Simulation Work Plan for 2024

Alberto Ribon
CERN EP-SFT

On behalf of the Simulation team

CERN EP-SFT group meeting, 15 January 2024

Outline

- Highlights of the Simulation achievements in 2023
  - Part of these have been included in the latest Geant4 release, 11.2 (December 2023)
- Program of Work for 2024
  - Focus on HEP domain and EP-SFT team
    - The full version will be presented to the users in a dedicated Geant4 Technical Forum (in March/April 2024)

Notes:

- *A large fraction of the work goes to support, and regular testing & validation, which do not appear in the work program !*
- *Names are present only in the achievements – in grey for those not by EP-SFT*

Project Purpose

- Provide support for the Geant4 **detector simulations** of the CERN experiments and projects
  - According to the Lab's priorities (LHC, fixed-target experiments, future colliders, *etc.*)
- Contribute to the development and maintenance of the **Geant4 toolkit**
  - Large, international and multi-science effort

Our Approach

- Improving the accuracy of physics models, and extending their coverage
  - Based on thin-target data
  - Note: high-fidelity simulations are always needed (regardless of their speed)!
- Improving the computer performance of simulations
  - By reviewing implementations, algorithms and approaches
    - Without compromising on the physics accuracy
  - By using fast simulation solutions, *i.e.* simplifications with reduced accuracy
    - Many approximations are possible, matching different requirements and use-cases
    - Traditionally, shower parameterisations; now, growing interest and applications of ML-based solutions
  - By exploiting new technologies
    - *E.g.* GPU accelerators

Early Adoption of New Geant4 Versions

- We aim to reduce the gap between the most recent available version of Geant4 and the actual versions used in production by experiments
  - Because of the effort (for developers) of supporting multiple versions of Geant4, and waste (for users) of potential improvements

- To achieve this goal, there are two main ways:
  - Geant4 developers directly involved in the simulation framework of the experiments
    - To regularly perform integration tests and promptly following up eventual issues
  - Validation tests exported from the experiments to the Geant4 testing & validation suite, *geant-val*
    - To anticipate problems and work out possible solutions/fixes during the development, *i.e.* before a Geant4 release

1st Part

Highlights of Achievements in 2023
(part of these are included in G4 11.2)

Geometry

- Major progress towards a new surface-based GPU model in VecGeom *(A. Gheata)*
  - Contributions also by *J.G. Caminero and E.G. Stan*
  - Implemented all the features required by particle transport, for a subset of solids
- VecGeom : code simplification, improved portability *(A. Gheata, J. Hahnfeld)*
- New Quantum State Simulation (QSS) integration method *(R. Castro, L. Santi, A. Mignanelli, J. Apostolakis)*
  - Alternative method for transporting charged particles in magnetic field, potentially faster
    - Built-in interpolation capability and faster finding of the intersection of the trajectory with surfaces
  - Included in G4 11.2
- Refined control of very long steps (typically in vacuum) *(J. Apostolakis)*
  - Included in G4 11.2
- Simplification of the implementation of touchables *(G. Cosmo)*
  - Code optimisation: removal of unused specialisations and inheritance
  - Included in G4 11.2

Kernel

- First prototype of task-based sub-event level parallelism *(M. Asai)*
  - Event split in sub-events, with automatic merging of the hits at the end of the event
    - In “Phase I” : all tasks have the same physics processes and see the same detector geometry
      - Infrastructure included in G4 11.2, not yet functional
    - In “Phase II” : each task has only the necessary physics processes and sees the limited detector geometry which are needed for that particular task
      - Essential for heterogeneous simulation

- Feasibility study on parallelisation of initialisation stage *(V. Ivanchenko)*
  - Made progress on the robustness of the physics initialisation
    - Included in G4 11.2
  - Identified areas where parallelisation in the initialisation stage can be achieved

- Update of particle properties to PDG-2023 data *(S. Okada)*
  - Included in G4 11.2

