Opticks/CaTS and Integration to Geant4 and LArSoft

Hans Wenzel

FERMILAB-SLIDES-24-02

Opticks/G4CXOpticks/CaTS: and Integration with LArSoft

Opticks is an open source project that accelerates optical photon simulation by integrating NVIDIA GPU ray tracing, accessed via NVIDIA OptiX. Developed by Simon Blyth: https://bitbucket.org/simoncblyth/opticks/

It was developed for the Juno experiment→ Bound to NVIDIA hardware and software.

CaTS (Calorimeter and Tracker Simulation) is a flexible and extend-able framework. With respect to Opticks it interfaces Geant4 user code with Opticks and defines a hybrid workflow where generation and tracing of optical photons is optionally offloaded to Opticks (GPU) using the G4CXOpticks interface, while Geant4(CPU) handles all other particles.

CaTS an advanced Geant4 example based on legacy version of Opticks (based on Optix 6) was included in Geant4 11.0:
https://geant4.kek.jp/lxr/source/examples/advanced/CaTS/ → not maintained, since Opticks based on Optix 6 is not maintained.

Status:

- Opticks has been completely reengineered by Simon Blyth migrating to OptiX7.
- The new Opticks (NVIDIA OptiX7) has been fully functional (off and on) since January 2024.
- New Opticks APIs have been tested and successfully integrated with a modified workflow of CaTS!
- https://github.com/hanswenzel/CaTS. Optimization, physics validation and benchmarking of the new Opticks are ongoing!

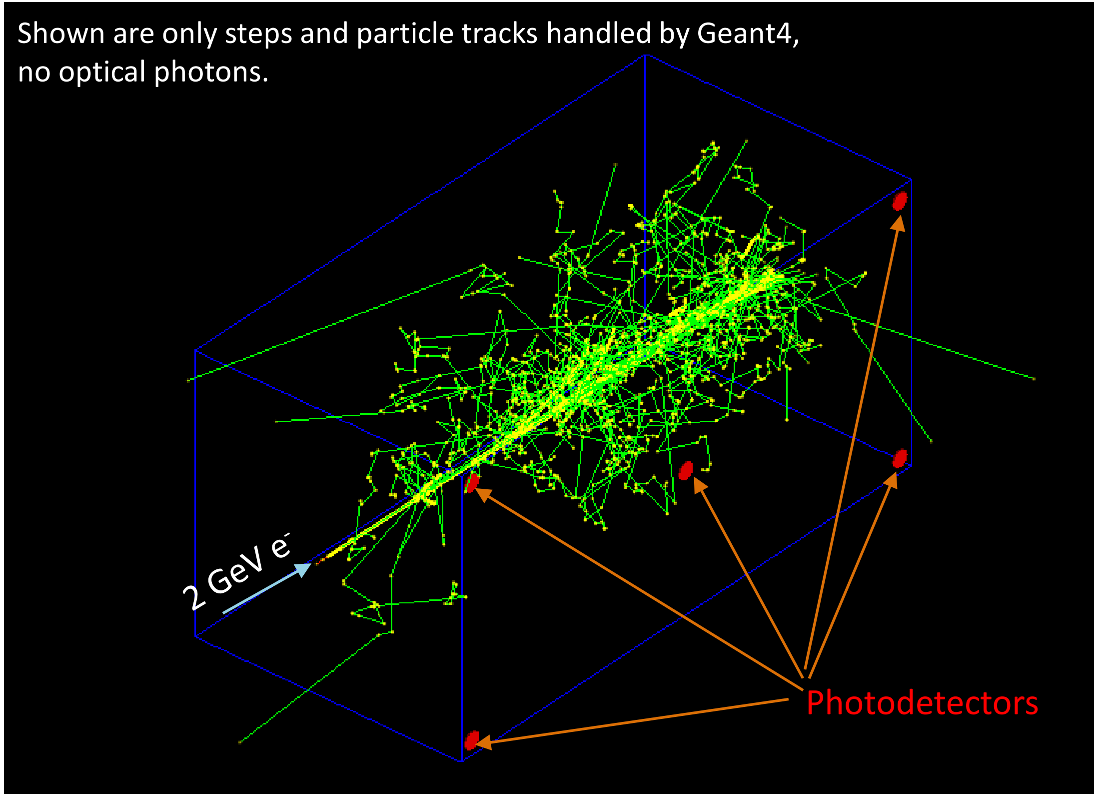
The computational challenge for TPCs based on liquid Argon (LArTPCs):

Test Detector Geometry:
Liquid Argon: x y z: 1 x 1 x 2 m (blue)
5 photo detectors (red)
photon yield (no E-field): 40000 γ/MeV With 1000V/cm e-field recombination factor ~ 0.7.
(low Z=18, low $\rho = 1.78$ g/cm^3).

