Session 8B: Vector & Accelerator

(co-chaired by John Apostolakis and Soon Yung Jun)

Extending GeantV to accelerators	*VALLECORSA, Sofia et al.*
Auto-vectorization: recent progress	*VALLECORSA, Sofia et al.*
Scheduling fine grain workloads in GeantV	*GHEATA, Andrei*
Progress in GeantV field propagation and EM physics	*APOSTOLAKIS, John et al.*
Neutron HPC proposal	*DOTTI, Andrea*
Aula Magna, Ferrara	10:00 - 10:15
Discussion	
Aula Magna, Ferrara	10:15 - 10:30

Extending GeantV to accelerators

S.Vallecorsa for the GeantV team

Auto-vectorization: recent progress

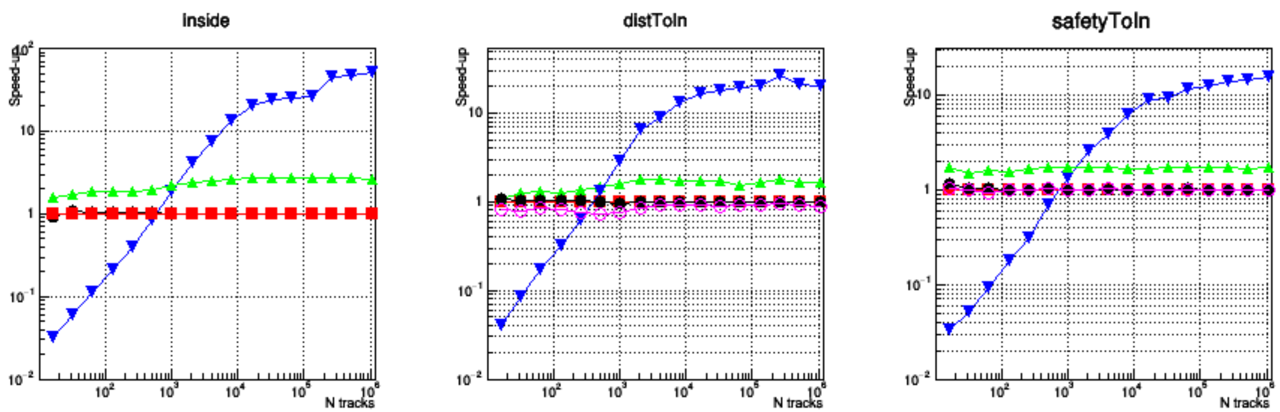
Guilherme Amadio, Sofia Vallecorsa

Geant4 collaboration meeting - September 2016

VecGeom on accelerators

Speedup for different navigation methods for the BOX (normalized to scalar)

Scalar (specialized/unspecialized)
Vector GPU (K20) ROOT

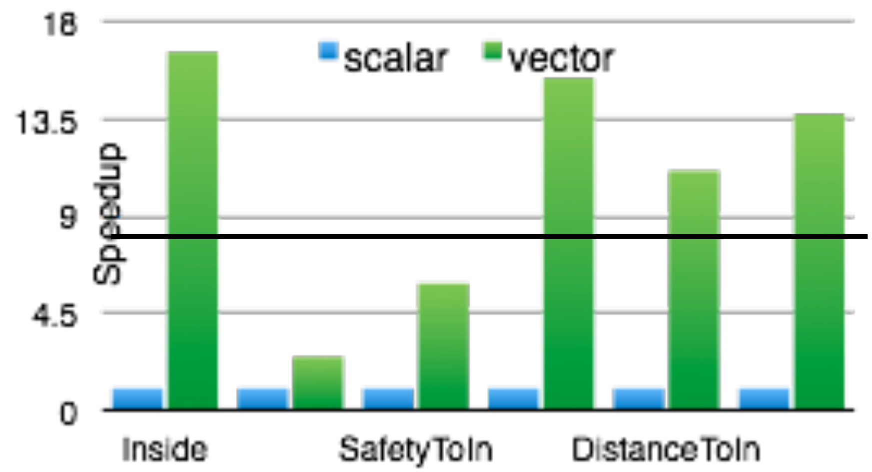


On KNL (AVX512) using UME::SIMD backend.

