliness, as some users may find it difficult to determine which models will be appropriate for their applications. Care should be taken so users do not get lost when choosing between various EM models and options.

Recommendation 1: Remove less accurate models and/or options when they do not provide significant benefits elsewhere. For example, eliminate the linear interpolation scheme unless spline interpolation is significantly more computationally intensive.

Recommendation 2: Provide guidance to users on the choice of models and options. This can be in the form of suggested physics lists for specific communities, or a more general decision tree that leads to a physics list.

The effort on computational performance is laudable. This is an issue that affects not only the EM domain, but the entire Geant4 toolkit. It should be addressed across the board by Geant4.

Recommendation 3: Geant4 management should use the mini working group as a model, and foster a coherent and broad based effort on continual computational performance improvements.

3.2. Hadronic Physics and FLUKA

Previous Recommendation 6: We recommend continuing the dialog with the FLUKA collaboration to find a cooperative way to provide users with the best features of Geant4 and FLUKA. One possible implementation could be a loosely coupled interface between Geant4 and FLUKA hadronic physics model.

Attempts to incorporate the FLUKA physics model into Geant4 have not succeeded. Alternatives such as Virtual Monte Carlo can fulfill some user needs, but do not provide a complete solution. We recognize the effort made by the Geant4 collaboration to resolve this problem, and believe that it is unrealistic to expect the FLUKA physics model to be interfaced to Geant4 in the foreseeable future.

The FLUKA hadronic physics model represents a huge investment by the community, and the code has undergone extensive validation. By understanding the physics difference between the different models used in FLUKA and Geant4, including assumptions and parameterizations, one can then improve the corresponding implementations in Geant4 without necessarily incorporating the FLUKA code.

Recommendation 4: Study the assumptions and parameterizations of the physics models in FLUKA. Improve the corresponding models implemented in Geant4 in cases where FLUKA physics models are superior.

Recommendation 5: Support the further development and maintenance of tools that allow the user to combine Geant4 and FLUKA, such as the Virtual Monte Carlo and FLUGG.

3.3. Other Hadronic Physics Issues

Previous Recommendation 7: We recommend that Geant4 develop detailed plans to improve the hadronic package, in particular to improve the simulated shower shapes.

Previous Recommendation 8: We recommend putting in place a set of simple hadronic benchmarks which allow quickly identifying very basic problems like disagreement with well known shower shapes.

Previous Recommendation 9: We recommend that Geant4 accelerate the development of physics models to address these needs.

Much progress has been made in the area of hadronic physics.

- Differences between data and simulation in longitudinal shower development have been reduced from ~25% to the few % level.

- Development of new physics models is on-going.

- There is now a more accurate modeling of double and triple differential cross-sections.

- Prioritization has been done and a resource plan is available.

However, there has been little benchmark of fragment production cross sections from nucleus-nucleus reactions. Inclusive fragment production cross sections from thin target data give a rough estimate of the quality of the models. If the inclusive cross sections are wrong, the double and triple differential cross sections cannot be correct. However, this gives less information about the quality of the hadronic physics models than when comparing simulated double and triple differential cross sections with measured ones.

Recommendation 6: We recommend that comparisons of differential quantities which are more sensitive to the potential physics differences between data and simulation are performed.

3.4. Computing Performance

Previous Recommendation 10: We recommend extending the scope of the computing professionals to review and optimize all Geant4 code.

Previous Recommendation 11: We recommend that Geant4 encourage users to monitor their applications, and provide feedback so additional "hot spots" can be identified.

Previous Recommendation 12: We recommend the creation of a performance optimization guide. It is likely that such information already exists and just needs to be collected into one document.

Previous Recommendation 13: We recommend providing a simple mechanism for users to turn off "irrelevant" processes for a given region.

Previous Recommendation 14: We recommend providing guidance on the use of VRT in Geant4 so users can better take advantage of them. References to *publications and data regarding the validation of the VRT methods in Geant4 should be made available in the performance optimization guide.*

Previous Recommendation 15: We recommend systematic tracking of code performance for each part of the code, and for each physics model. Comparisons with previous versions should be an integral part of the release notes.

Previous Recommendation 16: We recommend that Geant4 keep itself abreast of developments in the area of multi cores and advanced instructions, so it can take advantage of them when there is sufficient infrastructure and support to do so.