- 2×10^6 photons are produced/event. For a single 2 GeV electron shower (not fully contained).
- Using Geant4 to simulate photon generation and propagation using a single core on an Intel® Core i9-10900k@ 3.7Ghz takes :

  ~ several minutes/event

(Compared to **0.034 seconds/event** without optical photon simulation) → LArTPC-Experiments use look up tables and parameterizations instead of full simulation for photon response.



The computational challenge for TPCs based on liquid Argon (LArTPCs) (cont.):

Contribution from Cerenkov Radiation is significant due to refractive index increasing in VUV and $1/\lambda^2$ dependency of Cerenkov spectrum. Cerenkov radiation even more relevant in Electromagnetic showers due to contribution from ultra relativistic light particles (e^+, e^-).

Minimum ionizing particle MIP 2 GeV/c μ:

18% of Photons from Cerenkov radiation,

82% Scintillation Photons.

2GeV electromagnetic Shower: $2x10^6$ photons.

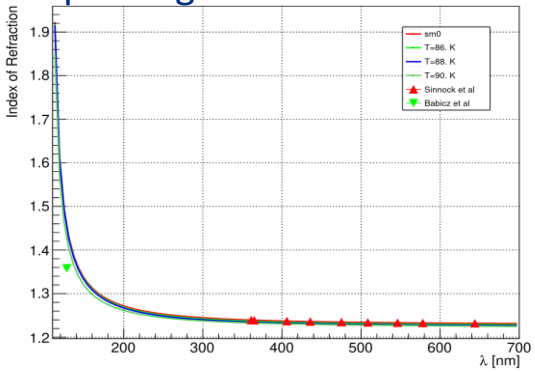
64% of Photons from Cerenkov radiation,

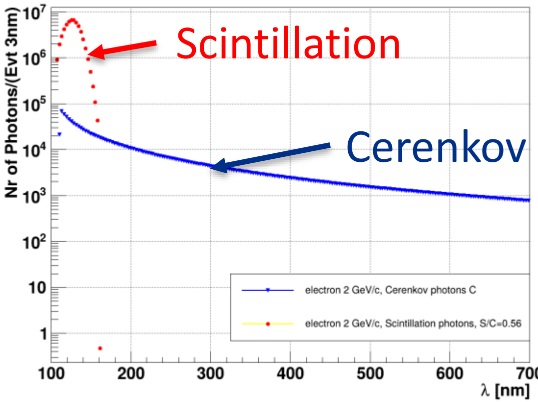
35% Scintillation Photons.

Currently the photons are mainly used to provide a t_0 for the electron drift and as trigger, but one can imagine to improve the energy resolution by combining the Ionization and optical signal or separating the Cerenkov and Scintillation signal might allow to apply dual read out corrections to improve the response of hadronic showers significantly!

→ in general, detailed simulation will allow to explore various ideas!!!