KNL systems use 512 bit registers corresponding to 16 SIMD lanes in DP (and 32 in SP)

Observe super-linear speedup for some methods

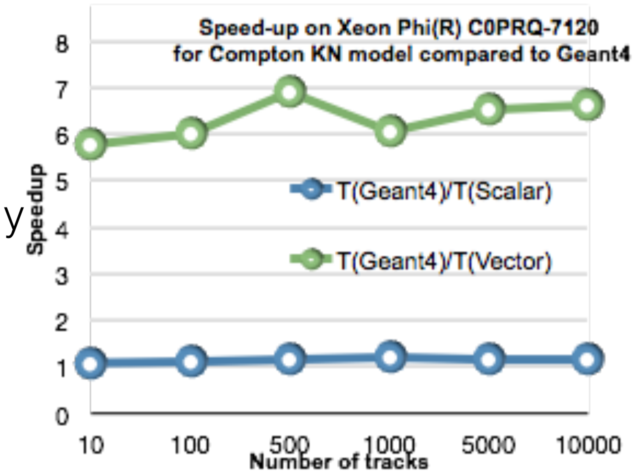
Intel® Xeon Phi™ CPU 7210 @ 1.30GHz, 64 cores

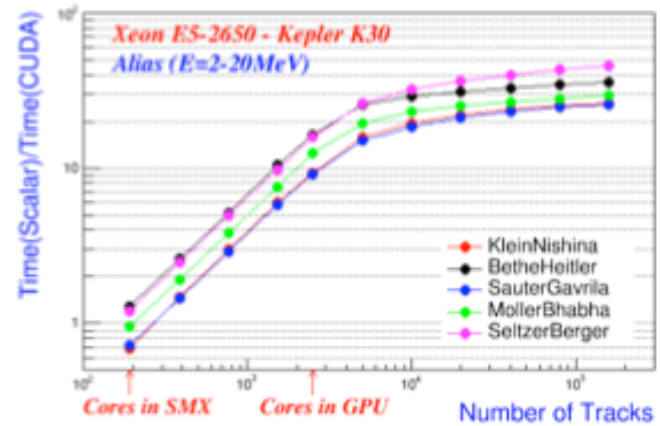


Beyond Geometry

Electromagnetic processes are vector/accelerator friendly

The vectorised Compton scattering shows good performance gains





