HEP software community meeting on GeantV R&D
Panel Report

Gloria Corti, Peter Hristov, Chris Jones, David Lange, Tom LeCompte, Zach Marshall, Pere Mato (ex-officio), Martin Ritter, Andreas Salzburger, Torre Wenaus (ex-officio), Dennis Wright, Avi Yagil

November 7, 2016

Contents

Executive Summary

The “HEP software community meeting on GeantV R&D” took place Oct 25-27 2016 at CERN. The open meeting, organized by the HEP Software Foundation (HSF) at the request of the GeantV Project, was the first exercising of the HSF’s mandate to organize (at project request) peer reviews of important common software projects. The GeantV effort has reached the point at which community feedback is important to assess progress and as input to future directions, and they accordingly proposed this meeting. As described by the GeantV team:

> The goal of the GeantV project is to develop a high performance detector simulation system integrating fast and full simulation that fully exploits advanced features of modern CPU architectures, including CPU accelerators. After more than three years of R&D the project has produced a prototype capable of transporting tracks in complex geometries exploiting micro-parallelism, SIMD and multithreading. It is therefore timely to organise an assessment of the project, to assess the progress that has been made in addressing the original goals, and to demonstrate that the required performance improvements can be met for the simulation of complex events in complex detector geometries such as those found at the LHC. One important byproduct of the assessment should be to inform the community more widely of the work that is being done and for this reason everyone interested in the topic is encouraged to attend and to participate actively in the event.

A panel of ten experts from the community was assembled and provided with a charge to guide their assessment. The panel has produced this report offering findings and recommendations as advice to the GeantV team to help it improve its work and set the future direction.

The panel benefited from carefully prepared, comprehensive slides and presentations, and thanks all involved in their preparation. Slides and discussion did not fit within the allotted agenda, but we were all complicit in that with the extensive discussions that took place. More than 40 people were in the meeting room, and about 10 connected remotely. The meeting was very successful in promoting dialogue between developers and users.

The panelists found the meeting to be very successful and productive. We learned a lot and commend the GeantV team on an effort that has accomplished a great deal. With this meeting the GeantV activity and the HSF have taken the commendable step of engaging the community in planning where the GeantV activity goes in the future.

This document is structured following (primarily) the charge. The scope of each section is *summarized in italics*. Findings and recommendations address the charge points (which ***are shown in the relevant section***) as well as other topics identified by the panel. Findings are significant observations and inferences. Recommendations suggest a particular action or direction, often in light of a finding.

As an executive summary, a brief summation of our responses to the charge points is included here. The detailed assessment constituting the basis for the summation is given in the following sections.

Charge points and summary response:

1) ***The degree to which the project has demonstrated it can meet its goals in a realistic application (realistic events in a realistic detector)***

The activity has delivered several components for use in Geant4 and expects to continue this integration in the near to medium timescale (2017-2019). These include VecGeom geometry and new EM physics, both targeted for inclusion in Geant4 10.3 with release in December 2016. In the assessment of the panel, experiment migration to a fully validated GeantV vectorized track transport framework will not be a realistic prospect before LS3 and Run 4 (roughly 2026). Concerning meeting the goals of performance increase with respect Geant4 and the seamless integration of fast and full simulation, the panel was presented with insufficient information to conclude whether the activity will be able to meet them.

2) ***The ability of the new track transport framework to fully exploit parallelisation in all its forms and to deliver a substantial performance gain as compared to the existing simulation engine, Geant4***

Performance gains have been demonstrated in geometry basketization, but overall gains and cost/benefit cannot be assessed until the framework is more complete, including realistic physics and a more complete handling of MC truth.

3) ***The suitability of the new track transport framework to perform fast and full simulation with the possibility of mixing the two modes in different regions of the geometrical setup and/or for different particle types and momenta***

The presentations and materials did not address plans and designs for fast simulation sufficiently to be evaluated.

4) ***The impact of any changes to the implementation of physics models that may be needed in order to achieve the stated improvements in software performance***

Integration of physics into the transport framework is at too early a stage to assess how physics implementations will impact framework performance. In EM physics, unit tests of new codes for EM processes show speed-up factors of 3-7, but it is not yet established how this will scale to an application. It is recognized that hadronic processes pose a large challenge to vectorization. New physics model implementations can benefit from the years of C++ evolution, experience and modern libraries developed since their Geant4 forbearers were introduced, thereby bringing some performance improvements irrespective of vectorization.