Liquid Argon: Refractive Index

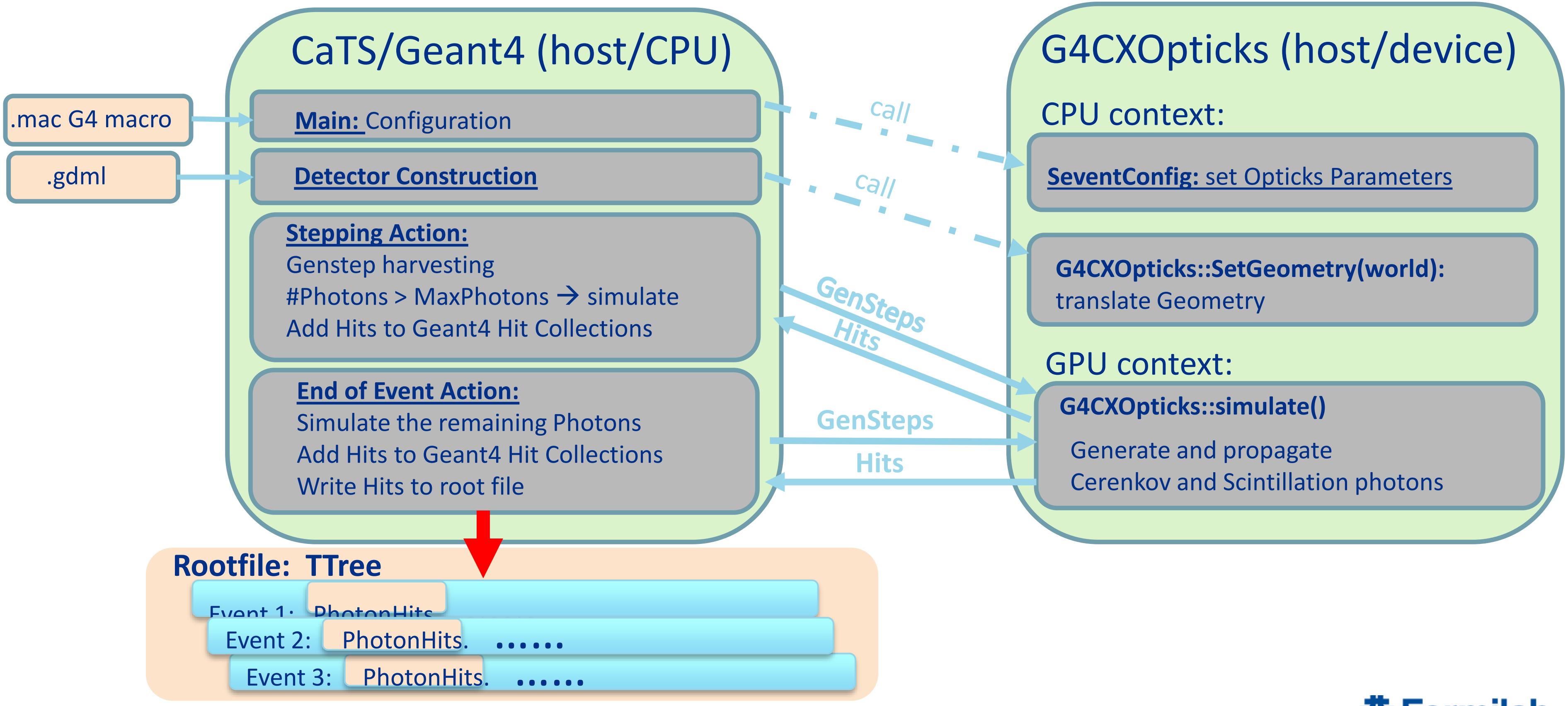




Simulation of optical photons: an ideal application to be ported to GPU's.

- **Massive parallelism:** Only one particle type is involved (optical photon), but many of them (~$2x10^6$/event). Only optical photons are produced.
- **Simple algorithms:** Only a few physics processes need to be implemented on the GPU. These processes are: to generate photons: G4Cerenkov), G4Scintillation (Reemission),
  to transport photons: G4OpAbsorption, G4OpRayleigh, G4OpBoundaryProcess, G4OpWLS (not yet implemented in Opticks yet but needed for LArTPCs).
- **Little data transfer from host to device and vice versa:**
  - GenSteps for the Cerenkov and Scintillation processes → host to device
  - PhotonHits → device to host.
- **Optical ray tracing is a well-established field:** benefit from available efficient algorithms (OptiX®).
- **Optimal use of NVIDIA® hardware and software:** (NVIDIA® CUDA, NVIDIA® OptiX ®, hardware accelerated raytracing RTX when available).

CaTS workflow using the new version of Opticks based on OptiX® 7:



Note!

