# Geant4 and Related Projects

Philippe Canal, Krzysztof Genser, Soon Yung Jun

CSAID Roadmap Meeting

March 7, 2024

# Outline

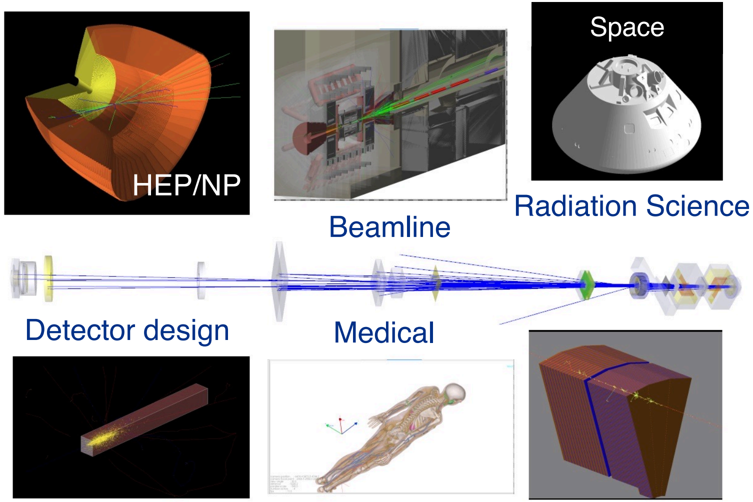
- Geant4
  - History and 2024 Census
  - Fermilab contributions, core activities and strategies
  - Funding outlook and sustainability
- Related R&D Projects focusing on GPUs
  - Optical Photon Simulation
  - Celeritas
- Plans

# Geant4: A Simulation Toolkit

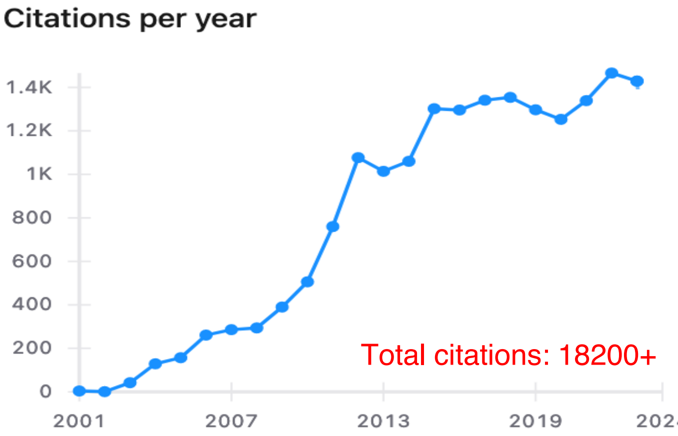
- 30 years of evolution and being the main simulation toolkit for HEP

| 1994 | 1998 | 2013 | 2021 | 2024 |
|---|---|---|---|---|
| CERN RD44 | V0.0 Geant4 | V10.0 Geant4 MT | V11.0 G4Tasking | |

- Where are Geant4 applications used?




S. Agostinelli *et al.*, NIM **A** 506 (2003) 250-303
J. Allison et al., IEEE Trans. Nucl. Sci. 53 (2006) 270-278
J. Allison et al., NIM **A** 835 (2016) 186-225



# Geant4 Collaboration

- 2024 Census: 144 members (N) from 24 countries: 33 FTE total

| Country | FTE (N) | Country | FTE (N) |
|---|---|---|---|
| CERN/Swiss | 8.6 (21) | Japan | 1.2 (11) |
| US | 5.5 (19) | UK | 1.0 (4) |
| France | 5.0 (23) | Spain | 0.5 (4) |
| Italy | 2.7 (13) | Australia | 0.4 (1) |
| Russia | 2.1 (5) | Portugal | 0.3 (2) |
| Korea | 1.4 (7) | Other | 4.7 (34) |

US: FNAL, JLab, LLNE, ORNL, SLAC and 7 other institutes
Other countries: Canada, Sweden, Denmark, Netherland, Greece, Hungary, Slovakia, Tunisia, South Africa, India, Vietnam, Argentina, Saudi Arabia

- Geant4 is organized by
  - Working Groups
  - Task Forces (R&D, Physics validation, …)
  - Steering Board
  - Oversight Board (Institutional/Regional)

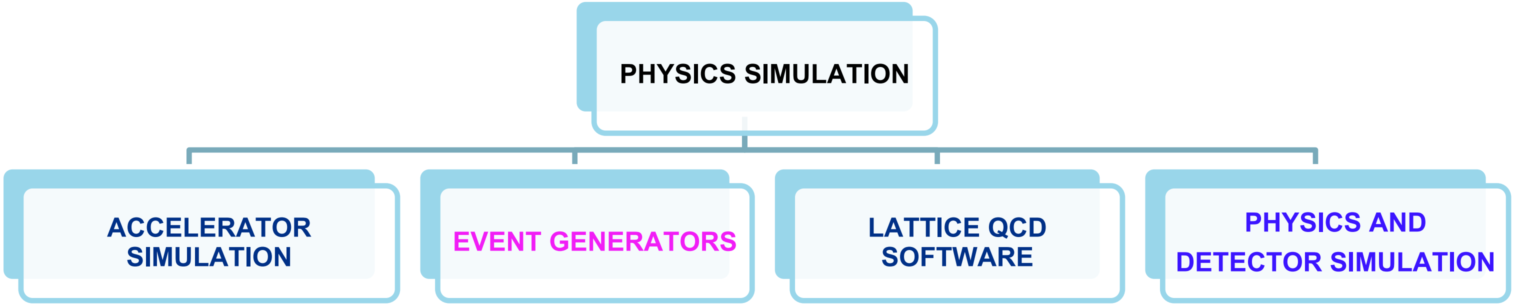
# Geant4@Fermilab

- Strong participation in many working groups (WG) for more than two decades
- 2024 Fermilab members and contributions: 8 members and 3.54 FTE including R&D (1.40)
  – Oversight Board (D. Elvira), Steering Board (S. Y. Jun) and User requirements (K. Genser)

| Working Group | FTE | Member |
|---|---|---|
| EM Physics | 0.63 | K. Genser, S.Y. Jun, H. Wenzel |
| Examples | 0.04 | H. Wenzel, J. Yarba |
| Hadron Physics | 0.68 | K. Genser, R. Hatcher, K. Lynch, S. Y. Jun, H. Wenzel, J. Yarba |
| Geometry | 0.10 | G. Lima |
| Physics Lists | 0.04 | K. Genser (Deputy), R. Hatcher, H. Wenzel, J. Yarba, |
| Testing and QA | 0.43 | P. Canal, K. Genser, S. Y. Jun (Coordinator), J. Yarba |
| R&D Task Force | 1.40 | P. Canal, S. Y. Jun, G. Lima |

# Geant4 members in Physics Simulation Department in DSSL Division

- Krzysztof Genser (Dept. Head), Philippe Canal (Assistant Dept. Head for Geant4 R&D), Soon Yung Jun (PDS Leader), Guilherme Lima, Julia Yarba, Hans Wenzel, Robert Hatcher, Kevin Lynch (Beam Div.)