5) ***The credibility of the project plan (deliverables, timeline, and human resources) for delivering a new simulation toolkit on a timescale that matches the schedule of the experiments allowing for the time that will be needed for integrating user code and validating physics performance***

The panel was presented with high-level objectives and an effort-level-dependent timeline, but not with a overall project plan. The panel makes recommendations towards establishing project planning in areas that are developing concrete deliverables to the experiments, and establishing realistic timescales for the deliverables. One important deliverable to define the interface that GeantV will provide to experimental frameworks and user applications. It is critical that the user interactions, for example for sensitive detectors, are discussed with stakeholders as early as possible.

6) ***Assess the extent to which opportunities to leverage GeantV work in the near term are being sought and exploited, for example by integrating GeantV-directed geometry improvements in Geant4***

The panel was pleased to see the degree to which such opportunities are being exploited. These include VecGeom and new EM physics integration into Geant4, and (as clarified in the meeting) integration of Geant4 hadronics in GeantV (which can be expected to bring improvements to the hadronics code) as a first step in delivering GeantV hadronics.

7) ***Assess whether the most important risks have been identified and whether adequate attention is being given to addressing them***

Identified risks and their mitigation approaches include:

- Deliverable risk. R&D inherently has the risk of failing at achieving the targeted objectives. Risk is mitigated through a component model, spreading the risk among components of varying inherent risk levels and thereby making it likely at least a subset will be successfully delivered.
- Schedule risk. The panel judged the schedule risk of delivering a migration-capable GeantV toolkit for Run 3 as very high, and recommends deferring this objective to LS3 and Run 4 (roughly 2026). For shorter term objectives the projectization recommended by the panel would contribute to tracking and managing schedule risk. The panel makes recommendations towards less aggressive near- and mid-term milestones which would also lessen schedule and manpower risk.
- Manpower risk. Funding and the availability of expertise are substantial risks towards attainable milestones. Projectization, careful attention to building career paths, and engaging active participation of the experiments by offering valued near/mid-term deliverables are mitigation directions.

8) ***Assess the project's approach to ensuring experiment framework compatibility, ease of migration to GeantV, and interoperation of GeantV with other toolkits (particularly Geant4).***

Overall and in many specific respects the activity's approaches towards supporting compatibility, migration and interoperation are excellent. The activity has adopted a strategy of delivering modular components independently usable and with their own individual release schedules. Where possible these components will be usable via Geant4, thereby supporting a clear path to operation within and validation against the experiments' simulation codes. The GeantV and Geant4 physics validation repositories (databases) are common, an excellent development that will allow much easier validation of new physics models, new versions of GeantV, and new code optimizations. While many of these approaches are good, there remains plenty of work to do before the cost of migration can be determined, e.g. the interfaces are for the most part not yet in place.

General

Findings and recommendations not associated with a particular charge item or section below.

Findings

1. The panel benefited from carefully prepared, comprehensive slides and presentations. The panel thanks all involved in their preparation. Slides and discussion did not fit within the allocated agenda but we were all complicit in that with the extensive discussions that took place.
2. The panel found the meeting to be very successful and productive. We learned a lot and commend the GeantV team on an effort that has accomplished a great deal. With this meeting the GeantV activity and the HSF have taken the commendable step of engaging the community in planning where the GeantV activity goes in the future.
3. The agenda did not entirely reflect the guidance provided to the panel, in that areas of emphasis came late in the agenda and so were compressed in the discussion, in particular the detailed architectural presentation at the end, whereas topics not emphasized in the guidance such as physics came earlier.
4. There appear to be developments in many different directions, but it is not entirely clear how they come together.
5. Among these many developments, it was not clear to us what the activity's priorities are and to what degree defined priorities drive the work program.
6. The panel was charged to assess the GeantV effort as an R&D activity. As such it is producing substantial value to the community through components that in the near term will provide new capability through integration with Geant4 for example.

Recommendations

1. In the judgement of the panel the activity has reached a stage where it should transition beyond exclusively R&D to aim at concrete deliverables for the benefit of the simulation community. Many of the comments in this report are in the direction of formalizing the project in this way. Such a 'projectization' could mean scrutinizing the whole activity to prioritize and select what is carried forward and in what form. The scrutiny may result in deciding some activities should be carried forward as formalized deliverables, others should remain at an R&D level, and/or some activities should be stopped and the resources switched to more productive directions.
2. The panel believes that use of GeantV developed codes in production simulation prior to Run 4 will derive primarily from contributions reaching the experiments through Geant4. The GeantV and Geant4 programs in these areas will then necessarily be interdependent, and GeantV and Geant4 planning should accordingly proceed in close concert as these areas are projectized in GeantV. Simulation development effort is a scarce commodity in terms of both expertise and funding that should be used according to agreed priorities.
3. A natural question arising after this successful meeting is the need for a follow-up. The panel believes there should be one, perhaps in 18 months or so, which addresses a list of deliverables set in advance. Given the tight coupling between GeantV and Geant4, some thought should be given to how to structure and scope a follow-up. In particular, the panel believes that if there is a follow-up meeting it should include participation from Geant4 (participation that goes beyond attendance and the overlap between those two groups).
4. Given the inevitable time pressures of such a meeting, it is suggested for the future that the agenda better reflect the emphases and priorities of the guidance.
5. The panel endorses event throughput as the definitive bottom line metric for evaluating performance and added value.
6. It is gratifying and important that a talented new generation is coming into simulation development. Career support is vital. The activity should aggressively pursue such support within the institutes and with funding agencies.