Previous Recommendation 17: We recommend that Geant4 publish a plan regarding the expected computing performance of the toolkit over the next five years.

The Geant4 collaboration has addressed the issue of computing performance and the communication of performance topics relevant to the user community. The preparation of a dedicated document, in the form of a Twiki page with hints on improving the performance, the monitoring of the performance at every release, and the tracking of these metrics in release notes are very valuables results.

The model of improving the code based on the needs and problems of the user community is positive. We believe this should be complemented by proactive systematic code improvements.

Recommendation 7: We recommend that Geant4 continue to monitor for hot spots, conduct periodic code reviews, and address code revisions needed for improvement.

Many Geant4 applications involve the detailed tracking of many independent particles through complex geometries. The code can be very memory intensive. When several applications are run on multi-core machines, one application per core, the total memory requirement multiplies. However, if an application can simultaneously track multiple particles on multiple cores, the overall memory requirements can be reduced significantly. Hardware costs should see a corresponding decrease for the same overall computational performance. We are happy to see that users who initiated such a study have been brought into the Geant4 collaboration to provide closer interaction. Adapting the Geant4 toolkit to utilize multiple cores is a major undertaking that can span several Geant4 releases. There was not a coherent plan presented to address the risk that multi-core multi-thread development based on a frozen version of Geant4 could result in incompatibilities with the latest version of Geant4.

Recommendation 8: Development of multi-thread capability must be coordinated with the other developments of Geant4 to minimize potential conflicts and/or divergences in code changes. We recommend that multi-threading capability is made a fully supported standard core feature of Geant4 as soon as possible.

We welcome the suggestion of a major code design assessment and revision. Geant4 development started many years ago, and it is appropriate to reassess its design at this time. In addition, what might have been a prudent decision to avoid using all the features of the C++ language at the time may not be the best option today; the language is more mature and there is much greater familiarity in both developer and user communities. A well thought out redesign and judicious use of more advanced language features as well as modern hardware capabilities can make Geant4 more maintainable and help keep Geant4 vital as the preferred simulation toolkit for the future. The formulation and enforcement of coding rules help to keep the code readable for future generations of developers and users. It also improves the quality and efficiency of the code.

Recommendation 9: We recommend conducting the code design assessment and evaluating changes to the Geant4 code to keep it vital. The code design assessment should be held as soon as possible. It is a good opportunity to integrate multi-core capabilities into the Geant4 toolkit.

The wide diversity in the user community both in their areas of application and in their sophistication in the use of Geant4 makes it difficult to provide usage guidelines. We also understand the limitations due to manpower constraints. Nevertheless, we encourage the Geant4 collaboration to keep this suggestion in its work plan.

3.5. Usability Issues

There were no specific recommendations in the usability section of the previous review report. The issues were presented in other sections.

3.6. Physics Validation

Previous Recommendation 18: We recommend aggressively populating the database with all relevant experimental data, as well as validation results provided by others.

Previous Recommendation 19: We recommend continuously and systematically benchmarking against other Monte Carlo transport codes.

An impressive program of systematic physics validation has been carried out. For example, hadronic Test30 is for the low energy range, Test35 is for the medium energy range, and Test45 looks at neutron yields. On the electromagnetic side, there have been validations of Geant4-DNA processes and therapeutic applications. The hadronic results are available in two complementary web sites, one hosted at CERN and the other hosted at Fermilab. There is also a web page detailing standard EM validation results.

Recommendation 10: The hadronic validation web pages should be consolidated into one single entity, and a similar web page for low-energy EM group should be created.

Some comparisons with experimental data and with other Monte Carlo transport codes have been performed. For example, Geant4/FLUKA benchmarking has been performed as part of the LCG simulation validation project. In addition, Geant4 plans to participate in the 2009 Hadronic Shower Simulation Workshop, and in a number of head-to-head comparisons with MNCPX, PHITS, etc.

Recommendation 11: Continue the regular benchmarking and comparison with other Monte Carlo transport codes.

Recommendation 12: Provide pointers to benchmarking summary information when available, and create the summaries when they do not currently exist. Given the widespread user interest in FLUKA for example, it would be particularly helpful to users to have ready access to that set of benchmark results.

3.7. Release Validation

Previous Recommendation 20: We recommend defining and automating a common validation procedure to be run for every release, monitoring a comprehensive set of variables and exploiting the comparisons with the collected experimental results.

