The Monte Carlo Method: Versatility Unbounded In A Dynamic Computing World
Chattanooga, Tennessee, April 17–21, 2005, on CD-ROM, American Nuclear Society, LaGrange Park, IL (2005)

THE Geant4 KERNEL: STATUS AND RECENT DEVELOPMENTS

John Apostolakis, Gabriele Cosmo
CERN, Geneva, Switzerland
John.Apostolakis@cern.ch, Gabriele.Cosmo@cern.ch

Makoto Asai
SLAC, Menlo Park, CA, U.S.A.
Makoto.Asai@slac.stanford.edu

On Behalf of the Geant4 Collaboration

ABSTRACT

Geant4 is used in production for several large HEP experiments and is utilized for applications in many other fields including space and medical applications. Its capabilities have been continuously extended, as its performance and modeling are enhanced. An overview is given of recent developments in kernel areas of the toolkit, and briefly of the new modeling physics options. These include, amongst others, performance optimization using production thresholds in different geometrical regions for scalable use in very complex setups, the extension and improvement of the propagation in external electromagnetic fields, and the inclusion of event biasing and tallying.

Key Words: Simulation, Particle interactions, Geometrical modeling, Software engineering, Object-oriented technology

1 INTRODUCTION

The Geant4 toolkit [1] provides comprehensive physics modeling embedded in a flexible structure provided by its kernel. The choice of physics modeling and the functionality of a robust kernel enable users to adapt it for diverse applications. We provide a short summary of key developments undertaken by the Geant4 Collaboration [2] over the past four years, with emphasis on the Geant4 kernel and its relation to physics processes.

The Geant4 kernel provides key services for the toolkit: tracking, geometry description and navigation, abstract interfaces to physics processes, material description, management of events, run configuration, stacking for track prioritization, a framework for the creation of hits and digitization, a framework for fast simulation (shower parameterization), event biasing options, and interfaces to external frameworks.

Physics processes included cover the range of physical interactions, and are organized as electromagnetic (EM), hadronics, decay and optical processes. Particles tracked include leptons, mesons, hadrons and photons. Photons of optical wavelengths are treated separately from gammas. Different implementations of physics process are offered in several energy ranges for many physical processes. These typically use complementary modeling approaches and provide different trade-offs between accuracy and computing requirements.

For each established application area, a set of physics simulation packages is provided. Each package is a configuration of physics processes (a ‘physics list’) that is tailored to address the specific requirements for accuracy in the most relevant physics observables, with varying levels of CPU resources.

Physics performance has benefited significantly from the extensive comparisons with established experimental data undertaken in collaboration with and the independent validation efforts of several HEP experiments, groups and numerous other users.

In addition to the kernel and physics, Geant4 provides interfaces to enable users to interact with their application, and save their results. Visualization drivers connect to OpenGL, VRML and the high-quality DAWN system. Text-based and Graphical User Interfaces (GUIs) and a flexible framework for persistency are included.

In order to meet the computing requirements of shielding and other applications, commonly-used efficiency-enhancing techniques are offered in the toolkit: variance reduction (event biasing) techniques and options for fast simulation (shower parameterization) are built-in.

2 RECENT DEVELOPMENTS

Geant4 release 3.0 (December 2000) enabled the shooting of general ions as primary particles or by a radioactive source, a new cross-section and energy-loss parameterization for energy loss of ions, cross-section biasing options for hadronic interactions and transition radiation models, Release 3.1 (April 2001) introduced the ‘assembly’ volume, grouping volumes in arbitrary manner and placing them as a single combined volume, an improved interface to physics generators handling the proper decay time of initial particles, and the storing/retrieving calculated physics vectors for better performance in large setups, including HEP experiments.

Geant4 release 4.0 (December 2001) included enhanced stacking mechanisms for applications with a multitude of primary particles, a new model for proton-induced fluorescence, handling of time-dependent electromagnetic field and of fields local to a volume, run-time tests for volume overlapping, volume creation by reflection, Doppler broadening of low-energy neutron cross-sections, and the Chiral Invariant Phase-Space (CHIPS) hadronic model [3]. New options were added in revision 4.1 (July 2002) for X-ray transition radiation, gamma conversion to muon pairs, improved 3D optimization for parameterized volumes, high q^2 part of electro- and gamma-nuclear reactions, quark-gluon string model, and geometry importance biasing option.

Among a number of technical transitions made, most notable were the migration to the C++ Standard Template Library (STL), and the revision of examples and test programs to utilize the Abstract Interface for Data Analysis (AIDA) [4] and its implementations.