7. The group is engaged in both developing a new toolkit that may offer a significant performance gain, and in implementing features, models, and optimizations in existing code. As such, it is important that the performance is always judged by placing side-by-side the throughput of a fixed Geant4 MT version (10.2 is an acceptable choice); the latest Geant4 MT version including all the GeantV group's developments that have been ported; a scalar version of GeantV; and a vectorized version of GeantV. These comparisons should be prioritized to focus on 'standard' silicon, as the potential gain on GPUs, while interesting, will require significant effort both to achieve and to realize for the experiments (including, potentially, significant investment in the acquisition of new hardware). The throughput should be tested on a dedicated machine similar to those used in production by the major experiments.
8. The panel recommends that GeantV take it as a general project guideline that wherever possible components be developed such that they are usable in Geant4 as well as GeantV. As well as contributing early value to the community, this will facilitate the integration and validation of GeantV simulations in the experiments, and will draw valuable early feedback from the experiments. The approach is already having success in areas like geometry and EM physics, and should be extended. The answers to our questions contained a concise expression of such a guideline: GeantV in the final picture as "a set of physics, geometry and service modules common to GeantV and Geant4 (and beyond)."
9. Delivering capability and improved event throughput to the experiments through the integration of GeantV components in the experiments' production Geant4 simulations is a GeantV activity success and should be regarded as such. Delivering for Geant4 and today's production simulations is not "working against ourselves", a statement (perhaps ironic) heard in the review, it is a win for all.
10. The panel endorses the group's expressed desire to increase the dialog with the Geant4 collaboration, for example to coordinate effort in developing common components (especially common 'new'/revised physics models), and to better leverage a broad range of experience. The panel agrees that the development of modular components that can be used in GeantV and Geant4 (and possibly beyond) will naturally foster the collaboration with the Geant4 community.
11. The panel agrees that (as expressed in the answer to a question) GeantV has to be a toolkit, as contrasted to an application, or a more monolithic "logical toolkit" like Geant4. GeantV should be a "physical toolkit" allowing individual modules to be independently released and used.
12. A dictionary of terms and concepts (in particular relating GeantV concepts to the familiar Geant4 classes) would be useful to avoid ambiguity and confusion on e.g. what is meant by a hit. In some talks this was provided and was very useful.
13. As a guideline in developing random number generation strategies, event reproducibility is an essential requirement. This can have important performance implications.

Project goals

The degree to which the project has demonstrated it can meet its goals in a realistic application (realistic events in a realistic detector)

The goals as presented to the committee before the meeting were:

1. Develop a detector simulation framework with a multithreaded transport engine that exploits data locality and explicit SIMD vectorization.
2. Design the code from the ground up for maximum performance on modern computing architectures. The code will be portable to different CPUs and accelerators such as (but not limited to) GPGPUs and Xeon Phi machines.
3. Start either from Geant4 physics models and improve them in performance and accuracy or develop entirely new models, whatever is the most appropriate.

4. Achieve computing time speedup factor of 3-5 on complex (LHC-size) HEP detectors for improved physics performance with respect to Geant4. Investigate and understand the limitations to reach a factor of 10.
5. Integrate full and fast simulation capabilities in a single detector simulation framework.

These goals were not fully reflected in the program of talks put before the panel. Nonetheless, a consistent set of goals is critical to driving the program forward.

Findings

1. A factor of 10 speed up is desired but 2-5 is considered likely by the team. This goal is judged to be realistic by the panel if the factor is to be evaluated with respect to the current version of Geant4. One aspect of the team's goals is to understand the limitations to reach a factor of 10 speed improvement. The only discussion in this line was in the context of GPU accelerators.
2. The team targets the beginning of LS2 for the start of commissioning of a GeantV toolkit by experiments. Specifically, a complete toolkit release with equivalent or better physics demonstrated for a LHC-sized problem by the end of 2018 is a major milestone.
3. Migration of the LHC experiments' production simulations to GeantV on the timescale of Run 3 start is not a realistic objective in the judgement of the panel. Such a migration could only be realistically considered for LS3 and Run 4 (roughly 2026). Nevertheless, having the interfaces clearly defined by Run 3 would not only be a worthy goal but a critical step towards this migration.
4. The integration of full and fast simulation was listed among the goals of the program, but was almost unaddressed by the talks.