Previous Recommendation 21: We recommend defining quantitative metrics for validation results.

Previous Recommendation 22: We recommend that all validation results, both the quantitative metrics and the underlying distributions, be made easily accessible to the user.

The current release validation procedure started in June 2006 and has incrementally improved. The present test suite runs on large statistics and makes use of Grid resources. The submission and evaluation procedures are partially automated. They report any deviation from previous results (significant at P-level of 1%), provide summary information of key observables and summarize the number of deviations found. Comparisons are generated as plots, so that it can be checked if a pattern of deviations is present.

We believe that the level of automation is about right. Checking for patterns of small but systematic deviations is an important part of the testing that cannot be easily automated. We agree with the plan to increase the number of tests for release validation.

It is important to quantify and then track the validation results. For this reason the previous review has recommended to define quantitative metrics for validation results. Except in areas where models describe the data well, the Geant4 tests still do not have such a metric. We recognize that in cases where models do not describe fully the relevant data, metrics such as chi-squared tests are not very useful. However, empirical metrics such as MC/data ratio can still be very useful.

Recommendation 13: Adopt a range of metrics to characterize as broad a range of validation results as possible.

In cases where there are discrepancies between data and simulation, it is important to probe parameters closest to the underlying physics of the deviation. This may not be in the original suite of tests, and would require defining better comparisons and finding more sensitive metrics.

Recommendation 14: Look for better metrics as required by the data, and add them to the suite of tests to be tracked over time.

The accessibility of the results of the validation has been improved since 2007. The hadronic validation results have been collected in two points of access and the electromagnetic tests into one point of access, both accessible from the main Geant4 web page. However, documentation (explanation) of the tests, automation of the publication of the results and completeness of the published results still need significant improvements.

Recommendation 15: We recommend pursuing the outstanding issues in this area expeditiously.

3.8. Documentation

Previous Recommendation 23: We recommend updating the documentation on the web right away, and formulating a plan to have periodic reviews of the web site to keep it up to date.

Previous Recommendation 24: We recommend that Geant4 document the limitations, and validity and applicability ranges of the different EM and hadronic models, and the physics lists. Where models have overlapping validity ranges, document the tradeoffs.

Previous Recommendation 25: We recommend that Geant4 improve its code documentation. Example approaches by other groups include Doxygen, and README files that can be browsed in LXR.

Previous Recommendation 26: We recommend improving the release notes with the addition of expected changes in physics and computing performances, while keeping the current extensive list of code change descriptions. We encourage strengthening the editorial coordination of the release note preparation.

Previous Recommendation 27: We recommend that Geant4 review the current installation procedure with the aim of adding conveniences for users, such as additional defaults, self configuring procedure, a template `.spec` file to create an RPM in the user's environment.

There have been improvements in various areas of documentation, e.g. the validation of EM and hadronic physics, and new medical physics pages provided by new contributors. The priority given to physics list documents is appropriate when the effort is resource limited. Proper documentation is fundamental to the usability of the Geant4 toolkit; it is important that adequate resources be found for it.

Recommendation 16: Acquire or redirect resources to improve the documentation and keep up with updates.

Recommendation 17: We recommend the adoption of periodic reviews of the documentation.

The Physics Reference Manual is the primary source of information on physics models. A prototype web page linked to individual models is available, but has been difficult to maintain to avoid inconsistencies and duplications of other documents. We see this as indication of potential confusion of the user, and endorse Geant4 plans to improve these pages. The Geant4 team also has plans to improve this documentation to include a description of the tradeoff between accuracy and speed.

Recommendation 18: Improve documentation design to achieve greater consistency across its sections.

The Geant4 Software Reference Manual (SRM) contains annotated header files. An LXR installation with search capability and cross-linking between classes is linked from the Geant4 web site. There are Doyxgen documentation pages hosted at KEK; however, they are limited by the source code format needed by SRM's automatic extraction. We do not see strong reasons to change source code format to be Doxygen compatible. Currently, LXR, Doxygen and SRM pages appear in separate places, and their different uses are not clear.

Recommendation 19: Provide a clear explanation to the users, for example in the user support page, of the roles of these three documentation tools (LXR, Doxygen and SRM).