Geant4 release 5.0 (December 2002) included scoring, the Binary Cascade [5] and an implementation of the Bertini cascade [6], providing new theoretical hadronic models for energies between hundreds of MeV and about ten GeV. Revision 5.1’s key feature was the ability to define geometrical regions, and to set a different value for production thresholds (or cuts) for photons, positrons and electrons in each region (“cuts per region” functionality). This new feature is to be discussed in the following dedicated chapter.

Geant4 release 5.2 (June 2003) provided improvements in the cuts per region functionality, a performance optimization of the propagation of charged particles in a field, and revisions of the pion reaction cross-sections.

A key feature introduced in release 6.0 was a re-factored implementation of ‘standard’ electromagnetic (EM) processes. This enables refinements and specializations, and improves performance and maintainability. Release 6.0 was the first release to include our configurations of hadronic physics processes (‘physics lists’). Further physics models and new functionality were provided. As at each major release, a number of interface changes were mandated for user code.

A consolidation release (6.1, March 2004) provided enhancements to aid the stability of large-scale productions, including new tools to identify infrequent problems, which impact large productions yet can also disrupt other use. Geant4 release 6.2 (June 2004) included new models for ion reactions and improvements to the new EM process implementation.

Geant4 release 7.0 (December 2004) included the first implementation of a concrete shower parameterization, following the approach of GFLASH [7] for Geant3, an improved way of storing/retrieving physics tables and new methods to calculate volume of a solid and mass of a part of a setup. Improved handling of tracks stuck on the boundary of two volumes and a migration to the <cmath> standard mathematical library were also included.

3 THE KERNEL

The Geant4 ‘kernel’ consists of the modules that enable the simulation of physics processes, but does not include physics or processes. These include the geometry and field, particles, materials, tracking and run & event management. This section examines new developments and improvements in affecting a number of these modules, including primarily the geometry, field and run & event management. Tracking and particles are described elsewhere [1]. In addition we present a summary of the options for event biasing and variance reduction.

3.1 Geometry

The Geant4 Geometry module enables the description of models of detectors and setups and the navigation inside these models. Basic abilities include the creation of a hierarchical geometry, in which volumes are described in terms of ‘solid’ shapes. Volumes can be placed singly or in repetitive patterns. Complex shapes are created by Boolean operations on basic CSG solids, or using BREP solids. The Navigator computes the straight-line path of tracks in a volume model, and finds the intersection with the next boundary.

Capabilities for creating geometry models were enhanced, enabling a user to more easily create complex setups and change them, and to use different geometrical views of a single setup (‘ghost geometries’) in order to describe aspects of the problem. Additional capabilities enable a user to create a part of a setup, and reflect it in order to create a mirror image. The mirror includes all daughter volumes, with material and other properties.

The geometry of a setup can be created, used and later modified in a single computer job. At the start of a simulation the geometry is optimized by creating an auxiliary set of ‘voxels’ for navigation. These abilities are provided by the geometry module, and coordinated via the simulation state, which is managed by the Run module.

Several different geometries can be created in a single Geant4 application, describing complementary aspects of a setup. The primary one is the ‘mass’ geometry, which contains the ‘principal’ volumes with the materials of the setup, in which the physical interactions will be simulated. In addition to this mass geometry, a user could create a parallel geometry that can be defined for triggering fast shower parameterizations, e.g. in a simplified geometry with averaged material. Yet another geometry can describe the detector readout.

In addition it is now possible to create a ‘parallel’ geometry for importance values, in order to aid in geometrical variance reduction. This provides the basis for event biasing steered within a simpler setup, which is important in the case of complex ‘mass’ geometries.

An improvement makes optional the voxelization of a volume. This enables the use of geometries of medium complexity that change frequently or vary within a single Geant4 ‘run’ configuration (e.g. a moving target).

A new feature in the Geometry, since version 6.0, is the abstraction of the Navigator. This enables an advanced user to replace or change the active Navigator. In addition, a first consolidation simplifies the Navigator interface. Finally the state of the geometry system is simplified, and is maintained primarily in the ‘touchable’, that enables simultaneous Navigators for the same geometry.

A single volume can represent several copies, displaced or rotated, in order to save memory in case of large repetition. Two options have been available: replicas which slice their parent along an axis and Parameterized volumes which allow a user to define auxiliary set of positions and rotations of sub-volumes. New ‘division’ volumes also slice a parent volume, extending the capabilities of ‘replicas’ by providing offsets. All Constructed Solid Geometry (CSG) solids, including boxes, tubes, cones, and polyhedra, can be sliced along an appropriate axis, which is usually a symmetry axis. Divisions are implemented in terms of existing parameterized volumes, and thus can be visualized better in the case of radial slicing than existing replicas.