Electromagnetic Physics

- Electromagnetic physics in crystals (channelling) *(A. Sytov)*
  - Implemented via fast simulation interface
  - Included in G4 11.2
- X-ray surface reflection process *(H. Burkhardt)*
  - Needed for applications in accelerator physics (*e.g.* FCC-ee) and space science
  - Included in G4 11.2
- Technical refinements in bremsstrahlung *(V. Ivanchenko)*
  - To fix configuration settings
  - Included in G4 11.2
- More robust and cleaner configuration of ionisation classes *(V. Ivanchenko)*
  - For proton, alpha and ions of different energies
  - Included in G4 11.2

Hadronic Physics

- Progress on the ATLAS energy resolution problem *(L. Pezzotti, D. Konstantinov, A. Ribon)*
  - Close agreement between hadronic showers of Fluka and Geant4
    - This seems to point towards an unknown problem in the simulation of the test-beam set-ups...
  - Found a tune of FTF string model that recovers a good agreement with ATLAS test-beam data; included in G4 11.2
    - Applied by default only in the physics list used by ATLAS (FTFP_BERT_ATL)

- New hadronic example *FlukaCern* *(G. Hugo)*
  - Showing how to use the new interface between Geant4 and (Fortran) Fluka-Cern; included in G4 11.2

- Extension of INCLXX model to antiproton annihilations *(D. Zharenov)*
  - Important for the CERN antiproton experiments; included in G4 11.2

- Several developments in nuclear physics (all included in G4 11.2)
  - Improved Fermi Break-Up model *(V. Ivanchenko)*
  - New LightQMD model, relevant for medical physics *(Y. Sato et al.)*
  - Implemented Doppler Broadening Rejection Correction (DBRC) for a more accurate modeling of low-energy neutron elastic resonant scattering in heavy nuclei *(M. Zmeskal, L. Thulliez)*

- Revised charge-exchange process *(T.L. Chau, V. Ivanchenko)*
  - Not used by default, but can be activated on top of any physics list; requested by NA64; included in G4 11.2

geant-val.cern.ch

- Essential web application tool for testing & validation of Geant4 *(D. Konstantinov)*
  - Constant effort to improve, maintain, run and use it
- Added first FNAL validation test in *geant-val* *(L. Pezzotti, J. Yarba)*
  - Lepto-nuclear test (*test75*)
  - Other tests are expected to be added directly by the FNAL team
- Run existing test-beam set-ups with the new *FlukaCern* interface *(L. Pezzotti)*
  - Important comparison between Fluka and Geant4 for the simulation of hadronic showers
- Added ATLAS LAr barrel EM calorimeter test-beam in *geant-val* *(L. Pezzotti)*
  - Old test, without real data, but very useful for regression testing and CPU performance studies of Geant4 electromagnetic physics

Fast Simulation

- Core work on *Par04* and MetaHEP *(D. Salamani, A. Zaborowska)*
  and new ML architectures (transformer) *(P. Raikwar)*
  - Collaborations with LHCb (MetaHEP), Openlab & IBM (new ML architectures)
  - Work done on the Geant4 example *Par04* included in G4 11.2

- ATLAS-related work
  - Validation of the new fast-sim (Atlfast-3, used in Run 3) *(D. Salamani et al.)*
  - Co-developing a VAE for low-energy pions (256 MeV - 16 GeV) *(D. Salamani et al.)*
  - Inference with ONXXruntime *(D. Salamani et al.)*
  - On-going work to make experiment-independent the classical parameterisation used by ATLAS *(A. Zaborowska*, co-supervising a EP-RD fellow, *J. Beirer)*

- Community work
  - Co-organisation of the CaloChallenge *(D. Salamani, A. Zaborowska)*
    - First calorimeter simulation for ML challenge
  - Open Data Detector *(A. Zaborowska et al.)*
    - Benchmark detector for algorithmic studies

AdePT (Accelerator demonstrator of electromagnetic Particle Transport)