- Core Geant4 activities at Fermilab and related R&D
  - Geant4: Enhance accuracy of physics and detector simulation and maximize throughputs
  - R&D efforts: Exploring and leveraging heterogeneity (GPUs)
  - Support experiments and users

# Current Activities of the Fermilab Geant4 Team

- Geant4
  - Coordinating the testing and quality assurance working group (S. Y. Jun)
  - Computing profiling and performance evaluation (using the Wilson cluster) (J. Yarba)
  - Hadronic physics validation and physics model parameters tuning (J. Yarba)
  - Management of user requirements (K. Genser)
  - Taking shift for system integration tests (merge requests and nightly tests) (S. Y. Jun)
  - Participation in working group activities (All)
- Geant4 related R&D focused on co-processors (GPUs)
  - CaTS/Opticks/OptiX integration (H. Wenzel)
  - Celeritas: development and integration (P. Canal, S.Y. Jun, G. Lima)
  - VecGeom2 (surface-based geometry model) (G. Lima)

# Support Fermilab Experiments and Users

- Consulting on optimal use of Geant4 (physics and computing)
- Geant4/VecGeom ups packaging (J. Yarba, G. Lima)
- Dedicated support/participation of Geant4 members in local experiments
  - Mu2e: K. Genser
    - Enabling transitions to new versions of Geant4 by proactive testing/debugging and absorption and implementation of locally developed needed features into Geant4 whenever needed/feasible (e.g., creation and monitoring of the ShieldingM (a variant of the Shielding) physics list used by Mu2e).
    - Advocating for Geant4 code changes/patches as needed and enabling the use of new features of Geant4 when beneficial.
    - Guidance on optimal use of Geant4 and code reviews; Dissemination of Geant4 knowledge.
    - Direct implementation of Geant4 related experiment's code requiring expert level Geant4 knowledge.
  - NOνA: R. Hatcher
    - Creation and maintaining AltPhysicsListFactory.

# Support Fermilab Experiments and Users (2)

- (continue) Dedicated support/participation of Geant4 members in local experiments
  - DUNE, proto-Dune, EMPHATIC, ICARUS, etc.: H. Wenzel
    - Maintaining LArG4(LArSoft interface to Geant4)/artg4tk and code reviews.
    - Changed LArG4/artg4tk so that the ionization potential of material can be changed and checked effect on energy deposits. Consulting migration from EdepSim to LArSoft/LArG4.
    - Worked on "Modeling Quasi-elastic Scattering in 1GeV/c $\pi$+ Liquid Argon Interactions with Geant4".
    - Consulting EMPHATIC (magnetitic field implementation), ICARUS (geometry description), Calvison (Optical Photon simulation and modeling of hadronic showers).
    - LArIAT Computing Liaison: helped to estimate Geant4 uncertainties on total cross sections.
- Other direct helps on a best effort basis
  - Running physics model validation jobs for Fermilab related programs (J. Yarba)
  - PIPII (Faraday cup design), DARPA (gamma-ray imaging), g-2 (Implementing symplectic integrators in Geant4, collaboration with J. Apostolakis/CERN and R. Fatemi/U. Kentucky), etc. (S.Y. Jun)

# Outlook: From Snowmass to P5

- Computation Frontier Summary by D. Elvira (CHEP 2023 talk and arXiv:2210.05822)
  - Findings: Long-term development, maintenance, and user support of essential software packages cutting across project or discipline boundaries are largely unsupported.
  - Resources for Geant4 physics models, software maintenance and support severely reduced in the USA.
  - Recommendation: the US HEP community should take a leading role in long-term development, maintenance, and user support of essential software packages with targeted investment. → CPSC
- P5 Town Hall (BNL) Meeting: Geant4, a focal HEP simulation tool, presentation by K. Genser
  - While being widely and successfully used in various contexts, it also has its limitations, in part, because of the lack of a sufficient number of people working on it - not sufficiently supported in the US.
- PAC Report (Jun 2023) Recommendation 4 (Strategic plan for Software and Computing)
  - We recommend that CSAID continues to contribute to major community HEP software tools like ROOT, GEANT, taking a leadership role for the Intensity Frontier aspects, and focusing on the needs of FNAL Physics Centers.

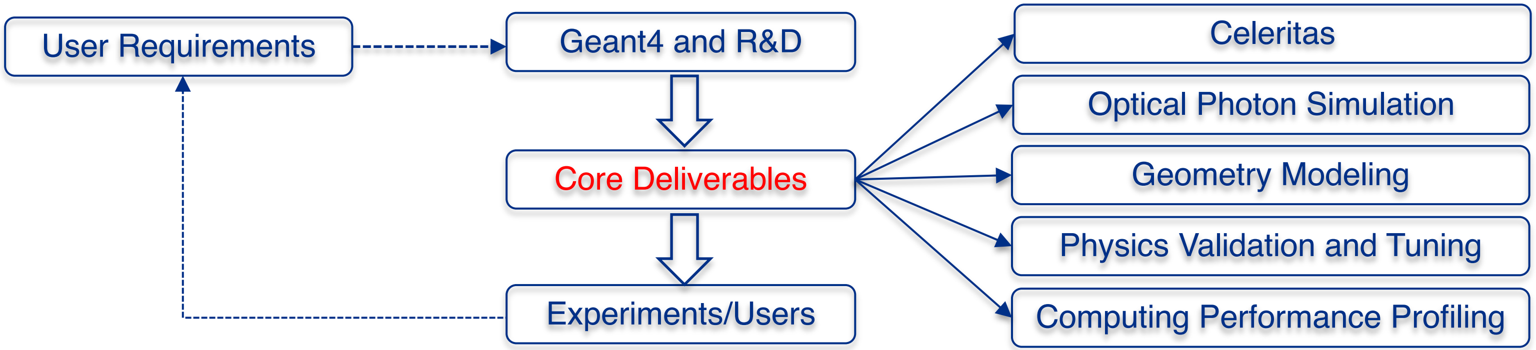

Fermilab

# Outlook: Related P5 Recommendation

- Area Recommendation 17:
  - Based on experience from the last decade, we recommend adding support for a sustained R&D effort—at the level of $9M per year in 2023 dollars—to adapt to emerging hardware and other computing technologies. This should include efforts to transition the products of software R&D into production. We also suggest that DOE's Office of High Energy Physics and NSF's Directorates for Mathematical and Physical Sciences and Astronomy coordinate with other programs within the Office of Science and NSF to ensure that the profile of computing resources available matches the needs of particle physics experiments.
- Area Recommendation 18:
  - We must ensure sustained development, maintenance, and user support for key cyberinfrastructure components, including widely-used software packages, simulation tools, and information resources, such as the Particle Data Group and INSPIRE. Although most of these shared cyberinfrastructure components are not specifically tied to projects, nearly all scientists in the field rely on them. A significant investment—at the level of $4M per year in 2023 dollars—for this type of shared cyberinfrastructure with dedicated personnel is appropriate.