Recommendations

1. Detailed goals should include estimated GeantV-to-Geant4 ratios for the various major components of GeantV simulation. These estimates should then be produced for several examples of HEP detector configurations.
2. Define realistic near- and mid-term project goals for bringing capability to the experiments. Consider defining LHC Run 3 project objectives and corresponding timeline and deliverables focused on contributing to an N-fold improvement in Geant4 performance through the integration of GeantV project components in the experiments' production simulations.
3. If the close integration of fast simulation is to be a significant selling-point of GeantV, the collaboration must invest significantly more effort into designing the interfaces to fast simulation. The GeantV group organized a successful workshop to collect requirements from experiments. The panel recommends additional effort in reviewing what the experiments have done in this regard while designing these interfaces.
4. A priority for the activity's various components should be to get interfaces out soon and into the hands of people who will use them. Cultivate early participation in prototyping and feedback.
5. The collaboration is currently working on a great number of diverse areas, including different accelerator applications, NUMA architecture, and genetic algorithms to optimize basketization. While these are interesting efforts, work should be focused on the primary goals of the group. This is true whether the primary goals are constructing a full-fledged vectorized prototype or developing a large number of back-portable new models and modules for use in both Geant4 and GeantV.
6. The panel encourages defining as a priority goal getting a 'full' simulation running early (e.g. using Geant4 physics models) in order to get early feedback from experiments.

Architecture, geometry and navigation

Structure, concurrency, interfaces, geometry, navigation, transportation, MC truth, fast sim

The ability of the new track transport framework to fully exploit parallelisation in all its forms and to deliver a substantial performance gain as compared to the existing simulation engine, Geant4

The suitability of the new track transport framework to perform fast and full simulation with the possibility of mixing the two modes in different regions of the geometrical setup and/or for different particle types and momenta

An impressive array of concepts and prototypes has been presented where many parts (geometry, navigation, geometry basketization) seem to be in a good state. However, a combination of all these concepts into one coherent tool was either not presented clearly or is still in its early stages.

Findings

1. The team has demonstrated a working implementation of a simulation system based on the basketization scheme. At the moment the system only basketizes based on geometry, but the team showed plans to also test basketization of physics processes. They have shown excellent utilization of vector systems (on KNL they reached a peak throughput when processing around a 100 tracks and on GPUs hit a peak at 10,000 tracks). The scaling limit with increasing number of threads was not given.
2. The backbone of the geometry implementation, VecGeom, uses VecCore for the vectorized math and algebra operations. VecCore has three back-ends for the actual implementation: scalar, Vc and UME::SIMD. An important aspect of the development of VecCore/VecGeom is currently on supporting track baskets, i.e. vectorization of various simulation needed distance calculations as well as vectorized ray based navigation.
3. No detailed concept for fast simulation was presented. The absence of a detailed plan for fast simulation integration in both presentations and milestones raised concern amongst the panel members. It's difficult to judge the feasibility of the intended approach at the current state. A filtering mechanism for feeding into fast simulation modules has been suggested, but not further explained. It is unclear how their design for fast simulation integration expands beyond the current functionality available in Geant4.
4. The specialized geometry/navigation mechanism recompiles the geometry description into concrete code. While this sounds promising, it is not clear if this would allow for small displacements (e.g. misalignment) without recompilation or if this would require a new library for every change to the geometry. This could also have an effect on deployment as precompiled binaries for all possible geometries need to be shipped or a full JIT system needs to be present.
5. The track transport implemented as a Runge-Kutta integrator is based on (an earlier version of) VecCore, with vectorization on track level in mind. Intrinsic vectorization of the transport equation is also possible and it was demonstrated to achieve reasonable speedup when implemented (in ATLAS) in the past. Caching some part of the current field cell information is under investigation.
6. The design for handling user defined hits involves the use of data factories. The factories produce hit blocks where there is a hit block per hit type per thread per event. The hit blocks can therefore be updated safely since no threads are sharing the same block. Once GeantV finishes an event, the planned system will merge the relevant hit blocks for the event automatically. It is unclear at that point if the design of the merging step is already done and how user code gains access to the final merged container of hits. There is some concern that the merging across threads could become a serialization point.