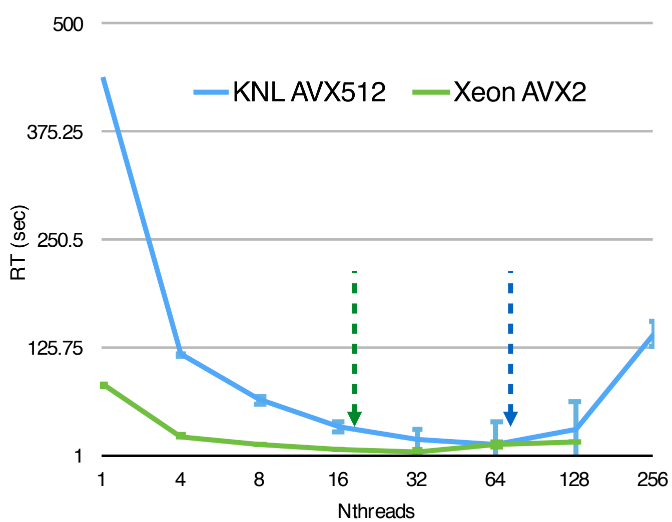
First GeantV full navigation on KNL

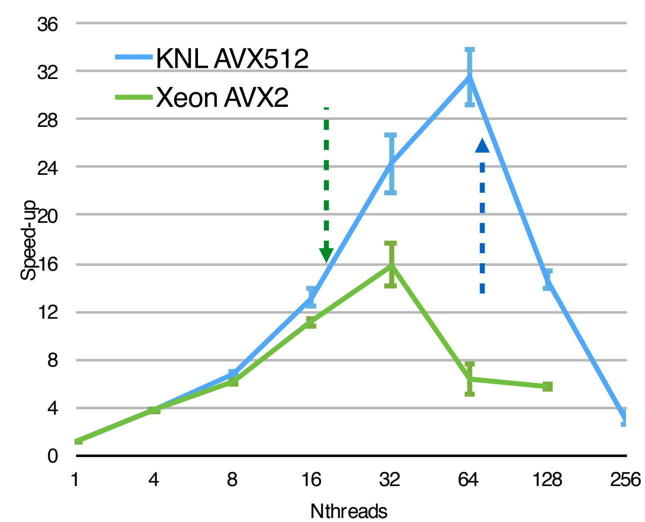
Tabulated physics

Simplified detector geometry

Full track transport and basketization procedure

Use UMESIMD backend for AVX512

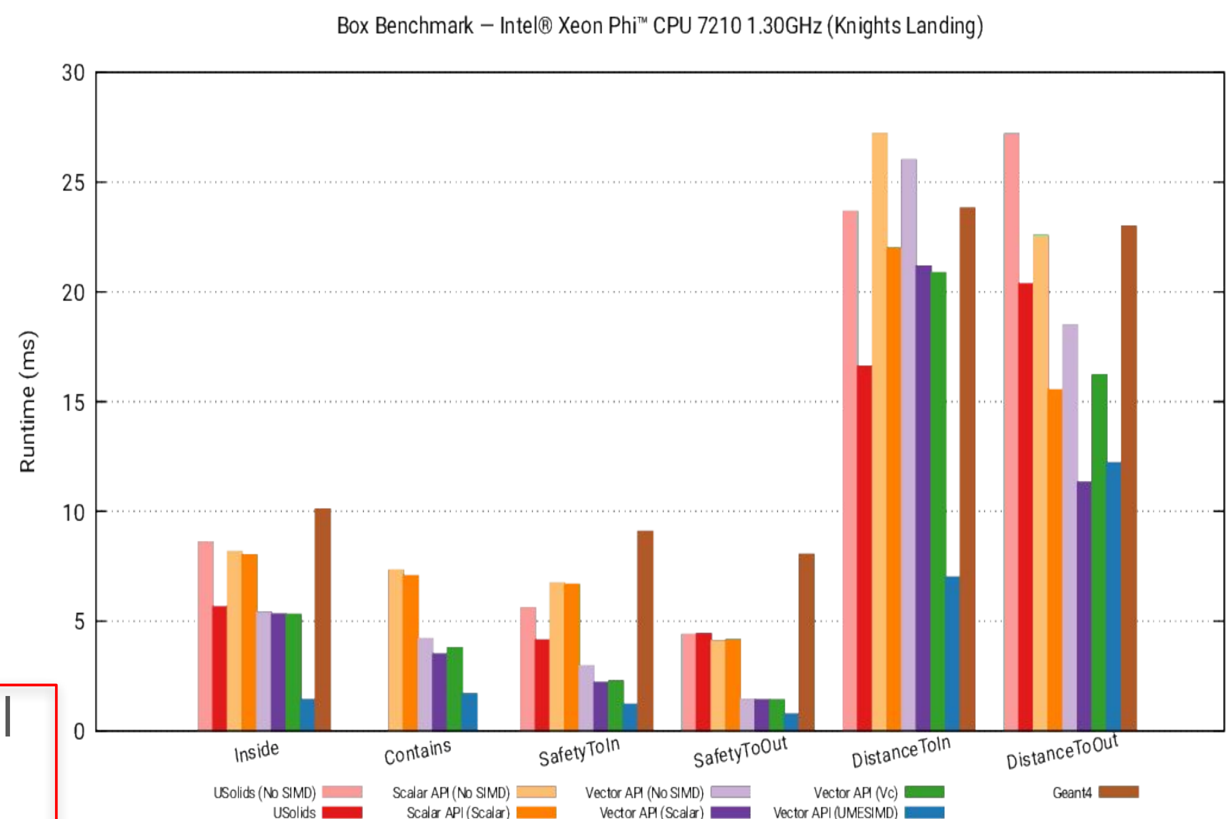




Auto-vectorization

VecGeom Benchmarks on Intel® Xeon Phi™ (KNL)