# Fermilab Geant4 Team Strategy

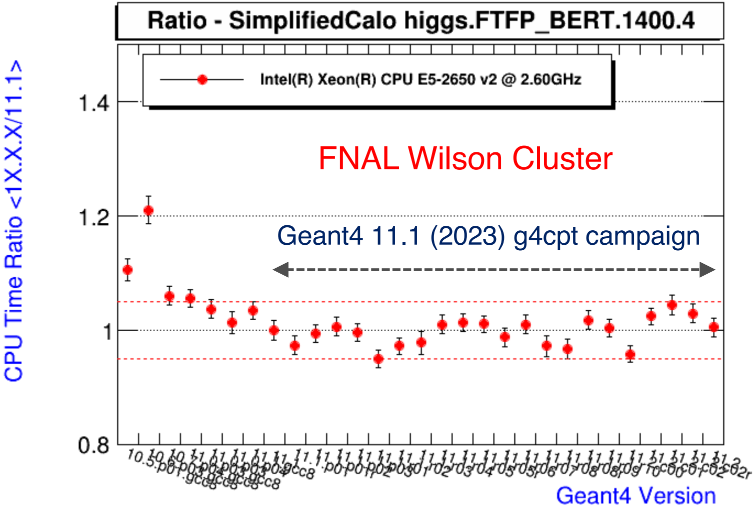
- Follow CSAID/DSSL/PhysicsSim vision and mission, and seek OHEP funding opportunities.
  - Contribute core Geant4 developments in close collaboration with domestic and international partners.
  - Lead Geant4 R&D for high performance computing, especially with co-processors (GPUs).
  - Support Fermilab experiments and user community.




- Maintain a strong focus on adapting cutting-edge technologies of future scientific software and computing for HEP detector simulation and provide deliverables to stakeholders.

# Geant4 Computing Performance Task (G4CPT)

- The Fermilab Geant4 team is responsible for benchmarking, profiling, monitoring CPU and memory performance for all Geant4 releases including internal reference releases by application {Geometry I Particle, Energy, Magnetic Field, Physics List} ~ 50+ samples.



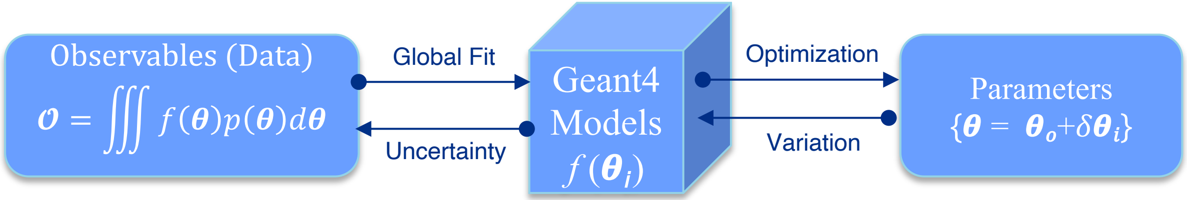

**Geant4.11.2.r02 SimplifiedCalo** Open|SpeedShop

| Sample | Physics List | B-Field | Energy |
|---|---|---|---|
| Higgs->ZZ | FTFP_BERT | ON (4.0T) | 14 TeV PYTHIA |
| | | OFF (0.0T) | 14 TeV PYTHIA |
| 100 MeV e- (5K e-/event) | FTFP_BERT | ON (4.0T) | 100 MeV |
| | Shielding | ON (4.0T) | 100 MeV |
| | Shielding_EMZ | ON (4.0T) | 100 MeV |
| Electrons | FTFP_BERT | ON (4.0T) | 1 GeV 5 GeV 10 GeV 50 GeV |
| | | OFF (0 T) | 1 GeV 5 GeV 10 GeV 50 GeV |
| Pions- | FTFP_BERT | ON (4.0T) | 1 GeV 5 GeV 10 GeV 50 GeV |
| | | OFF (0 T) | 1 GeV 5 GeV 10 GeV 50 GeV |
| | QGSP_BERT | ON (4.0T) | 1 GeV 5 GeV 10 GeV 50 GeV |
| | QGSP_BIC | ON (4.0T) | 1 GeV 5 GeV 10 GeV 50 GeV |
| | FTFP_INCLXX | ON (4.0T) | 1 GeV 5 GeV 10 GeV 15 GeV |
| Protons | FTFP_BERT | ON (4.0T) | 1 GeV 5 GeV 10 GeV 50 GeV |
| | FTFP_INCLXX | ON (4.0T) | 1 GeV 5 GeV 10 GeV 15 GeV |
| | FTFP_BERT_HP | ON (4.0T) | 1 GeV 5 GeV |
| | Shielding | ON (4.0T) | 1 GeV 5 GeV |
| Anti-Protons | FTFP_BERT | ON (4.0T) | 1 GeV 5 GeV 10 GeV 50 GeV |
| Gamma | FTFP_BERT_EMZ_AugerOff | OFF (0 T) | 250 MeV 1 GeV |
| Gamma | FTFP_BERT_EMZ_AugerOn | OFF (0 T) | 250 MeV 1 GeV |

- Report a summary to the category coordinators (~20x per year) and the Steering Board (~5x per year)

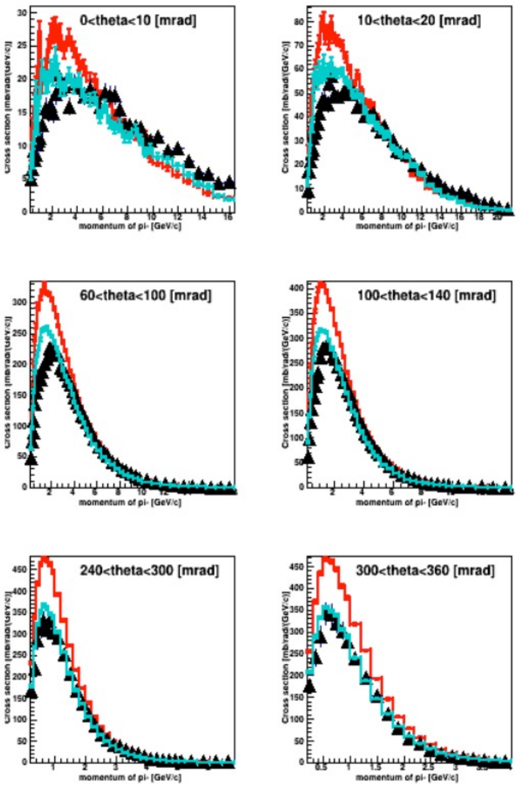
# Optimizing Geant4 Physics Models

- Improving simulation by tuning physics model parameters ($\boldsymbol{\theta}$)
  - Provide best sets of model parameters ($\boldsymbol{\theta}_o$) to match experimental data.
  - Reduce simulation uncertainties ($\delta\boldsymbol{\theta}_i$): helps to reduce systematic errors on experimental observables ($\mathcal{O}$) (model uncertainties).




  - Variation of parameters produces substantially better agreement with some datasets, but more degrees of freedom are required for full agreement [(GEANT4 parameter tuning using Professor)]().
  - Example: Julia Yarba's FTF Tune vs. the default Geant4 v11.1
  - Unfortunately, some widely used hadronic models have not been actively developed over the last few years due to the lack of personnel.