7. The MCTruth manager only described handling of the track tree and procedures for cutting parts of it away. No indication was given how to handle hits if the track creating them is omitted, or how to handle relations between hits and track information in general. The back navigation of mother-daughter relations that may be processed in different threads has not been demonstrated nor a clear concept presented.
8. The MCTruth manager interface seems to require the user code to be completely multi-threading safe as different threads/tasks might process tracks from the same event concurrently. It's not clear if there will be any additional services or interfaces to help with this effort.
9. So far GeantV doesn't support pre-assigned decays and treats all primary tracks equally. While the team is confident that this feature should not pose any problems no detailed plans were presented.
10. User interfaces (e.g. scoring, user actions) seem to not be defined yet.
11. No design for general abort handling of events or parts of events that are spread over different threads or processes was shown (it was mentioned that loopers are caught, but no further demonstration was shown).
12. The ongoing effort of fully debugging the transport in Geant4 took on the order of ten years.

Recommendations

1. The architectural design appears to still be in flux. It is recommended to formalize the definition of the main components, including the interfaces between them and towards the user code, taking into account all the known use cases. Integration with the experiment frameworks needs to be considered as part of the user interface design.
2. A common use case for hits is to not keep all hits that occur in a sensitive detector and instead only keep a 'sum' of the hits. A standard mechanism within GeantV should be developed to support this case.
3. Both the MCTruth and the hits are the components of the system for which users must plug in their own code. At the moment the proposed interfaces are quite different. It is recommended that the team review these two use cases and form common concepts which then can be used to develop user interfaces which reuse those concepts. This will lower the barrier of entry for the users as learning how to handle hits can carry over to learning about MCTruth.
4. In general, MCTruth and hit handling should be handled in unison to show that hits can be associated to the correct (ancestor) track information even if some tracks are omitted from the output. This might introduce a bottleneck (e.g. if remapping on event closeout is needed) or might not cover all use cases (e.g. decide upon keeping a track at the end of its life). While most experiments will implement these parts themselves a fully coherent MCTruth handling including hit mapping has to be demonstrated.
5. The interfaces for fast simulation have not been developed, which makes understanding the integration of fast simulations into the full framework very difficult. In particular, significantly more thought needs to be invested into the interfaces for fast simulations that will write hits in a similar way (and, ideally, into the same hit collections) as the full simulation.
6. Given that the number of cores is projected to continue to increase, it is important to demonstrate that the GeantV design is capable of scaling to the number of cores expected in the medium and long term.
7. In a multiple-track transport module, a clear strategy of handling navigation failures/faults of one of the tracks in a parallel processed basket has to exist.

Physics

The impact of any changes to the implementation of physics models that may be needed in order to achieve the stated improvements in software performance

Some of the speed-up of GeantV must come from optimized, vectorized physics models. Physics quality should not be compromised in this speed-up. Numerous branches in physics code, especially in hadronics, pose a major challenge to vectorized implementations. Because EM processes dominate showers and have relatively fewer branches they are the first targets of optimization and vectorization. Low energy neutron and elastic hadron scattering processes are also more amenable to vectorization. There is a concern that tabulated hadronic physics or libraries of pre-generated events will compromise physics performance.

Findings

1. A large amount of work has been proposed to select likely physics code, review it on theoretical and implementation bases, vectorize it, validate it and integrate it into GeantV. Their estimate for the remaining work to implement a first version of the EM physics is of order 1.5 FTE-years, and the work to implement the first version of hadronic physics they estimate at 10.5 FTE-years. The team expects that at least double this amount of effort would be required to bring the EM physics and hadronic physics to comparable levels of diversity when compared to Geant4. The work to wrap Geant4 hadronic physics described as 'Plan B' is not in common with the re-implementation and has been described as requiring at least two years, which should be viewed as in addition to the 10.5 FTE-years already required for hadronics.
2. Significant improvements in physics processes have resulted from code review. These include the re-written Goudsmit-Saunderson multiple scattering model and improvements in LPM suppression in bremsstrahlung. These improvements have been propagated to Geant4 and will appear in version 10.3.
3. Due to the dominance of EM processes in showers, priority has been given to the review and optimization of these processes. Among these, e+/e-/gamma processes will be developed first, to be followed by muons, hadrons, etc.
4. Unit tests of EM processes show speed-up factors of 3-7. It is unclear how this will scale to an application.
5. The hadronic interface is in its early design stage. Many of the interface changes indicated may also be useful in Geant4.
6. Hadronic processes pose a large challenge to vectorization, leaving elastic scattering and low energy neutrons as the best candidates for early development. Other processes currently rely on a large amount of branching.
7. A modest amount of speed-up may result from wrapping existing Geant4 processes so that optimized CLHEP methods can be used.
8. Some additional speed-up may result from developing a hadronic framework which is more streamlined compared with that in Geant4. Templating may replace virtual inheritance when more code is completed.
9. Tabulated physics will aid in vectorization and speedup, although this was regarded as only a stopgap until better models become available.
10. Most pieces of transportation in field are vectorizable; field look-up is the critical speed issue.
11. The panel views very positively the proposed direct collaboration between the GeantV activity group and the Geant4 team at SLAC on a GPU-enabled neutron physics library, interoperable in GeantV and Geant4.
12. The GeantV and Geant4 physics validation repositories (databases) are common. This is an excellent development that will allow much easier validation of new physics models, new versions of GeantV, and new code optimizations.