- New surface model *(A. Gheata)*
  - Based on bounded surfaces; header library implementation for GPU performance & portability
  - Major task, took most of the development effort
  - Targeting a realistic complex set-up test (*e.g.* CMS) for 2024
- New test tools for benchmarking and validation *(J.G. Caminero)*
- Integration with experiments *(J.G. Caminero, A. Gheata, J. Hahnfeld, W. Pokorski)*
  - Test AdePT in more complex set-ups (geometry, particles input/output, *etc.*); study the requirements for realistic sensitive detectors
    - Standalone ATLAS TileCal test-beam *(D. Costanzo, S. Lachnit, L. Pezzotti)*
    - Integrated G4HepEm on CPU in CMSSW; discussion on required simulation output in HGCal (*J. Hahnfeld*)
    - Started LHCb Gaussino integration *(J.B. Benavides, W. Pokorski)*
- Geant4 assessment of AdePT & Celeritas on December 13th - 14th
  - Check-point of these projects; more synergies between these two projects
    - Assessment report expected for January 19th

Other Activities

- Replacement of ATLAS EMEC custom solid with Geant4 solids *(E. Tcherniaev)*
  - Useful to speed up the simulation, and to extract the GDML geometry of the entire ATLAS detector (which then can be used in stand-alone applications, as we are doing with CMS)

- Review of the interface between Geant4 and the ATLAS framework *(A. Dell'Acqua)*
  - The current interface was made a long ago, with a pure sequential Geant4. Fresh look with modern, task-based Geant4. Possible synergies with Gaussino.

- Geant4.jl : Julia interface to Geant4 *(P. Mato)*
  - Evaluate interoperability of the Julia language with existing large C++ libraries in HEP
  - Exploit Julia language to provide a simple and ergonomic user interface to Geant4
  - Status: complete functional prototype, already useful for small simulations and training events

- Differentiable programming in HEP detector simulations *(M. Aehle, M. Novak)*
  - Useful for gradient-based detector optimisations
  - Created *HepEmShow* package; one paper in preparation

Geant4 Tutorials and Schools

- Two CERN training courses per year
  - Geant4 Beginner Course
    - In Spring, run by M. Novak
  - Geant4 Advanced Course
    - In Autumn, run by several SFT members
      (J. Apostolakis, G. Cosmo, V. Ivanchenko, A. Ribon, A. Zaborowska, + a few other non-SFT collaborators)

- Simulation course at the ESIPAP school
  - European School of Instrumentation in Particle and Astroparticle Physics, Archamps
  - Run by A. Zaborwska, once per year (in February)

- Participation to some Geant4 courses outside CERN
  - *E.g.* in South Korea, by L. Pezzotti (November 2022 and 2023)

2nd Part

Work Plan for 2024

Geometry, Field and Transportation

- Completion and validation of the ATLAS EMEC implementation with Geant4 solids
- Reduce geometry initialisation time (voxelisation) using optional multi-threading
  - Requested by CMS
- Investigate use of multi-threading to speed up overlap checking and volume calculation
  (if enabled)
- Explore additional opportunities for memory and speed optimisation
- Investigate alternative implementations of navigation history
- Optimisation of QSS field stepper, enabling QSS3

Kernel and Software Management

- Complete "Phase I" of task-based sub-event level parallelism
  - Infrastructure already included in G4 11.2, but not yet functional
    - Reminder: "Phase I" means that all tasks have the same physics processes and see the same detector geometry

  Note: "Phase II" is expected after G4 11.3
    - Reminder: "Phase II" means that each task has only the necessary physics processes and sees the limited detector geometry which are needed for that particular task

- Modularisation of Geant4 libraries
  - Identify libraries and modules for merging or splitting, including optional modules that a user may choose to drop or add to the build

Electromagnetic Physics

- Optimise initialisation, and reduce initialisation time by using optional multi-threading
  - Requested by CMS
- Implement positron annihilation into 3 gammas
  - Relevant for medical physics, but might impact also HEP analyses
- G4HepEm *Specialised EM physics library for electron, gamma and positron for HEP (CPU & GPUs)*
  - Extension to provide the full functionality of native G4 EM physics
    - Configuration per detector region (*e.g.* multiple scattering for CMS)
    - Gamma / electron / positron – nuclear interactions
  - Additional features aimed to gain extra computing performances
    - Data re-structuring (*e.g.* of the macroscopic cross sections)
    - “General Process”-like handling of macroscopic cross sections
    - Woodcock Tracking for gamma particles
  - Integration and validation in ATLAS and CMS software frameworks
  - Maintain the GPU support of the library for AdePT