Contents of the release notes are considerably improved in the recent releases. Draft notes are now checked within the Geant4 collaboration prior to publication to insure higher quality. New sections have been added on user code migration, and computing and physics performance.

There has been little progress in the area of installation tools due to resource limitation, and the loss of key personnel in particular. The situation has become more urgent with the imminent withdrawal of support for the present underlying tool.

Recommendation 20: We recommend that Geant4 urgently come up with a new installation mechanism, and a work plan defined as soon as possible. Many HEP experiments have dealt with similar installation issues of comparable or larger scales for many years, and Geant4 should build on this expertise.

3.9. User Support

Previous Recommendation 28: We recommend setting up a user community supported Wiki as a collaborative tool. Geant4 should advise which documentations are appropriate to be moved to it.

Previous Recommendation 29: We recommend that Geant4 continue its vigorous support to implement and integrate user driven physics models.

Usage of Wiki has started in some areas. While the core documentation should be maintained by Geant4 to ensure accuracy and consistency, other documents can benefit from a broad base of community support. We expect that the maintenance cost of Web documents can be lowered by the broader adoption of Wikis in areas where developers and users can correct and update the contents.

Many physics models have been proposed and accepted. We understand that including a new physics model requires considerable effort on the part of Geant4 not only in the implementation phase but also in its maintenance. Prioritization is necessary in the face of limited resources, and should involve the user community.

Recommendation 21: We recommend the involvement of the user community in the long-term support of physics models, and suggest that Geant4 include this factor in the prioritization.

3.10. Other Issues

Previous Recommendation 30: We recommend that Geant4 develop manpower and resource plans for the next five years. They should be consistent with timely delivery of features and capabilities to the user community.

Previous Recommendation 31: We recommend that Geant4 seek the additional support implied by the manpower and resource plans from all available sources.

The project plans for 2009 and beyond were presented. There are detailed milestones in both EM and hadronic domains, some of which are contingent on additional manpower becoming available. These milestones will allow Geant4 management to monitor and assess its own progress, and to re-direct resources as appropriate.

The manpower situation was also presented. A large fraction of the existing manpower has been provided institutionally. It is understandable that institutions and their funding agencies would like to direct their resources to address issues critical to their applications. Individual contributors making up the rest of the manpower pool are also likely to focus on their specific needs. This leads to a situation where a particular domain may have needs that cannot be easily addressed with the available manpower.

Proposals have been submitted to several agencies in different countries. We commend Geant4 in its active pursuit of additional support, which if successful will bring in significant new resources. However, it appears likely that Geant4 will still be without complete freedom to re-direct the collective resources it has.

Recommendation 22: We recommend that Geant4 continue to seek additional support from all available sources, especially where the resource can be more flexibly utilized, such as students and interns, and seek to extend the Geant4 collaboration with new institutions.

7. Conclusions

The Geant4 collaboration has made significant progress in addressing the issues identified in the review of April 2007. Some topics, such as the reorganization of EM groups, have been dealt with thoroughly. Work on some other topics is far along, while work on some other areas has started but has not progressed as far. We suggest that the Geant4 collaboration develop a prioritized work plan with milestones in order to aggressively address the outstanding items.

Appendix A: Review Agenda

All the presentations are linked to the review web page at http://indico.cern.ch/conferenceDisplay.py?confId=44900.

January 19, 2009

09:00	Executive Session	
10:30	break	
11:00	Preparatory session	
13:00	Lunch	
14:00	Introduction	John Apostolakis
14:15	Hadronic Physics	Dennis Wright
15:30	break	
16:00	Hadronics – Validation	Dennis Wright
17:15	Computing Performance	Gabriele Cosmo Makoto Asai
18:15	Usability – User Support	John Apostolakis Makoto Asai

January 20, 2009

09:00	Executive session	
09:30	Usability - Documentation	Katsuya Amako John Apostolakis
10:30	Electromagnetic physics	Vladimir Ivanchenko Sebastien Incerti
11:00	break	
11:30	EM Validation	Vladimir Ivanchenko Sebastien Incerti
12:30	Release Validation	John Apostolakis Gabriele Cosmo
13:00	Lunch	
14:00	Manpower, resource planning	Petteri Nieminen John Apostolakis Makoto Asai
14:45	Executive session	
16:00	break	
16:30	Supplemental Q&A	
17:00	Executive session	
18:00	Close-out	