13. A new tool was presented for extracting neutron cross sections for use in GeantV and Geant4. This tool could replace several older tools and provide a common source for cross section information in the simulations.

Recommendations

1. The ability to integrate new GeantV physics models inside of Geant4 was viewed by the panel as absolutely critical for the validation of these physics models and optimizations. It is vital that this ability is maintained, even if the price is additional development time.
2. The proposed goal of complete EM physics by 2017 seems too ambitious given the size of the job and the available effort, and given the other priorities of the effort. More FTEs would be required to meet the goal and it is unclear that this would be the best use of additional effort.
3. Having complete hadronics by 2018 is even more optimistic. Geisha should be considered as an early vectorized hadronic physics model.
4. Consider hooking up to Pythia8 to generate hadronic interactions.
5. Integration of Geant4 hadronics via wrappers was presented as a 'Plan B', with Plan A being the long-term development of new hadronics codes. It is the impression of the panel that the availability of Geant4 hadronics in GeantV will be crucial for the experiments' physics validation of GeantV. Accordingly, 'Plan B' should be the first priority. It has the further advantage of putting a hadronics-capable GeantV into the hands of experiments as soon as possible. The panel agrees that the Plan B extraction from Geant4 will necessarily be accompanied by a revision of the interfaces and a substantial modularization and simplification of the code that could be expected to bring value to Geant4 as well.

Infrastructure and development

Development model and infrastructure, I/O, libraries, testing, validation, performance, metrics

Support software, including VecCore, I/O, testing and validation suites, and documentation will be essential in the development of GeantV. Project management tools such as Git, JIRA, and Jenkins are currently in use.

Findings

1. The activity has shown a well-established development model with best-practice tools and procedures for code integration, validation and testing.
2. VecCore/VecGeom define their back-end on compilation time and will create code which will only run if the instruction set (AVX, AVX2) is available on the client computer. This poses an issue for software distribution which is of course not unique to GeantV.
3. Current testing of GeantV seems to be limited to running a successful example job. Regression testing using unit tests and a web-service-based test suite is being worked on, but no clear time frame has been presented.
4. The implementation makes heavy use of templates. This is needed to maximize performance. However, it was mentioned that compilation of the code can require considerable amounts of memory on the machine and has caused one compiler to crash.
5. There are some minor issues to be addressed as the activity is exposed to a wider community.
    a. The source code of GeantV and VecGeom is available in GitLab repositories hosted at CERN. The VecGeom repository has a well prepared set of installation instructions, while the GeantV

repository is missing a README. A link to the main webpage of GeantV should be provided in both repositories.
   b. The installation instructions from GeantV are probably sufficient for developers, but may cause difficulties for new users.
   c. The described workflow on the website is very basic and differs from what was presented at the meeting. The coding standards are well established, however the description is long and difficult for new contributors.
   d. A code checker (ECLAIR?) was mentioned. It is not clear at which stage of the process this is used and if such checking is a prerequisite for merge requests/code integration, and/or part of the merge request review.
   e. The panel was not able to access the link to the VecGeom project in JIRA from the GeantV web page.
   f. "The best Geant ever" statement that appears on the main GeantV page still has to be demonstrated.
   g. A convention has been set up for branches with username/JIRA project/JIRA ticket. It is not clear when more than one person is working on common development if they each have their own branch and how they share their code.
   h. The Doxygen documentation for GeantV and VecGeom should be linked to the main Web page.
6. Regarding I/O it was unclear to the panel how events are being kept in memory and how they are streamed out. The panel asked for a diagram to be provided which shows the proposed structure of GeantV output, tree entries and branches, and the workflow of events moving from their memory representation to storage. The response provided by the team did not provide sufficient clarification.
7. The panel appreciated the benchmarks that were provided. In some cases, it was difficult to understand what the reference of the comparison was, and this made it difficult to judge the performance improvement provided by the new code. The recommendation above to standardize benchmarks would help to address this concern.