Hadronic Physics

- Review of the diffraction dissociation treatment in the Fritiof (FTF) string model
  - Aimed to improve the simulation of thin-target data; impact on hadronic showers to be carefully evaluated
- New hadronic datasets
  - Consistent, in particular in the treatment of nuclear levels with incomplete information, and with fewer (ideally without) unphysical nuclear levels
    - This should solved several open bugs in Geant4 hadronics
- Continue the code review of ParticleHP
  - Started last year, and the first part included in G4 11.2. Goal is code maintainability & CPU performance
    - Required by ATLAS for radiation background studies
- Long-term : **Fluka-Cern** C++ rewriting of the hadronic physics, compatible with Geant4
  - Fluka hadronic physics will be available in a dedicated physics list
  - Useful for developers to cross-check hadronic model implementations; useful for users for the evaluation of systematic errors (*e.g.* HL-LHC physics analyses)

geant-val

- Keep adding new tests (*e.g.* CMS HGCal), including those used by experiments to validate their migration to new Geant4 versions
  - This helps the developers to discover in advance unexpected changes and problems, reducing the effort and the delay on the users for moving to newer Geant4 releases
  - Validation against experimental data is the best, but even pure regression tests are very useful
- Review all tests, homogenise their structure, and collect them into a single place
  - Major, collective, shared effort which will take a few years

Fast Simulation

The ML-related work items will be integrated into the new ML activity

- Develop transformer-based ML models
  - Establish the best single-geometry diffusion model
  - Work on inference optimisation
  - Extend to different geometries and test adaptation capabilities, measure savings on training time
- Experiment-specific work (in collaboration with members of the experiments)
  - LHCb
    - Find the best working model for hadronic showers (possibly a transformer-based model)
  - ATLAS
    - New Fellow (Peter Mckeown) will continue the work of D. Salamani on ML for ATLAS, implementing a data structure that allows to test VAE and transformer-based models
    - Co-supervise work of J. Beirer on FastCaloSimV2-based classical shower simulation
  - CMS
    - Implement data production sample with structure that allows to test transformer-based models on HGCal
- Others
  - Speed-up simulation of oriented crystals detector
  - Community efforts : CaloChallenge and Open Data Detector

AdePT (Accelerator demonstrator of electromagnetic Particle Transport)

The work-items will be integrated with the assessment recommendations

- Enable simple integration of GPU R&D prototypes in experiments
  - Provide one interface for integration of GPU extensions in experiments' Geant4 applications and frameworks (common between AdePT and Celeritas, in collaboration)

- Optimisation of the workflow and components; experiment-specific optimisations
  - In particular for scoring, *i.e.* for the treatment of hits in sensitive detectors

- VecGeom
  - Completion of the surface-based geometry modeler targeting GPU acceleration
    - Implementation of conversion to the surface representation for the missing solids
    - Test it across challenging LHC experiment geometries
    - Implementation of the BVH-based acceleration for the surface model
  - Release of version 2 – portable and surface-aware
    - Refactoring the current CUDA implementation behind a common portability layer
    - Further simplification of the geometry interfaces

Team (as of January 2024)

- Staffs: J. Apostolakis, G. Cosmo, A. Dell'Acqua, A. Gheata, M. Novak, W. Pokorski, A. Ribon, A. Zaborowska

- Fellows:

  S. Diederichs, L. Pezzotti

  - SFT fellows; J. Hahnfeld contributed till the end of 2023

  J.G. Caminero, P. Raikwar

  - EP R&D fellows; D. Salamani contributed till the end of 2023; P. Mckeown will start in March 2024

- Important contributions from associated and visitors:

  D. Konstantinov, V. Ivanchenko, E. Tcherniaev

  - Permanently at CERN

  A. Bagulya, A. Galoyan, O. Kadri, V. Grichine, V. Uzhinsky

  - Regular visitors