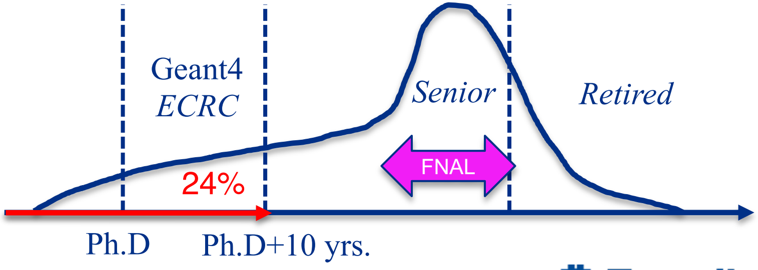
31GeV p + C → $\pi^-$ + X (NA61 data)




$\chi^2$/n.d.f = 27.7 (default) vs. $\chi^2$/n.d.f = 8.6 (tune)

# Sustainability of Geant4

- There are collaboration-wide concerns regarding generational imbalance.
- Sustainability of Geant4 was emphasized (by K. Genser) at the P5 Town Hall Meeting.
  - Physics Models: The people working on the physics models relied on by experiments should be specifically funded to do so.
  - Collaboration: As the needs of experiments and computing environments evolve, the Geant4 toolkit requires constant development, maintenance, and user support. All that requires an adequate number of people and stable funding.
- Many Geant4 senior developers already retired and may not continue beyond *O*(10) years from now, and all Fermilab Geant4 members have been affiliated with the lab for many years.
  - Geant4 Early Career Researchers (ECRC): only 24%
  - Some expertise is hard to find and takes time to transfer.
  - Transition to a new generation must start now.
  - Expertise disruption/loss risk is very high.

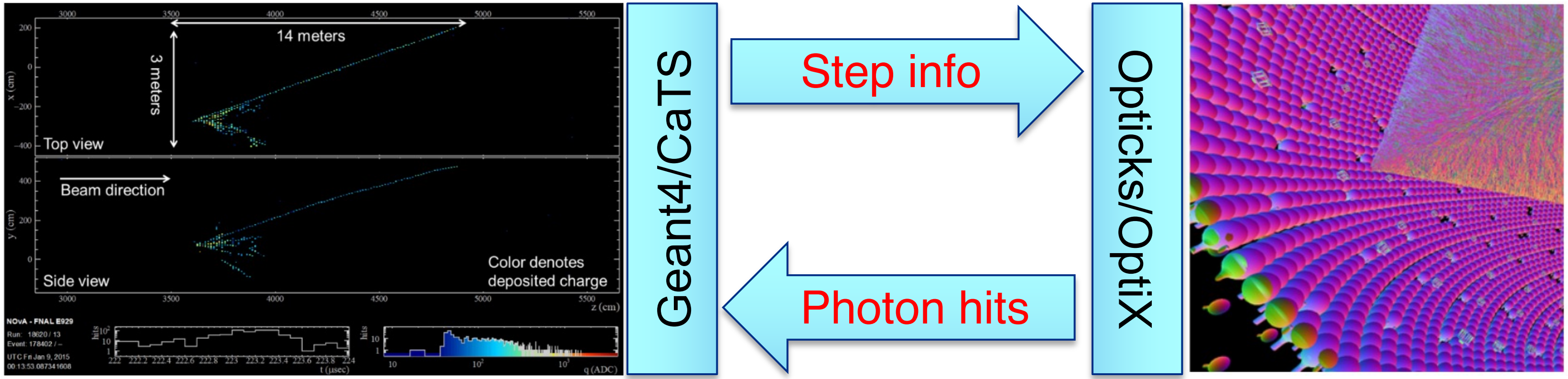


# Geant4 Related R&D Projects and Our Focus

- Mandate of Geant4 R&D Task Force
  - Promote longer-term R&D efforts to explore emerging technologies, computing architectures or software architectural revisions, and new or better physics ideas that would be beneficial to Geant4.
  - Make timely assessments on these R&Ds for their feasibility, benefits and required efforts.
- Main Development within Geant4
  - Improvement, optimization, modernization and refactoring of the existing Geant4 code
  - Development and integration of fast simulation techniques (e.g., Calo-Challenge/ML)
  - Investigation of the potential use of accelerators (CaTS, AdePT and Celeritas)
- Fermilab Geant4 Team focus
  - Optical photon simulation on GPU: CaTS/Opticks → LArSoft(artg4tk/Opticks) (H. Wenzel)
  - Accelerating simulation on heterogeneous architecture: Celeritas (P. Canal, S.Y. Jun, G. Lima)
  - Surface-based geometry models (G. Lima)

# Optical Photon Simulation on GPU

- Optical photons are copiously produced (ex. 45,000 for 1 MeV energy loss in LAr) and most neutrino and dark matter search experiments need an accurate simulation of optical photons.
- A full optical photon simulation for a charged particle of GeV-level energy requires significant computational resources (CPU and memory) → usually approximation by using lookup tables.
- A solution: GPU accelerated optical photon simulation using ray tracing (e.g., NVIDIA OptiX)

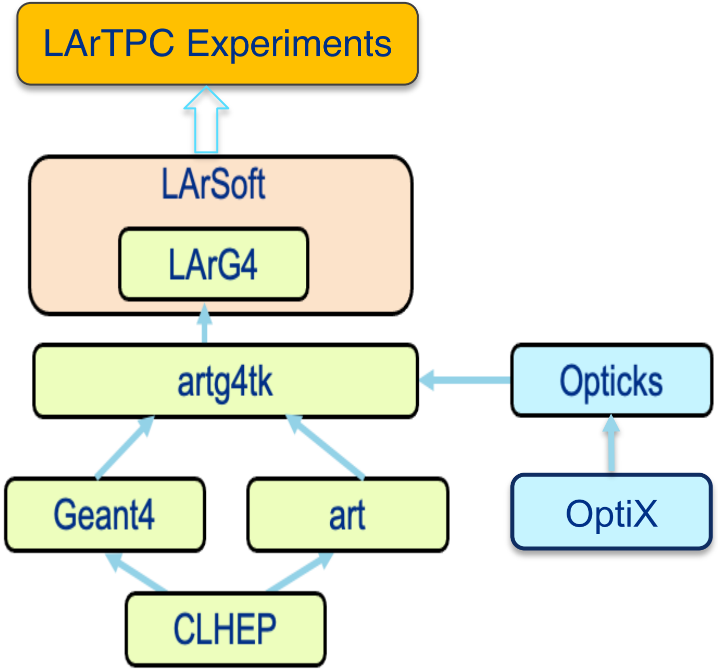



$T_{CPU} = O$ (1s) without Optical Photons

$T_{CPU} = O$ ($10^3$s) → $T_{GPU}$ = O (1s)

# CaTS/Opticks/OptiX Status and Roadmap

- Opticks was migrated to OptiX7 (Feb. 1, 2024) and CaTS was updated accordingly.
  - Opticks: GPU Accelerated Optical Photon Simulation using NVIDIA OptiX (Simon Blyth/JUNO)
  - CaTS: A Geant4 example that demonstrates how to use opticks within Geant4 for the creation and propagation of optical photons (Hans Wenzel)
- Integration CaTS/Opticks into LArSoft
  - CaTS functionality ➔ artg4tk/LArG4 (2024, Q1)
  - Benchmark and integration into one of LArSoft/LArTPC-based experiments (2024, Q2)
  - Big interest from external HEP groups (LZ, CalVision, EIC/ePIC, etc.)