Recommendations

1. A common effort to investigate appropriate distribution systems for instruction-set specific binaries, or preparation of fat binaries that contain different back-ends and select the appropriate one on startup, is needed. As this is not GeantV-specific it should be done at a higher level, possibly under the aegis of the HSF and related to the packaging and project template efforts.
2. Integration tests showing that track transport is working correctly for various conditions (e.g. different geometries, magnetic fields, etc.) need to be implemented. The tests should include checking that tracks are properly assigned to an event after the event has been closed.
3. A transition from "prototyping" to "product development" with extensive unit and regression testing as well as validation is recommended in order to keep the schedule.
4. The panel did not hear of a formalized code review process for new physics algorithms. This should be established, ensuring that unit tests and physics validation tests are included as the models are developed, rather than post-facto.
5. Benchmarks should clearly separate the effects of new algorithms, code design concepts, vectorization, concurrency and accelerators. The comparison should be done with respect to a fixed Geant4 version and also with respect to the latest version (which may already contain improvements provided by GeantV).
6. In order to do more extensive integration testing, CPU-burners could be implemented to take the place of components that have not yet been fully developed. This would allow more complete testing of, for example, the basketization system including both geometry and physics basketizing, and could allow an earlier confrontation of the question of the scaling (in CPU and memory) of the full system.

Project planning

Deliverables, timeline, manpower, consistency with experiment schedules/plans/priorities, project risks & mitigation

The credibility of the project plan (deliverables, timeline, and human resources) for delivering a new simulation toolkit on a timescale that matches the schedule of the experiments allowing for the time that will be needed for integrating user code and validating physics performance

Assess whether the most important risks have been identified and whether adequate attention is being given to addressing them

An impressive array of concepts and prototypes was presented to the panel. Many aspects of the research and development (including geometry, navigation, basketization) individually seem to be well-developed. Most of these have a schedule, or at least an estimated completion date. However, there was no overarching plan shown that incorporated these subprojects. Some subprojects, such as hadronic physics, have insufficient effort to meet a release date two years from now.

The proponents have identified areas of technical risk. At least as important is the issue of schedule risk, and what actions need to be taken if one or more sub-projects will not meet the release dates or the experiments' integration deadlines.

Findings

1. An integrated plan including deliverables, timelines and effort was not shown. Some of the sub-projects had one or more of these elements, but there was not a single overarching plan presented to the panel that incorporated all of them.
2. The panel did not get a clear sense of the overall priorities and how they come together to form a complete prototype from the material presented. (The factor of 2-5x speedup is a goal, not a priority.)
3. A set of milestones were provided in the Q&A. The milestones seem aggressive, and the early milestones do not always seem to lead towards an early public release.
4. The emphasis on CUDA, IBM Power, and other emerging architectures seems higher than needed to reach the initial GeantV release goals.
5. The primary risks discussed involved the necessary levels of effort either not materializing or proving insufficient. The effect of this will be to delay the project. The effect of such delays on the experiments is not clear but can be mitigated substantially through continued integration of R&D from GeantV into the Geant4 toolkit where possible.
6. Much of the effort needed to execute the project is highly specialized. Certain tasks and sub-projects rely on expertise that may be unavailable or difficult to supplement.
7. Project planning should clearly delineate R&D and production-directed activity. The latter requires a detailed plan and timeline to measure progress towards deliverables.
8. A manpower estimate was provided: the total number of FTEs working on GeantV is 12.

Recommendations

1. Produce a credible project plan that includes deliverables, milestones and external deadlines (such as when the large LHC experiments will need given deliverables). This plan should incorporate the interrelationships between the GeantV sub-projects and external milestones (such as alpha and beta releases) perhaps organized within a project management toolkit.

2. Prioritize the deliverables for the “in two years” release and potentially future releases. With help from the experimental community, determine which ones will, if delayed, force a delay in the release, and which can be de-scoped to a later release.
3. In the spirit of prioritization, if an R&D activity is not producing the desired results on a given time frame, the project should seriously consider reassigning resources to other activities that have demonstrated success and require more effort to deliver production quality components.
4. This plan should be realistic (i.e. achievable with the available effort). It may be valuable to consider alternative plans under scenarios with more or less effort available.
5. Identify projects for collaboration of potential interest to the experiments and Geant4 collaborators not currently involved with GeantV that can attract participation to well defined tasks and increase the level of effort.
6. A prominent open issue is the cost/benefit of the basketization and vectorization approach to concurrency and thus event throughput. As a matter of risk mitigation it is encouraged to define goals for the project that are not all contingent upon the timely success of this development path. GeantV components and deliverables have already been demonstrated to have independent utility. They should continue to be developed, able to be independently adopted by the simulation community to broaden the paths to success for the project.
7. As commented in the answers to questions, an important means of alleviating project risk is early joint efforts between the project and experiments to create good interfaces and to examine all potential issues in depth. This requires early direct engagement with the experiments, which is recommended (and again below in the context of growing the effort pool).