- Everything was compiled with Intel C/C++ compiler 16.0.3
- Used “-O3 -xMIC-AVX512”
- Contrary to AVX2 benchmarks on Skylake, UME::SIMD gives best performance on Knights Landing
- Scalar code under Vector API shows auto-vectorization in many cases



7

Summary

- Auto-vectorization is a powerful tool and compilers are getting better at it.
  - PRO: Almost “free lunch” provided the code is free of “vectorization hazards”
  - CONS: There are still differences among compilers, operations, architectures.
- However explicit vectorization using specific libraries still gives significantly the best result (ex. Vc for AVX2 and UME::SIMD for AVX512)

Scheduling fine grain workloads in GeantV

A.Gheata

Geant4 21st Collaboration Meeting

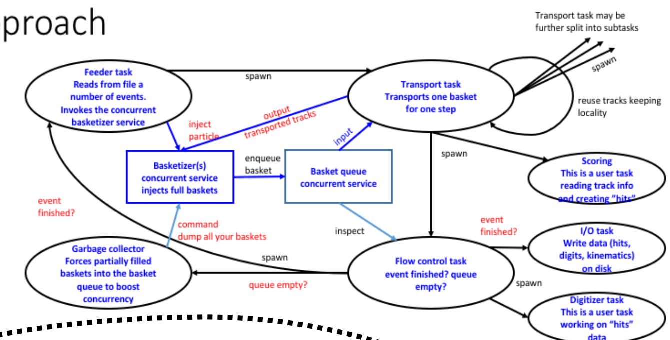
Ferrara, Italy

12-16 September 2016

Scheduling fine grain workloads in GeantV

- Migration to TBB tasks for better integration with concurrent experimental frameworks
- Partially implemented, more tasks to be added
- Good preliminary results

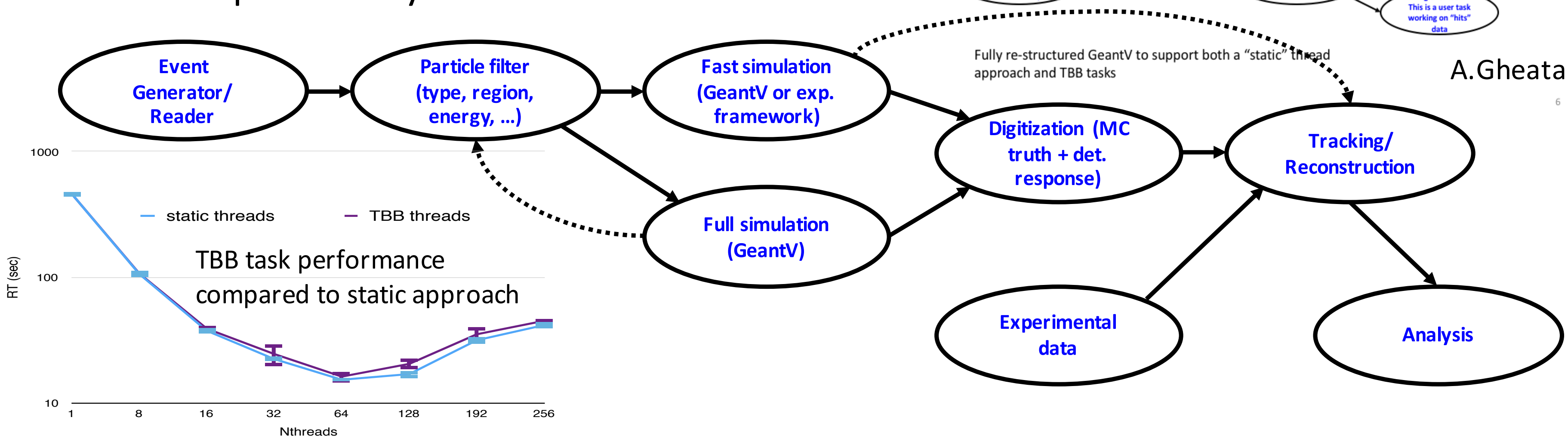
Framework: GeantV moving to a task approach