# Celeritas

- Celeritas is a new detector simulation code designed for computationally intensive applications on high-performance heterogeneous architectures
  - **Research and develop** novel algorithms for GPU-based simulation in HEP
  - **Implement** production-quality code for GPU simulation
  - **Integrate** collaboratively into experiment frameworks (for LHC experiments)
- Jointly funded by US DOE ASCR and HEP-ECP/SciDAC
  - Initial funding: EM particle transport on GPU interfacing through Geant4
  - Extension: Portable optical photon simulation and neutron transport on GPU
  - Partnership: ORNL, Fermilab, ANL, LBNL and U. of Warwick (UK/Excalibur)

CELERITAS

*Celeritas core team:*

Elliott Biondo *(ORNL)*,
Julien Esseiva *(LBNL)*,
Seth R Johnson *(ORNL)*,
Soon Yung Jun *(FNAL)*,
Guilherme Lima *(FNAL)*,
Amanda Lund *(ANL)*,
Ben Morgan *(U Warwick)*,
Stefano Tognini *(ORNL)*

*Celeritas core advisors:*

Tom Evans *(ORNL)*,
Philippe Canal *(FNAL)*,
Marcel Demarteau *(ORNL)*,
Paul Romano *(ANL)*

# Motivation

- GPU is current main HPC architecture and can no longer be ignored. Many HEP-CCE R&Ds are leveraging heterogeneity
- Take advantage of increasing heterogeneity in compute resources, especially Exa-scale computers under DOE flagship facilities with GPUs

NVIDIA (A100) GPUs


0.1EF(2021)

AMD (Instinct) GPUs


1.6EF (2022)

Intel (Max) GPUs


2.0EF (2023)

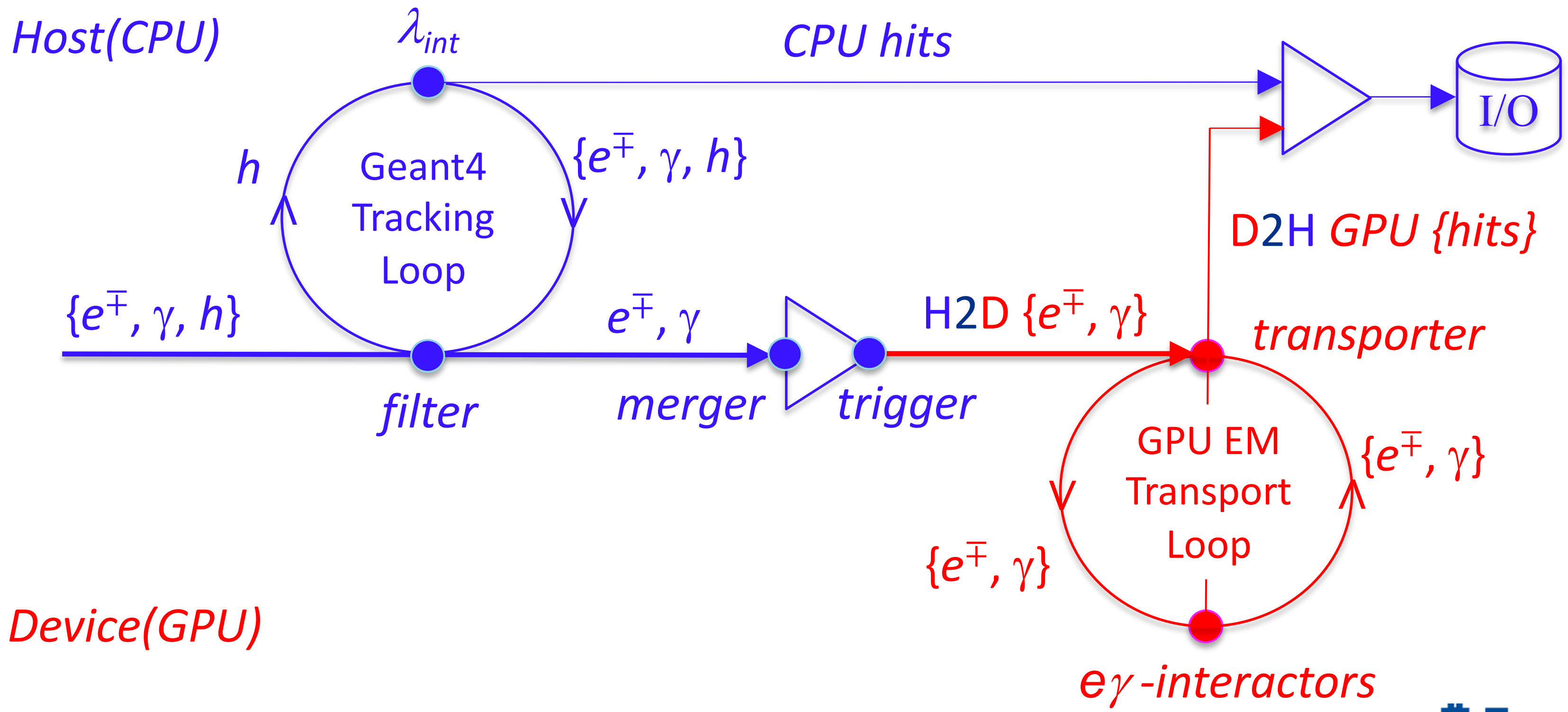
AMD (MI300) APUs


2.0EF (2024)

- Celeritas (Fermilab, LBNL/NERSC, ORNL, ANL) has access time for these facilities and targets architecture-independent detector simulation code (portability for CUDA, HIP, oneAPI)

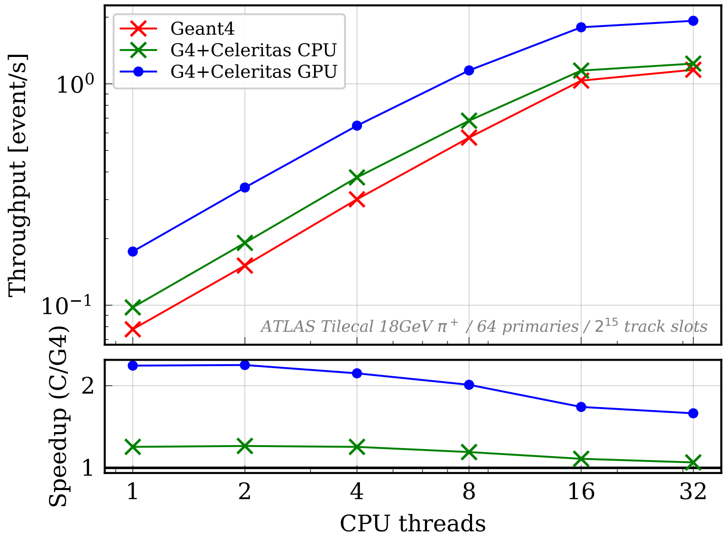
# Celeritas-Geant4 Application: How does it work?

