Acceleration of Sensitivity Map Calculation for PET Imaging Using Intel Xeon Phi Coprocessor

T. Dey, P. Rodrigues, Oncology Solutions, Philips Research, Eindhoven, the Netherlands

Contact: *Thomas.dey@philips.com*

Purpose:

- We aim at the evaluation of the Intel Xeon Phi co-processor for acceleration of 3D positron Emission Tomography (PET) reconstruction.
- We cover the calculation of sensitivity map, scatter estimation and back projection as computational hot spots of PET reconstruction.
- We use dynamic load balancing to share workload between two Phi cards and host system to reduce the runtime of each module (Fig. 1).
- Here sensitivity map generation is chosen as an example to show details of the implementation and results.

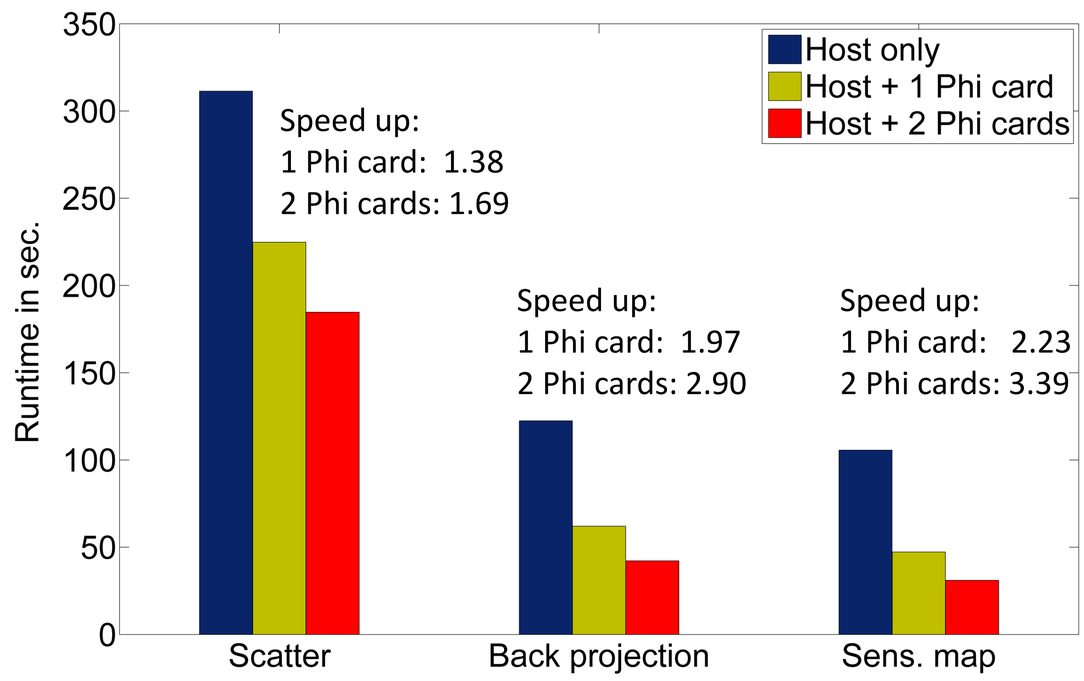

Fig 1: Runtime of PET reconstruction modules using Host and Xeon Phi.

Sensitivity map generation:

Algorithm:

The algorithm for sensitivity map generation is shown in **Fig. 2**. We use a voxel grid of 144 × 144 × 44 and 4096 ray samples per voxel. The radiological path is calculated by an algorithm of Jacobs et al[1].

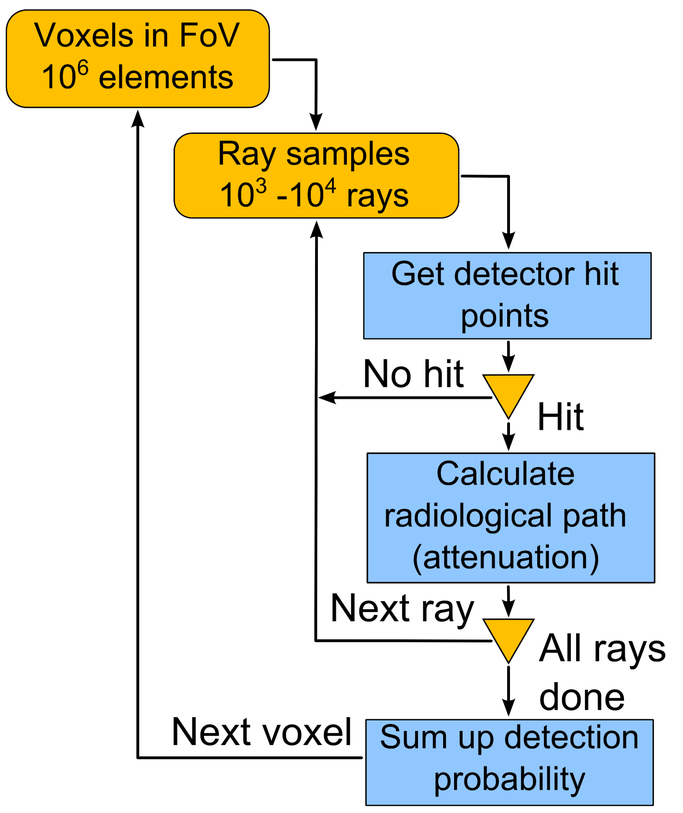

Fig. 2: Flowchart of sensitivity map algorithm.