Fully re-structured GeantV to support both a "static" thread approach and TBB tasks

A.Gheata

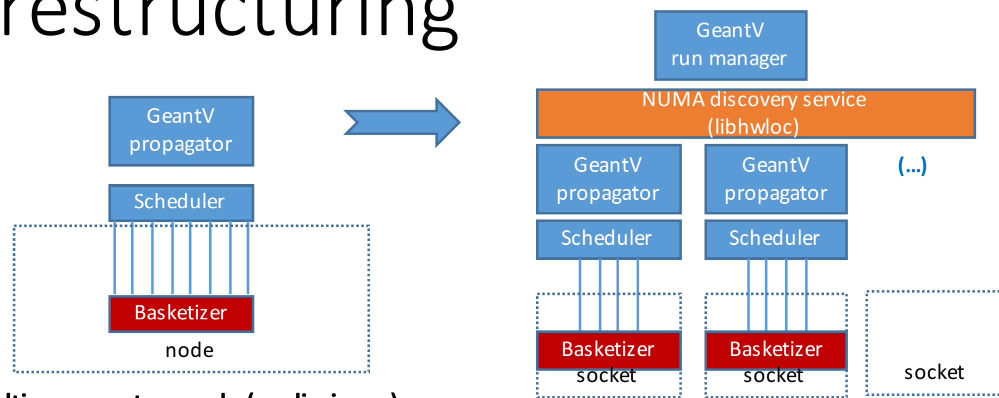
6



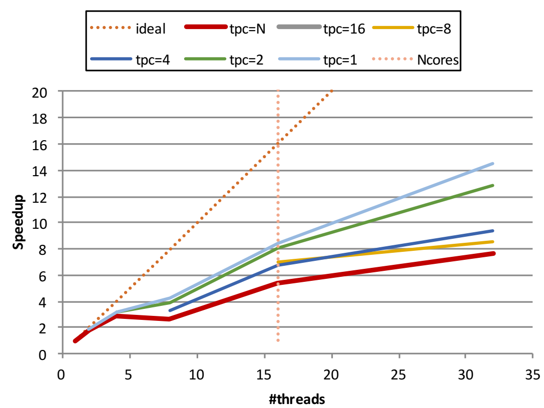
TBB task performance compared to static approach

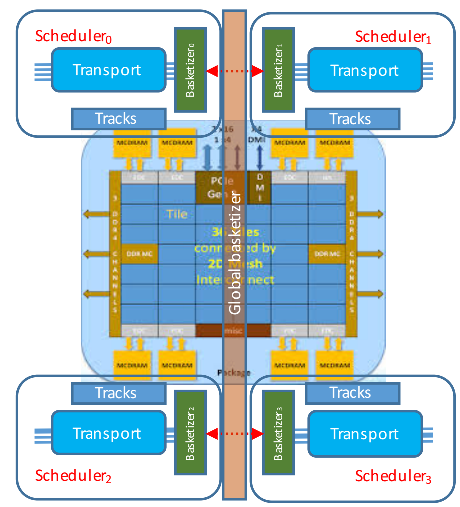
GeantV concurrency restructuring

- Work for better scalability on many-core (single process/multi cluster)
  - Introducing NUMA awareness
- Investigating concurrency optimizations oportunities in HPC environments