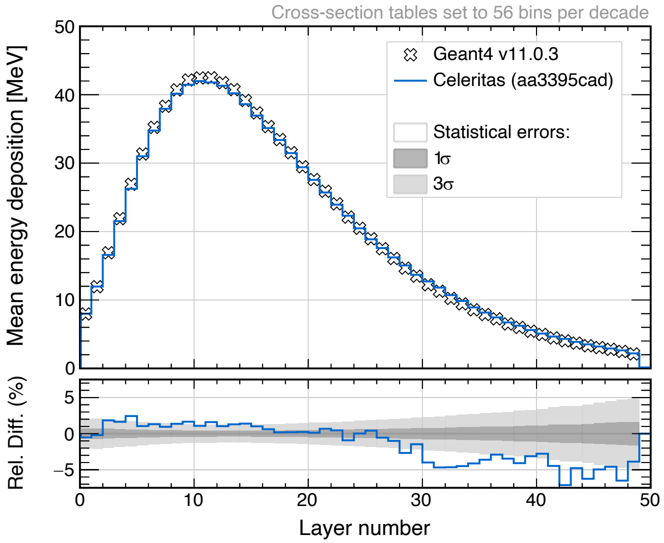
- A hybrid Geant4 workflow with selected tasks (e.g., EM particle transport) executed on GPUs



# Celeritas Status

- EM particle transport on GPUs has been completed and initial integration into CMS/CMSSW and ATLAS/Light-sim has been successfully demonstrated (toward v1.0).
- Computing and Physics Performance: ATLAS (tile calorimeter) as an example




- The relative speedup by Celeritas is over 2x up to 8 threads (the maximum speedup: ~3.0x)

# Celeritas Status by the Numbers

| | | |
|---|---|---|
| **100** | **lines of code** | to integrate Celeritas into a FullSimLight tile calorimeter test application, with no modifications to Geant4 |
| **1.8×** | **full-simulation speedup** | including hadronics and SD hits, by using 1 Nvidia A100 with 16 AMD EPYC cores for the ATLAS test beam application *[NERSC Perlmutter]* |
| **2–20×** | **throughput** | when using Celeritas on GPU (compared to Geant4 MT CPU) for EM test problems *[NERSC Perlmutter]* |
| **4×** | **performance per watt** | for TestEM3 (ORANGE geometry) using Celeritas GPU instead of Geant4 CPU *[NERSC Perlmutter]* |

# Geant4 Assessment of Simulation R&D Projects: AdePT and Celeritas

- Dec. 13-14, 2023 @CERN
- Findings
  - Physics kernels and CPU-GPU communication are not limiting factors.
  - Geometry is the limiting factor (96% of standalone runtime in CMS2018).
- Summary
  - Demonstrated capability of processing LHC scale detector simulation.
  - No show stoppers have been identified (physics, CPU-GPU scalability).
- Recommendation
  - The panel recommended to continue two projects and to collaborate on integrating developments into the HEP experiment's framework.
  - Explore the surface-based geometry approach, its memory consumption and scalability to large structures (delta-assessment by Dec. 2024).

**Gean4 R&D 2024-GPU Assessment Team**

Makoto Asai (Geant4/EIC, Jlab)

John Chapman (ATLAS, Cambridge)

Gloria Corti (LHCb, CERN)

Ivana Hrivnacova (Geant4/ALICE)

Vincenzo Innocente (CMS, CERN)

Matti Kortelainen (CMS, Fermilab)

Andrea Valassi (CERN IT)

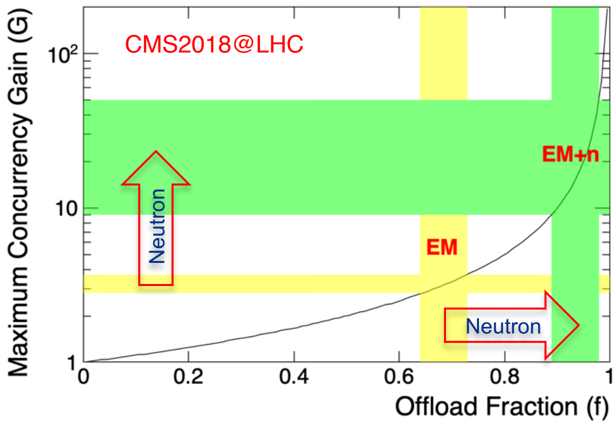
Sandro Wenzel (ALICE, CERN)

Ben Wynne (ATLAS, Edinburgh)

Marc Verderi (chair) (Geant4, LLR)

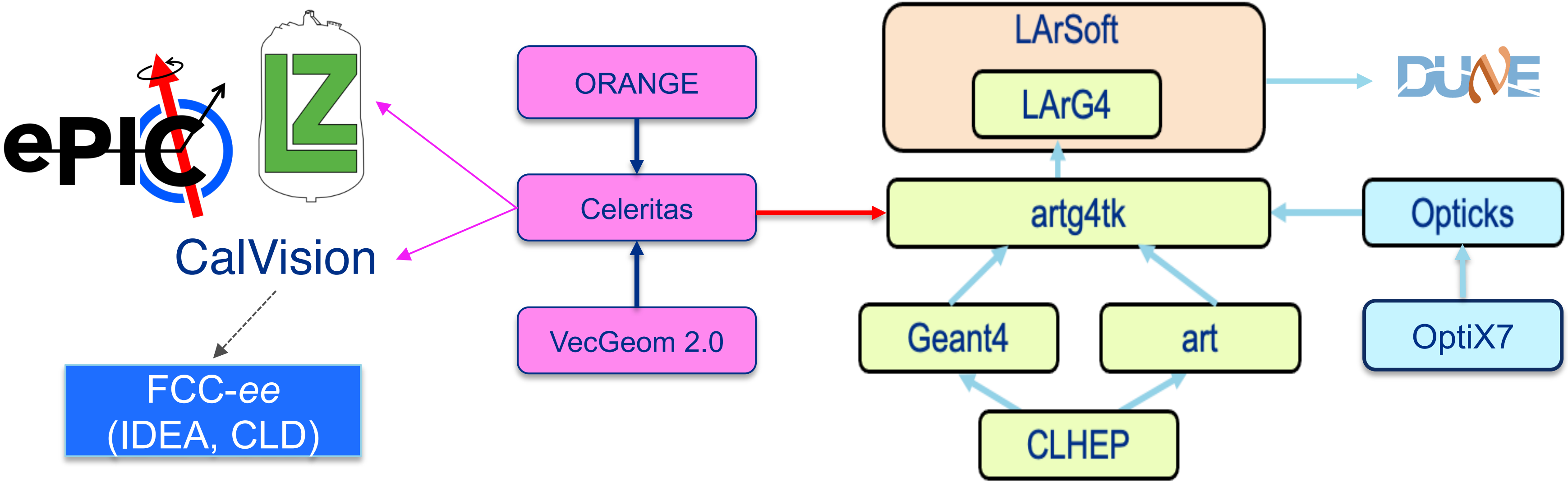
# Celeritas Roadmap and Fermilab Contributions