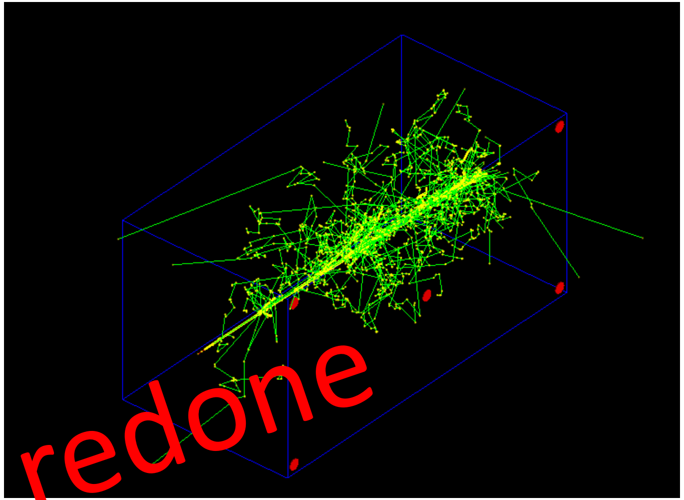
Note!

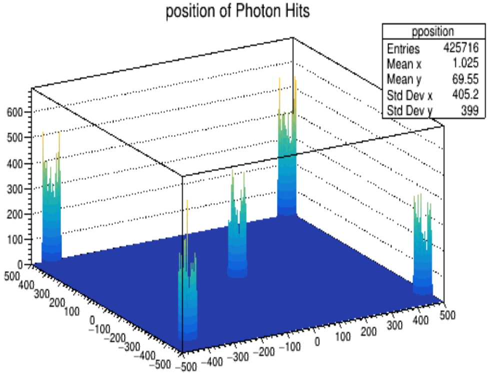
- Gensteps/Hits: provide a general interface between Geant4 and external photon generators/tracers. Not just Opticks! → make part of Geant4 API?
- Could the code for the various processes be rewritten so it applies to both CPUs and GPUs?

Performance: (Legacy Opticks, Needs to be redone)

Hardware:	
CPU	Intel® Core i9-10900k@ 3.7 GHz, 10 CPU cores
GPU	NVIDIA GeForce RTX 3090 @ 1.7 GHz, 10496 cores
Software:	
Geant4: 11.0, Opticks based on OptiX® 6	

To be redone





Number of CPU threads	Single threaded. Geant4 [sec/evt]	Opticks [sec/evt]	Gain/speed up
1	330	1.8	189x

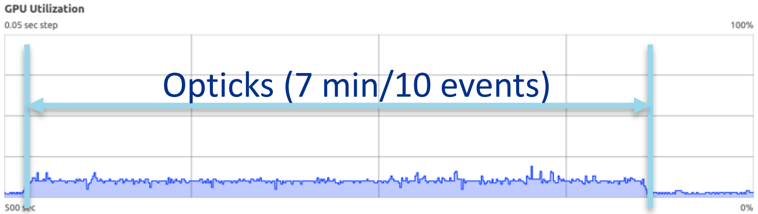
→It becomes feasible to run full optical simulation event by event! But comparison is to single threaded Geant4→ somehow unfair! Single geant4 threat can saturate the GPU and doesn't allow the use of multiple CPU cores.

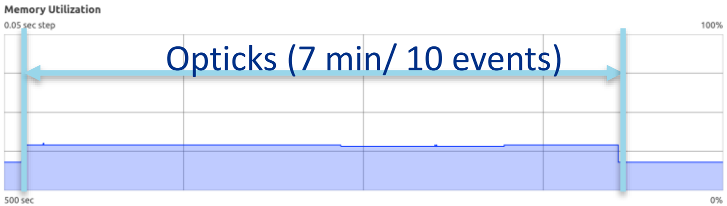
Status

The new Opticks (with NVIDIA OptiX7) has been fully functional (on/off) since January 2024.
New Opticks APIs are tested and successfully integrated with a modified workflow of CaTS (v2xx).
. https://github.com/hanswenzel/CaTS . Physics validation and benchmarking of the new Opticks are ongoing!

With the legacy version of Opticks (based on NVIDIA Optix 6) we achieved speed up in the order of a few times 10^2, this depends strongly on detector geometry, hardware and settings.
So far with the new version of Opticks (same computing hardware and detector geometry, 2 GeV e-shower) we found:
Compared to the previous version, the speed up by GPU over CPU is only a factor 10 times compared to several 10^2 times previously. We found the GPU (computing and memory) resources are underutilized. The graphs below show that only 10% of the CPU and memory resources are utilized → the optimization of the GPU kernel launch promises to improve the performance significantly.
The work integrating Opticks with the liquid argon TPC software framework (LArSoft) is ongoing.