Multi propagator mode (preliminary)





Vector Electromagnetic Physics Models & Field Propagation

Guilherme Amadio (UNESP),
Ananya,
John Apostolakis, Marilena Bandieramonte, Mihaly Novak (CERN)
Soon Yung Jun (Fermilab)

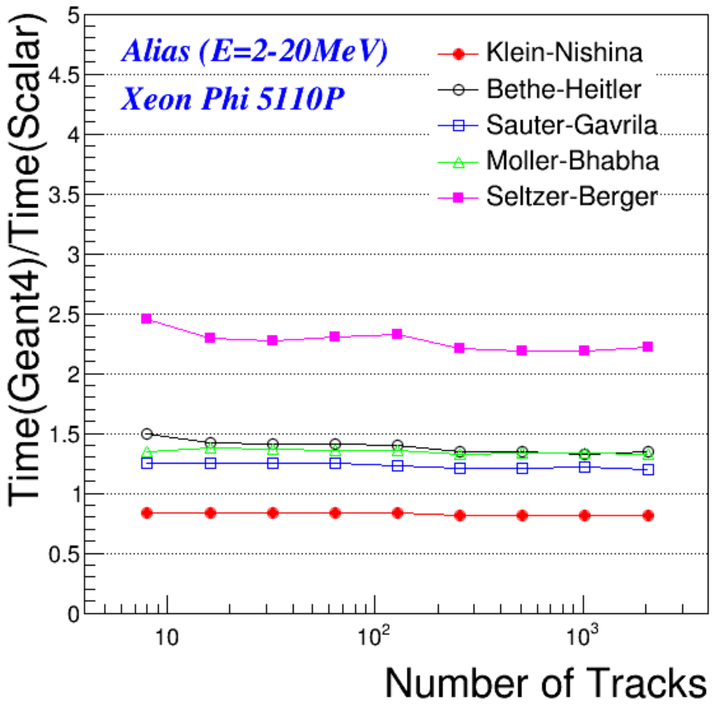
21th Geant4 Collaboration Meeting

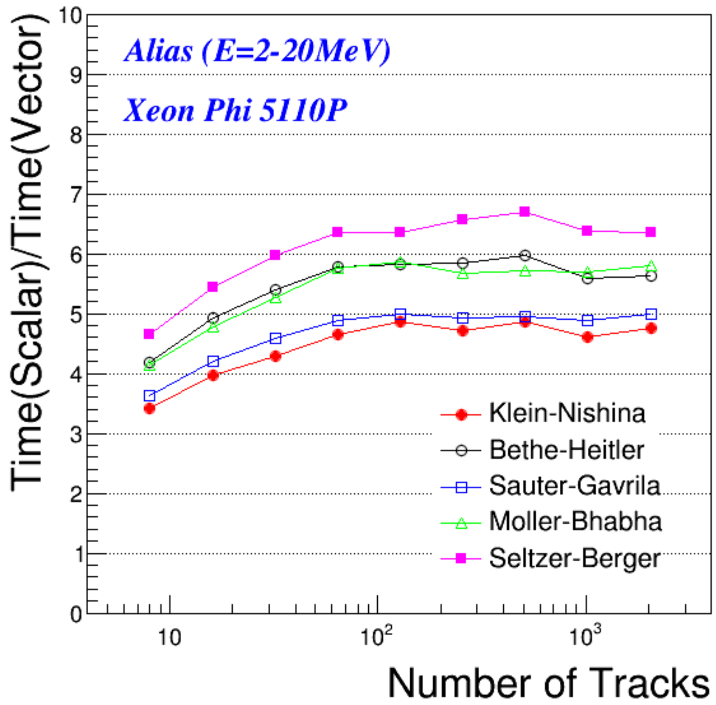
Magnetic Field Propagation