- Align with DOE ASCR/HEP-CCE and recommendation by Geant4 (2024).
  - Complete integration into HEP experiments for production and evaluate computing performance.
  - Participate in surface-driven geometry model developments (VecGeom2.0 and ORANGE).
  - Add remaining EM physics models (high energy multiple scattering and photo-/electro-nuclear models).
- Add as much work as possible on GPU → toward a full fidelity simulation toolkit.
  - Support optical photon transport (2024, HEP-CCE)
    - Remove NVIDIA/CUDA and OptiX dependencies.
    - Use generic HEP geometry and navigators.
  - Add neutron-nucleus (elastic, inelastic, capture) interactions.
    - Increase the maximum projected gain significantly, ~$O$ (10).
    - 0.5 FTE for three years (SciDAC extension) (2024-2026)



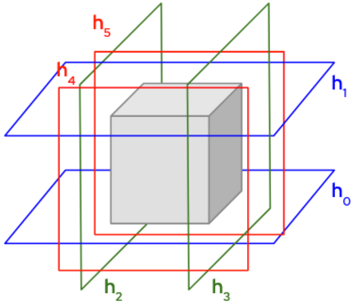
# Roadmap: Celeritas Optical Photon Simulation

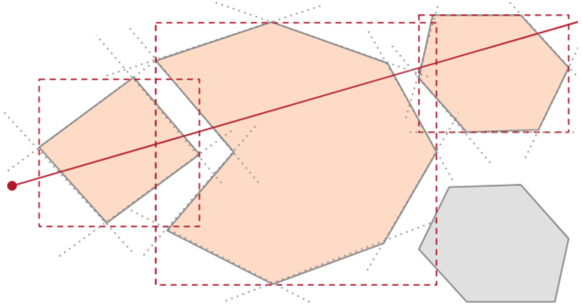
- Funded by HEP-CCE for one-year development (started from Feb. 2024)
  - A platform-independent full optical photon simulation (CUDA, HIP, etc.)
  - Generic geometry engine based on surface models (VecGeom2.0 or ORANGE)



# Geometry Roadmap: VecGeom 2.0 and ORANGE

- The Fermilab Geant4 Team has participated in the VecGeom/VecCore development from the beginning and was one of the major contributors to the project.
- Efficient geometry algorithms will be essential components of simulation on modern hardware.
  - Take a leadership role in relevant R&D efforts.
  - Acquire expertise in the surface-based models, associated navigation (moving through geometrical volume hierarchy) engines and their applications.
- The plan is to participate in the surface-based shape implementation in close collaboration with the Geant4 geometry development team (Guilherme Lima).
  - CERN: VecGeom2.0 (bounded model)
  - ORNL: ORANGE (unbounded model)
  - No custom navigation needed per solid type (efficient for GPU)



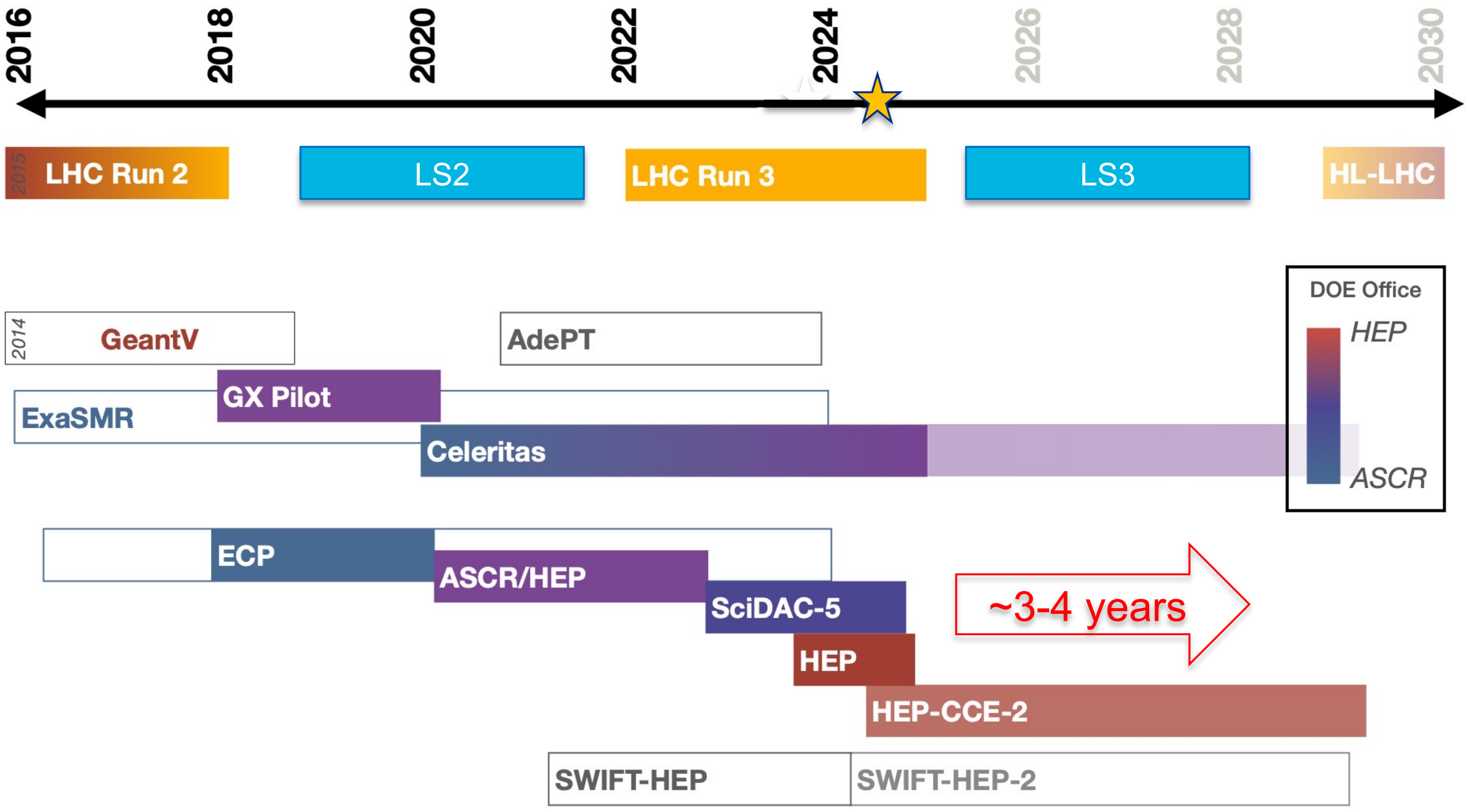
# Plans

- Continue contributing to core activities of Geant4 and support local experiments.
  - Computing performance monitoring, profiling and optimization
  - Physics validation and tuning physics model parameters relevant for Fermilab programs
- Be actively involved in relevant R&D efforts to improve Geant4 and related software.
  - Complete the CaTS/Opticks integration into LArSoft/LArTPC-based experiments (2024, Q2).
  - Continue to take leadership roles in Celeritas development.
    - Complete implementation of EM physics models, integration into the HEP experimental frameworks, optimization, physics validation (2024-2027, HEP-CCE2).
    - Implement architecture-independent optical photon simulation (2024, HEP-CCE).
    - Implement neutron transport (2024 – 2025, SciDAC).
  - Participate in VecGeom 2.0 development (and ORANGE) (HEP-CCE2, Operations) .