To be redone





Integration to experimental frameworks/LArSoft

- artg4tk/larg4: are art/LArSoft modules. LArSoft: software package for liquid Argon TPC's (generation, simulation, response, reconstruction, analysis). LArSoft is based on the art event framework.
- artg4tk/larg4: are based on CaTS but:
  - Use art Event for persistency.
  - Geant4 UserActions are accessed via art services.
- Currently transition is ongoing to new software packaging (spack) and build tool (mpd).

Plans:

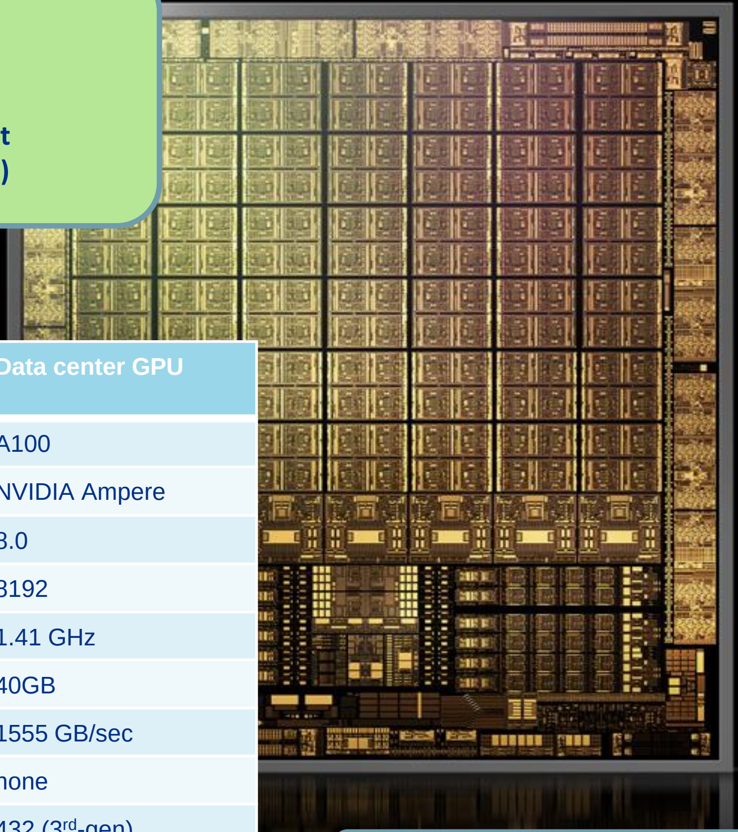
- Update the Geant4 advanced example CaTS to use the new Opticks. This will be part of the next release.
- Complete LArSoft integration.
- Provide detailed full-scale example of a liquid Argon Time projection Chamber (LArTPC) for optical simulation. (simple example exists).
- Provide detailed documentation of Geant4 optical physics processes and material properties (from literature) relevant to LArTPCs .

  "Liquid Argon optical properties for Geant4 Simulations":

  DUNE-doc-31579 → make the information available to the Geant4 code base.
- Enhance Opticks functionality (implement Wavelength shifting process).
- CaTS: Improve RootIO using Root TBufferMerger when running in multi-threaded mode

Extras

Opticks will only run on:
NVIDIA® hardware and NVIDIA® software
Software: NVIDIA® CUDA, OptiX
OptiX 6: allows to select/deselect RTX
OptiX 7: RTX cores are used when present
(RTX is not usually available on HPC systems)



	Graphics card	Data center GPU
	GeForce RTX 3090	A100
architecture	NVIDIA Ampere	NVIDIA Ampere
Compute capability	8.6	8.0
CUDA cores	10,496	8192
Boost Clock	1,7 GHz	1.41 GHz
Memory	24 GB	40GB
Memory bandwidth	936 GB/sec	1555 GB/sec
RT cores	82 (2^{nd}-gen)	none
Tensor cores	382 (3^{rd}-gen)	432 (3^{rd}-gen)
Shared Memory size	64kB	up to 164 kB

RT core: based on bounding volume hierarchy (BVH), a commonly used acceleration structure in ray tracing, ray-triangle intersection.