Further use is made of the isotropic ‘safety’, the conservatively-estimated distance to the nearest volume boundary from the current point. This has led to improvements in solids in the implementation of the safety. The solids affected were the simple (CSG) solids and the Boolean solids, created using Boolean operations of other volumes. In addition a new solid, the G4Orb, addresses use cases for a ‘full’ sphere, enabling the modeling of spheres of large dimension, e.g. one the size of a planet.

A geometry description language, GDML [8] has been created to describe in XML the volumes and material composition of a setup. A module for interfacing with Geant4 reads and writes geometry models from/into XML text files. Support for more solids and for replicas was added.

A major challenge for users is to create a consistent description of the geometry of a setup. It requires effort to avoid overlapping volumes or ones that protrude from their containing volume. Tools are required to help identify overlaps and enable the user to check his or her model geometry. Checks are done now at the time of geometry construction to ensure that the geometry model respects the rules of Geant4’s geometry module. An example rule is that only one repeated volume can be placed inside a containing volume.

During tracking Geant4 has not checked for malformed geometries – for reasons of performance and simplicity. A new option now enables some simple checking of the user's geometry model and the navigation during tracking. When using this geometry 'check' mode, Geant4 provides information on candidate intersections with volume boundaries to help users or developers to identify an underlying problem.

The full challenge of detecting 'significant' overlaps or protrusions is addressed by special tools. These include the DAVID tool which intersects the graphical representations of volumes [9], an example program OLAP that uses the full Geant4 tracking / navigation [10] and a new verification sub-module inside the geometry [11]. All these tools have adjustable intersection tolerances. The verification sub-module can run different verification tests and is accessible in any interactive Geant4 application through User Interface (UI) commands. Enhancements now enable its use on a sub-tree up to a specified depth, instead of a full geometry or full sub-tree.

3.2 Propagation in Field

Charged particles in Geant4 are tracked in external electromagnetic fields, and the intersection of their curved trajectory with geometry boundaries is approximated to a user specified precision. A 'global' field is the default for the entire setup, and can be assigned to the world volume. Other volumes can override this, to obtain another field that can be varying or constant, even zero.

A field, its solver and accuracy parameters are grouped into a Field-Manager. A curved trajectory is approximated as a set of chords. Key parameters in propagation are the maximal sagita and the required accuracy of integration. A user can specialize these accuracy parameters for each Field-Manager, and thus potentially globally or even for each volume or field. The user can also select, using a track's properties, the accuracy parameters for tracks of particular importance or unimportance. This enables a user, for example, to track precisely all muons or electrons with energy above 100 MeV, or to treat tracks in a calorimeter more coarsely.

3.3 Runs and Events

The Run Manager configures a Geant4 run, bringing together a particle source, a geometry and a set of physics processes. A re-factoring of the Run Manager module has separated that functionality, which is necessary for the kernel, from that which is convenient. This refinement enables an advanced user to create a customized run-manager, which is compact and easy to maintain.

In order to improve the link between a primary particle, its pre-assigned decay products, the trajectories of all resulting tracks and their associated hits, new hooks for user 'helper' classes have been created. These optional 'helpers' can carry user information for a primary vertex, a primary particle, an event or a region. Previously only primary particles, which are of a particle type defined in Geant4 and listed in the current physics list, are converted into tracks. With release 7.0 any primary particles exotic to Geant4 (e.g. weak bosons, Higgs, SUSY particles) can become tracks that result to trajectories and be identified as ancestors of tracks that deposit hits in sensitive detectors.

A new module, the General Particle Source, provides for the ability to generate primary particles sampling from a different spectra, integral and differential.

In a large setup, with hundreds of materials, the time for initializing process tables can run into minutes. This can makes interactive, investigative use (when a user changes setups significantly) very slow. This is exacerbated in large HEP experiments that adopt “Cuts per region”, since each process table’s entries can be as large as the product of the numbers of materials and regions. A new option saves these tables or vectors in files, to be read in at a subsequent run for a significant speed up in initialization. The first implementation retrieved tables only from a setup that is an exact match. A new scheme extends this, by handling the physics vector for each material and production threshold individually. This significantly reduces the recalculation, when few materials or regions are added or modified on a later run even in a large-scale application.

4 OPTIONS FOR PERFORMANCE ENHANCEMENT

Several different options are available to obtain better computing performance. The use of thresholds for the production of secondary particles, at the process level and optionally enforced by the kernel is widely used. Fast simulation via the parameterization of showers was available as a framework option, with the user providing the creation of hits. It is now enhanced with a first ‘shower package’ for EM showering in a homogeneous volume. In addition options for event biasing are provided to address relevant use cases.