Implementation and test platform:

- **Embree**[2] ray tracing kernels, optimized for host and Xeon Phi system to get the detector hit points.
- Parallel processing of voxels via **pthreads**.
- Intel Single Program Multiple Data (SPMD) compiler (**ispc**[3]) is used to exploit the vector processing units (VPUs) of the host and Xeon Phi system to calculate multiple sample rays simultaneously.
- In addition, radiological path calculation is implemented with Xeon Phi **vector intrinsics** for comparison.
- **Host system:** HP SL250s Gen8, 64 GB RAM, 2× Intel Xeon (E5-2670) CPUs @ 2.6-3.3 GHz, 16 threads each (Hyper-threading), 256-bit VPUs (AVX).
- **Coprocessors:** 2 Intel Xeon Phi cards (5110P) with 60 cores @ 1GHz, 240 threads and 8GB RAM each. 512-bit VPUs,

Speed up on Xeon Phi:

Here the focus is on the pure Xeon Phi performance, hence a single Xeon Phi card is compared to the host system (**Fig.3**).

- Ispc implementation show a **speed up of 1.65**
- Using Xeon Phi vector intrinsics (int.) increase performance by 10% (**total speed up to host: 1.85**)

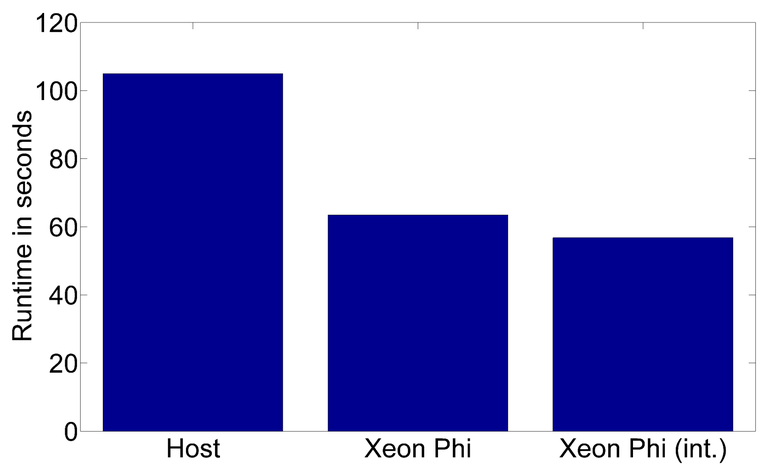

Fig. 3: Runtime on host and single Xeon Phi card.

Scalability:

- Host system scalability factor: 0.90 up to 16 threads and 0.17 and afterwards.
- Xeon Phi scalability factor: 0.83 up to 236 threads. Remaining 4 threads are reserved for memory mgt.
- Scaling behavior is shown in Fig. 4. Speed up is normalized to one host thread for better comparison, hence Xeon Phi scalability appears smaller in Fig. 4.

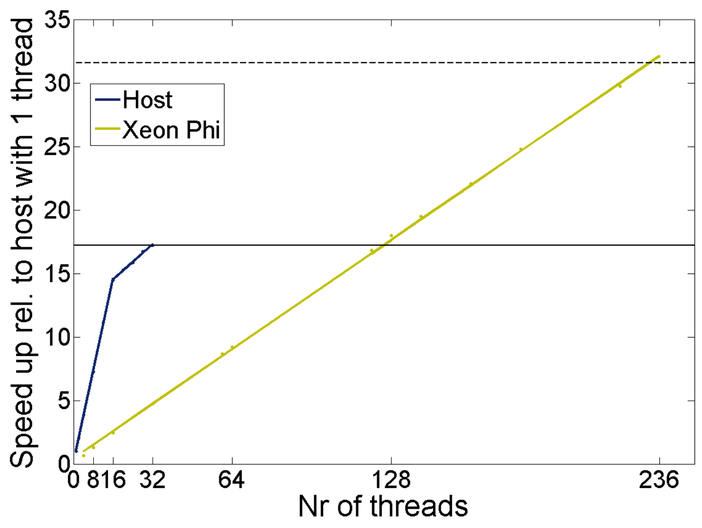

Fig. 4: Scalability graph, normalized to one host thread.

Key points and conclusion:

- Reasonable speed up can be achieved for computational hot spots of PET reconstruction by offloading workload to Xeon Phi coprocessors.
- Calculation of sensitivity map and back projection profit from vector processing capabilities of the Xeon Phi and show good scalability.
- Intrinsic vector programming of radiological path provides small additional speedup compared to (portable) programming with ispc.
- For scatter estimation also a runtime reduction is feasible, but it seems less suitable for acceleration by Xeon Phi, probably because of strong branching.

References:

[1] F. Jacobs et al. , “A fast algorithm to calculate the exact radiological path through a pixel or voxel space,” Journal of computing and information technology, vol. 6, no. 1, pp. 89–94, 1998.
[2] S. Woop, L. Feng, I. Wald, and C. Benthin, “Embree ray tracing kernels for CPUs and the xeon phi architecture,” in ACM SIGGRAPH 2013 Talks. ACM, 2013, p. 44.
[3] M. Pharr and W. R. Mark, “ispc: A SPMD compiler for high-performance CPU programming,” in Innovative Parallel Computing (InPar), 2012. IEEE, 2012, pp. 1–13.

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme FP7/2007-2013/ under REA grant agreement n° [317446] INFIERI "INtelligent Fast Interconnected and Efficient Devices for Frontier Exploitation in Research and Industry"