Draft Geant4 Assessment Report of the GPU R&D projects AdePT & Celeritas

Alberto Ribon

CERN EP-SFT

AF meeting, 1 February 2024

Disclaimer

- I am presenting *my* highlights of a ***draft*** version of the Geant4 assessment report of the GPU R&D projects AdePT and Celeritas

    - The final, official version will be ready in the coming days

Purpose

- The aim of the assessment is to check the state of the GPU R&D projects AdePT and Celeritas – **achievements**, **findings** and **difficulties** – and to make **suggestions** and **recommendations**

Assessment Process & Panel

- Presentations on 13/14 December 2023 at CERN

  - One week before, written documents have been provided

  - Open meeting, Zoom connection available, with speakers and most of the assessment panel physically present at CERN

- Assessment panel:

  - Makoto Asai (Geant4, EIC, Jlab)

  - John Chapman (ATLAS, CERN/Cambridge University)

  - Gloria Corti (LHCb, CERN)

  - Ivana Hrivnacova (Geant4, ALICE, IJCLab)

  - Vincenzo Innocente (CMS, CERN)

  - Matti Kortelainen (CMS, Fermilab)

  - Andrea Valassi (CERN IT)

  - Marc Verderi (Geant4, LLR) - chair

  - Sandro Wenzel (ALICE, CERN)

  - Ben Wynne (ATLAS, CERN)

4

General Considerations

- *"The panel congratulates the two projects for the impressive amount of work and investigations done. It has appreciated the open discussions and the constructive exchanges, including those on the issues and difficulties met by the projects."*

- *"The panel considers that **no key items have been overlooked**. Both projects showed they are well aware of the issues, of the developments and investigations to be carried on, and of the volume of effort to address these."*

- *"The projects are capable of LHC-scale detector simulations and **no show-stoppers have been identified today**, for what physics, performances or scalability are concerned."*

  - Note: so far, this is based on test simulations, not yet on production simulations

  "The projects still have important tasks to address, in particular on the surface-based geometry approach, its memory consumption and scalability to large structures."

Findings

- Physics
  - "*The **EM physics** is **not** at present **an issue** in terms of processing on GPUs*"
    - *"Despite its relatively complexity and stochastic nature"*
- Geometry
  - "***Geometry** is the **main limiting factor** today for computing performance*"
    - Due to the different complexity of the various 3D solids, causing thread divergence
- Scoring – *i.e.* handling of hits in sensitive detectors
  - "*The CPU-GPU communication is not the bottleneck it was expected to be. This is particularly interesting for the hit treatment... [that] may involve complex users' code, which adaptation or rewrite on the device may be difficult…* ***Sending back data from the GPU to the CPU to build steps for processing hits*** *with existing experiments' code comes with **low overhead**.*"
    - *"This will help trying out the GPU prototypes in experiments' frameworks"*

Common Recommendations

- Common activities, without yet merging

  - "*The panel does recommend continuing discussions and exchanges between the two projects about findings, difficulties and solutions. The panel does not recommend however merging – at least for the time being – the activities of the two projects into a single one.*"

- Hadronic particles in EM showers

  - Send back from GPU to CPU hadrons produced by photo- and electro-nuclear interactions

- Offloading below ~1 GeV and extending to EM particles in hadronic showers

  - Try out this, as alternative to the porting of full EM showers to GPU (as done so far)

- Include production thresholds (also known as "cuts")

- Profiling

  - "*The panel encourages using profiling as it may be a powerful way to detect inefficiencies.*" (*E.g.* 3D solid model *vs.* surface model)

- Interest of Detray/ACTS for the geometry representation

  - Evaluate suitability/synergy with surface-based Detray geometry

Recommendations to **AdePT**

- Processing on CPU

  - ”*The addition of a CPU-only processing mode would greatly help the integration into experiments, as allowing CPU only and CPU-GPU comparisons.*”

- Intermediate milestones

  - "*The AdePT team has conducted a number of investigations and found its path among the difficulties. The panel feels however that intermediate stages and results could be more precisely documented and recommends that intermediate milestones be made more explicit.*"

Recommendations to **Celeritas**

- Fix multiple scattering

  - *"Celeritas has shown that their implementation of the multiple scattering has some flaw… The panel recommends this problem to be fixed to ensure that no serious difficulties are at play here."*

- End of track and end of event signals

  - *"Celeritas proceeds today by dumping all steps, regardless of their track and event origin, and by applying a post-processing to re-aggregate these steps into proper tracks and events... The panel recommends having end-of-track and end-of-event signals present in the simulation data to demonstrate there is no obstacle here."*

- Technicalities in the report

  - *"The report provided is nicely done, but generally presents things at a high level, making it difficult for the reader to appreciate the technical aspects and possible difficulties of the approaches and solutions adopted. The panel recommends Celeritas people to not refrain from providing low-level information and plots."*

Delta-assessment

- *"The panel proposes a **delta-assessment by December 2024**."*
  - *"The two projects have demonstrated the need to evolve in depth the geometry representation, shown as the main limiting factor today.*
    The two projects anticipates they will be ready by September.
    December leaves then some time to perform batteries of tests."
    - In October there will be also the CHEP conference and the Geant4 Collaboration meeting
  - *"For this delta-assessment, the panel considers proposing a series of application benchmarks together with a series of distributions of various key observables and performance measurements."*