Ananya
John Apostolakis

Preliminary Performance: Alias Sampling Method

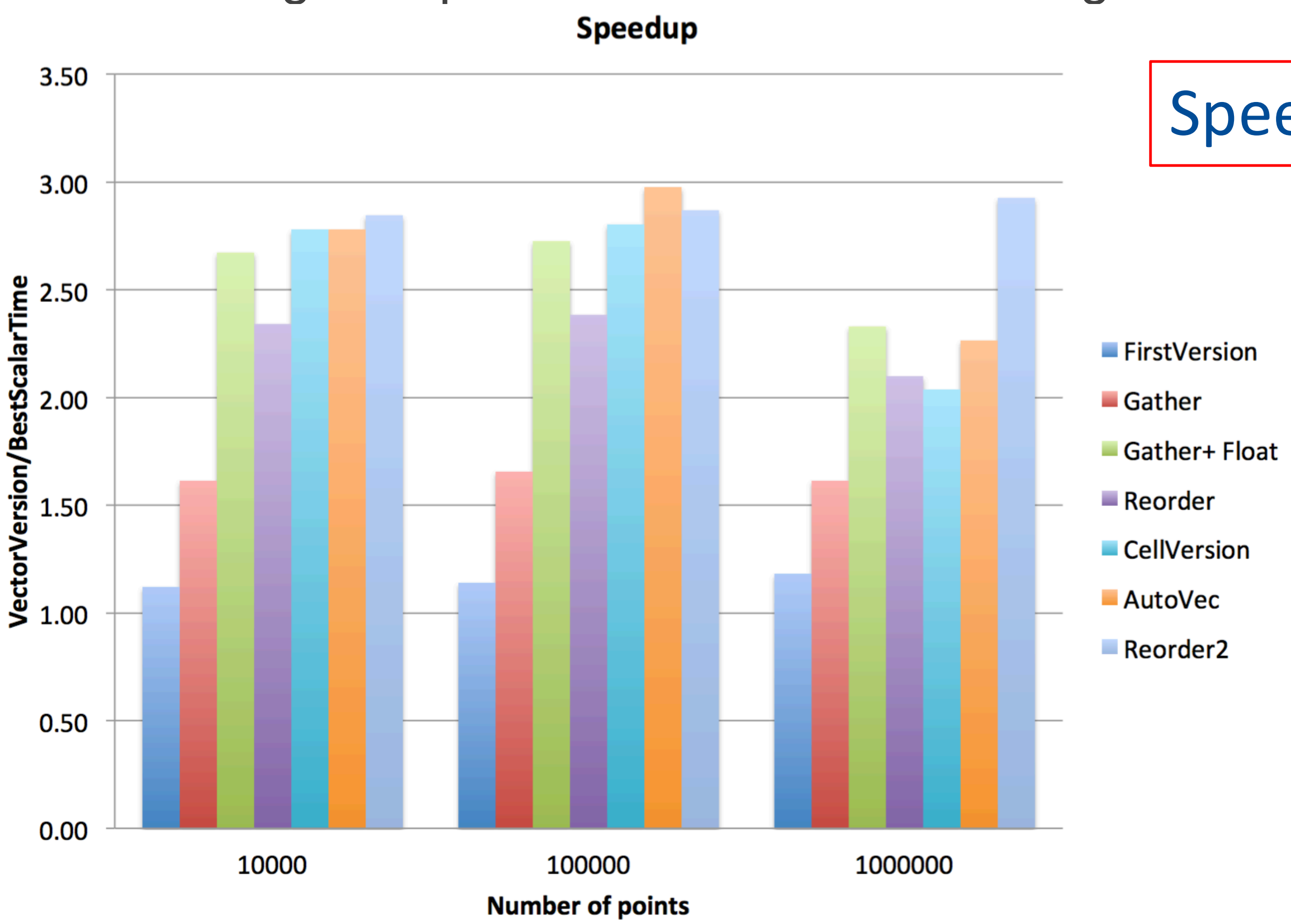
- Vector performance
  - input particle energy: 2-20 MeV (valid range for all models)
  - using 16 elements (random for each track)
  - MIC (Intel Xeon Phi 5110P 60 cores @ 1.053 GHz) - 8 vector pipelines for double precision – see also SSE/AVX in backup