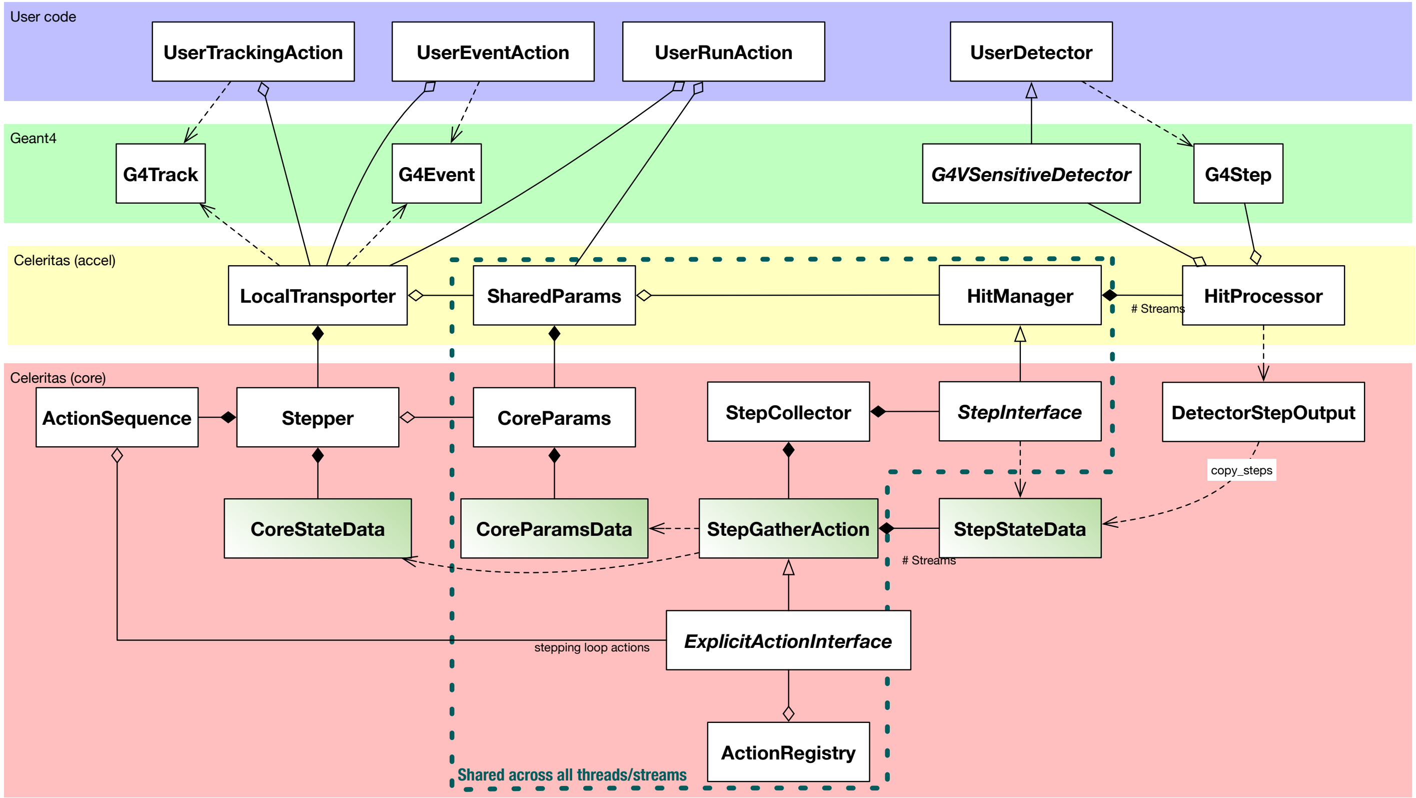
Plans of the Fermilab Geant4 team are aligned with Fermilab's mission and P5 recommendations.

# Backup

# Celeritas Timeline

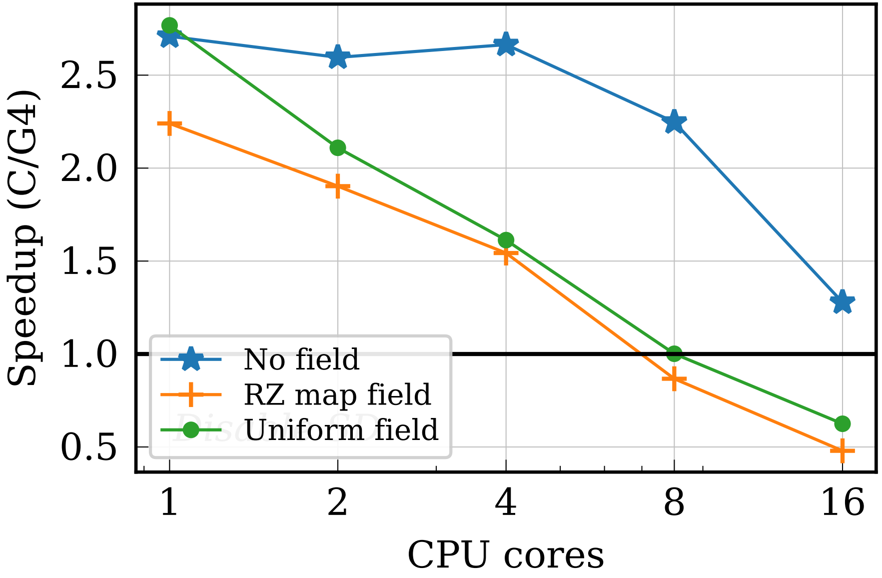


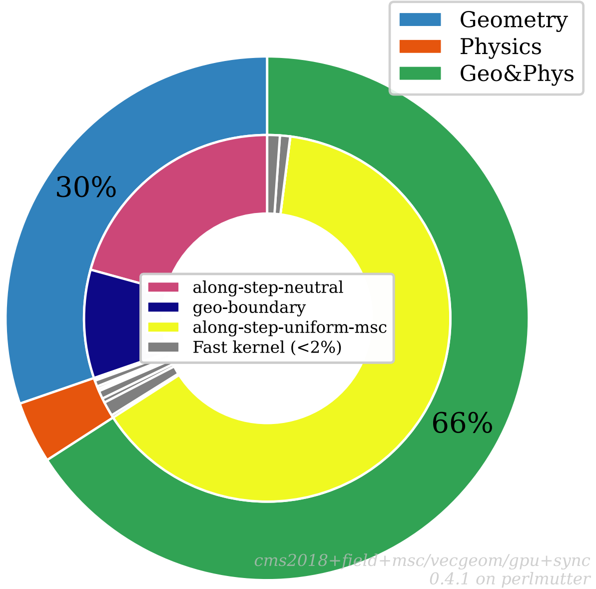
# Celeritas-Geant4 Workflow Chart



# Celeritas Computing and Physics Performance

- CMS HL-LHC detector geometry




- Close to the maximum gain with one thread/core, but need to investigate (N cores – 1 device) performance and optimize (N Cores – M devices)



Fermilab