4.1 Production thresholds for detector regions

Production thresholds for secondary particles are expressed in length (i.e. minimum range of secondary). This threshold is connected with the accuracy of position a user requires for energy deposition. Up to version 6.0, a Geant4 particle type had a unique production threshold (‘cut’), which was applied for all volumes. A different cut value was possible for each particle type. Yet this is ill-suited to a large detector, where the relevant length scales vary greatly between its regions. For example, a vertex detector requires spatial resolution of the order of a micron, whereas the structural supports tolerate thresholds of order of 1 cm. Having a unique threshold required the user to choose the lowest threshold needed for the most precise positioning detector device, and thus caused a performance penalty.

To address requirements from large-scale HEP experiments, we created the ability to define the production thresholds in a geometrical region in Geant4 version 6.1. A “per region” threshold can be common for all particle types or for each particle type separately. The thresholds are applied only when producing secondary gammas, electrons and positrons. Only processes, which have an infrared divergence or produce a large number of low energy secondaries, must respect this threshold.

A region can be any set of geometry volumes, and is typically a full sub-tree of a hierarchy, for example a detector sub-system. Volumes that are not physically connected can belong to the same region: a barrel and two end-cap calorimeters. Even a set of “deep” areas of support structures can be defined as a region. For each region, the user can define dedicated production thresholds for gammas, electrons and positrons. Daughter volumes at all levels will belong to a region (and share the same production thresholds), unless a daughter volume itself is assigned to another region. A default region and default thresholds are automatically created for the world volume. This enables existing or new programs without regions to be used without modification.

Test using the same threshold values in several regions showed no speed penalty compared with a unique default region with the same geometry. And experiments optimizing the regions and thresholds measured substantial performance improvements in their full detector simulations [13].

4.2 Event biasing

The toolkit has been enabling advanced developers to utilize event biasing in Geant4 applications using user code to vary the track weight since Geant4 0.0 (1998). But since Geant4 release 4.1 (June 2002), general-purpose biasing methods are available in the toolkit [12]. A module provides importance biasing, with splitting and Russian roulette tracks which go through boundaries of different importance. An importance value is associated with each volume. Either the 'mass' geometry or a dedicated 'parallel' geometry can be used for assigning the importance. An implementation of the weight-window method and a 'weight-cutoff' method are now available. Leading particle and cross-section biasing are provided for processes using the hadronic framework.

5 REMARKS

Continuous checking is undertaken of computing time to monitor CPU performance. Benchmark applications include simple setups, test beam and a use case with a complex magnetic field.

The interaction with Geant4 users is providing very valuable feedback. A new method for discussing key issues with users has been instituted in the past years: the Geant4 Technical Forum is open to all interested parties and individuals, and meets quarterly to discuss technical matters, including identifying issues and weighing priorities. Developers continue to emphasize identifying problems, and providing assistance to users in using Geant4 for established and new use cases.

6 ACKNOWLEDGMENTS

Many past Geant4 collaborators [14] have also contributed to the toolkit and developments mentioned, to the support and maintenance of the Geant4 toolkit. We express also our appreciation for the wealth and depth of feedback provided by Geant4 users, which was and is indispensable for improving the toolkit and in validating it for diverse use cases.

7 REFERENCES

1. Geant4 Collaboration (S. Agostinelli et al), Nucl. Instr. and Meth. in Phys. Res. A 506 (2003), 250.
2. http://cern.ch/geant4/
3. P. V. Degtyarenko, M. V. Kossov, and H.P. Wellisch, Eur. Phys. J. A **8**, 217-222 (2000).
4. http://aida.freehep.org/index.thtml
5. G. Folger, V. Ivanchenko and J.P. Wellisch, Eur. Phys. J. A 21 (2004), 407.

6. A. Heikkinen et al in *CHEP 2003 proceedings*, La Jolla, CA.
7. J. Tuli, *"Evaluated Nuclear Structure Data File,"* BNL-NCS-51655-Rev87, 1987.
8. http://cern.ch/gdml
9. S. Tanaka, K. Hashimoto and Y Sawada, *CHEP'98*, Chicago.
10. http://www.ge.infn.it/geant4/workshop2001/slides/arce.pdf
11. See the Geant4 Application Developer's Guide at, http://cern.ch/geant4/G4UsersDocuments/Overview/html/index.html, Sect 4.1.11, D Williams, G Cosmo.
12. M Dressel, CERN preprint CERN-OPEN-2003-048.
13. http://www.puhep.res.in/cms/indiacms04/indiacms_OCT04/Sunanda/OSCAR_2004.pdf
14. http://cern.ch/geant4/organisation/organisation.html