Field Propagation - Outlook

- Vectorized field propagation is in progress
  - Field & equation of motion vectorised
  - Helix and different RK steppers vectorised
  - ‘Driver’: different methods for keeping vectorisation when handling multiple tracks are under investigation



Speedup factor of ~3

Neutron HPC

Makoto Asai, Andrea Dotti (adotti@slac.stanford.edu), Tatsumi Koi ;
SLAC/SD/EPP/Computing

Geant4 21st Collaboration Meeting – Ferrara, 12-16 September 2016

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

FY2017 Proposal from SLAC team to DOE

SLAC team has proposed a ~0.5FTE work-plan for FY17+ (funding pending):

- to capitalize on the very successful experience of MPEXS
- to leverage expertise of Stanford NVIDIA Center of Excellence (ICME)
- to leverage SLAC specific expertise on neutron interactions, especially at lower energies

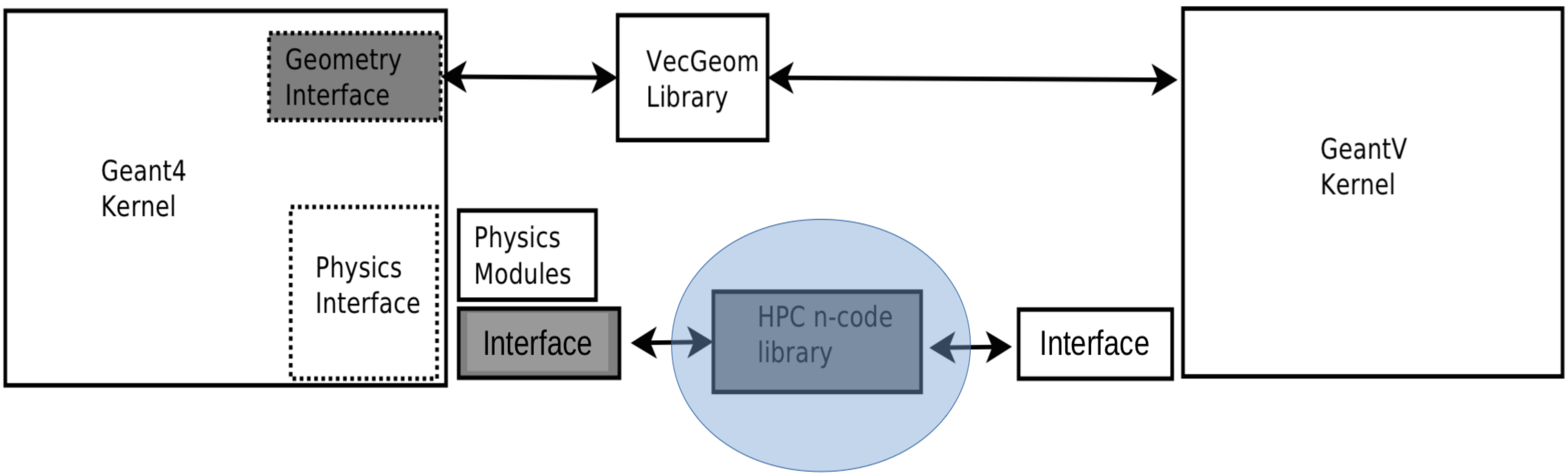
Develop a stand-alone, GPU friendly, neutron specific physics simulation library:

- outside of any specific “toolkit”, but with integration into Geant4 and GeantV in mind
- specialized code to deal with (low Energy) neutron interactions

Why (low-E) neutrons?

- for their nature they perform several very similar interactions: physics variety is relatively simple
- can reasonably limit variety of secondary species (pre-requisite for efficient GPU-style code)
- great local expertise that make the problem a success and a laboratory for possible future extensions

Preliminary Overview



3