Community integration

GeantV as common project(s)/modules, leveraging GeantV work outside GeantV proper, relation to Geant4, experiment framework compatibility, experiment software migration/integration, interoperation and experiment physics validation

Assess the extent to which opportunities to leverage GeantV work in the near term are being sought and exploited, for example by integrating GeantV-directed geometry improvements in Geant4

Assess the project's approach to ensuring experiment framework compatibility, ease of migration to GeantV, and inter-operation of GeantV with other toolkits (particularly Geant4).

Extensive work has been placed into making GeantV developments available in Geant4 and ROOT. Geometry and EM Physics developments allow for extensive checks of these features even without a full-fledged GeantV prototype containing all parts.

Findings

1. While there is some design for integration with task-based experiment frameworks, no indication was given on framework compatibility and migration efforts.
2. From the presentations it seems that large parts of the user interface will have to be re-implemented by the user (scoring, actions, field lookup, MC truth handling). This will probably lead to a high adoption barrier. It was not clear if and to what extent GeantV will offer services or interfaces to simplify migration and implementation of these items in a performant manner.

3. The design for integration with task-based (particularly Intel TBB) experiment frameworks was shown. Upon further review, the design was found to be insufficient. Consulting with an expert in the design of task-based HEP frameworks has lead to a modified design which needs to be tested.
4. In particular from the I/O discussion, it seemed that there is some ambiguity in vision regarding the use of GeantV as an application or as a toolkit. A strict binding of the I/O to a particular back-end technology should be avoided.
5. GeantV has contributed optimized physics models to Geant4 (see physics section) and expects to make more such contributions in the future.

Recommendations

1. GeantV has implemented its own set of lock-free multi-producer/multi-consumer queues. The actual benefits were not discussed in detail but it might be worthwhile to extract these into a separate library for concurrency utilities or feed the development back into an already existing library.
2. Several modules developed as a part of GeantV have the potential to have a broader community-wide impact with potential usage in event digitization, reconstruction, and analysis. The developers should consider whether such extensions ought to be part of the design of these tools, as these considerations may have significant implications. Examples where such expansion is encouraged include: VecGeom/VecCore, random number generation, field access, propagation (including Runge-Kutta transport), and geometry.
3. It should be ensured that the work is globally accessible. One excellent possibility is that the team follow the ROOT model: master git repository in GitLab, but public forked GitHub hosted repository which is automatically updated to match the master.
4. As much of the user interface is not yet designed it would be beneficial to reach out to the community for discussion on what interfaces are actually needed the most, and how to define them to benefit all experiments as early as possible.
5. The inclusion inside of future versions of Geant4 of a 'GeantV physics list' would significantly ease the validation of the GeantV physics modules. Such a physics list should be considered in collaboration with Geant4.
6. The GeantV developers should endeavor to view the product as an enhanced toolkit rather than as an application. Items like I/O should be viewed as parts of examples to give to the user, and as ways to understand and flesh out the interfaces, rather than as critical components needed for the first prototype.
7. Engage with experiments and framework developers in prototyping and testing activities. Work with them to identify collaborators on their side.
8. Continued investment in close and continuous collaboration with CERN OpenLab could provide significant benefit and should be explored as a potential source of assistance in some of the difficult computing problems being addressed by the new infrastructure.
9. One of the large risks of the activity is fragmenting the simulation community further with an additional, independent product. It is important to ensure coherence between the GeantV and Geant4 teams and close integration of the code in order to minimize this risk.

Links to other materials

Description of GeantV including goals:
https://drive.google.com/file/d/0B3H5nTdCzYK3QVV2OWgxczdXQ0ZxaVY5RzhGX0JNQ1JmNGNv/view

General document explaining the review:
https://docs.google.com/document/d/104EPbRpOC6cqCtDF4tzHLE0KSrvlL2Ic_xexiDfYtO0/edit#

Review Agenda:
https://indico.cern.ch/event/570876/

Overnight questions and answers:
https://docs.google.com/document/d/1rKxaj1g0pWhojZ4cb-Wc081Yv_SgIhcYoQ88wsN3AT0/edit?usp=sharing

GeantV website:
http://geant.cern.ch/

HSF website:
http://hepsoftwarefoundation.org/