Monte Carlo Transport on Modern Architectures: Quicksilver at the LLNL COE

Ryan C. Bleile*†, Patrick S. Brantley*, Dave Richards*, Shawn A. Dawson*, Matthew J. O'Brien*, Scott McKinley *, Leopold Grinberg‡, Hank Childs†
Lawrence Livermore National Laboratory, †University of Oregon, Eugene, ‡IBM,

Abstract:

The Center of Excellence (COE) programs at Lawrence Livermore National Laboratory (LLNL) were established to promote code advancement and research through vendor interaction and support, in order to enable code groups to prepare for the next generation of supercomputers arriving over the next few years. One COE focus involves the Quicksilver mini-app that is representative of Monte Carlo transport. Over the past year, efforts using the Quicksilver mini-app have: performed baseline comparisons of a coarse-grained threading model that matches the threading model used in its parent production application, implemented an initial fine-grained threading model using OpenMP 3/4.5 features, and been shown to run concurrently on CPU and GPU hardware architectures. Further study into efficient parallel tasking models, tracking algorithms, tally collection, and applying these changes to the production level application are the goals of this next year's efforts.

The purpose of the Quicksilver mini-app is to create a freely distributable code base that represents the key elements of the Mercury workload by solving a simplified time-dependent neutron transport problem.

- Problem Setup
  - Mesh: 3D Polyhedral
  - Nuclear Data: Setup simplified reactions per material isotope
- Time Stepping
  - Cycle Initialize:
    - Particle Sources
    - Population Control
  - Cycle Tracking:
    - Particle Tracking
    - Communication
  - Cycle Finalize:
    - Tally Reductions

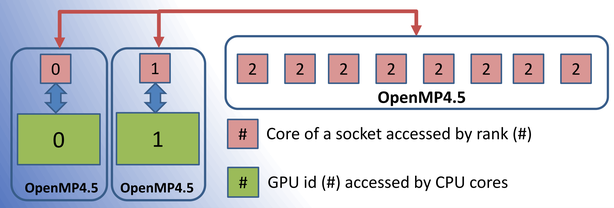
Course-Grain Threading Model:

- Threaded over particle vaults
- Tallies replicated per thread

Fine-Grain Threading Model:

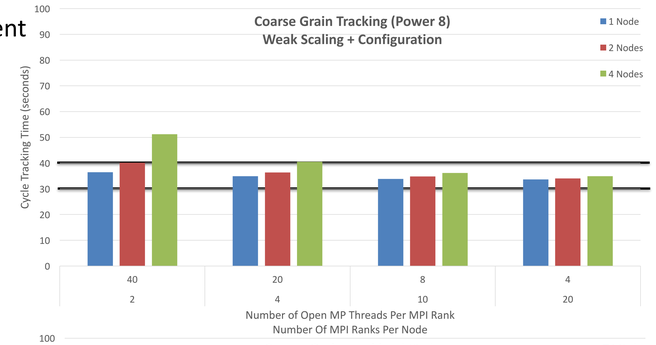
- Threaded over particles
- Tallies accessed through atomics
- CPU and/or GPU at runtime per rank

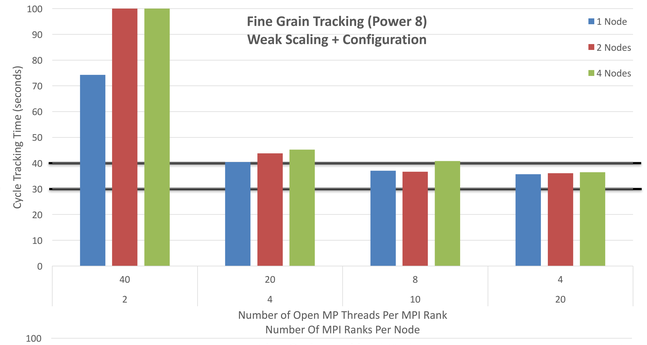
Parallelism Across an Entire GPU enabled Node

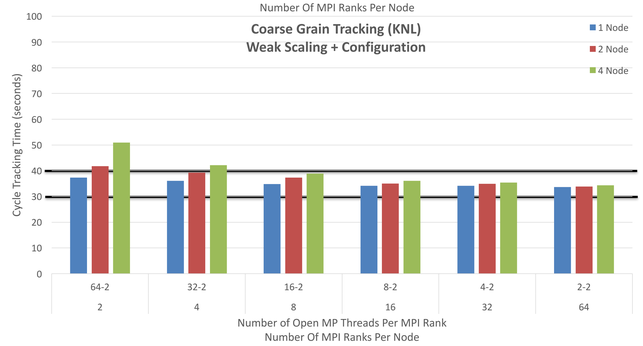


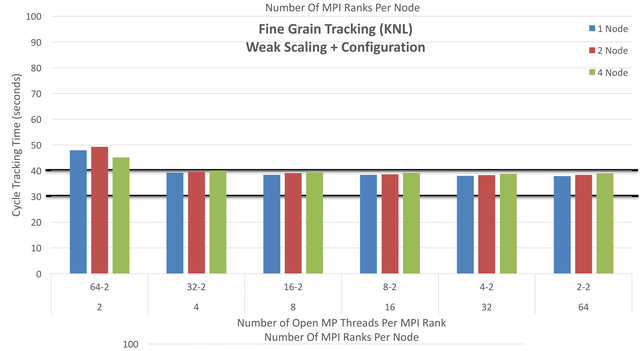
Performance Studies:

Understanding MPI/OpenMP Tradeoff and Weak Scaling

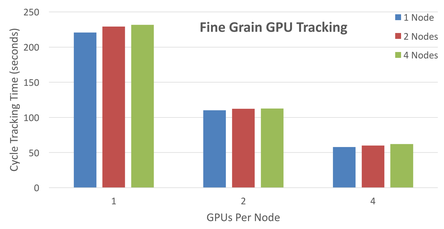


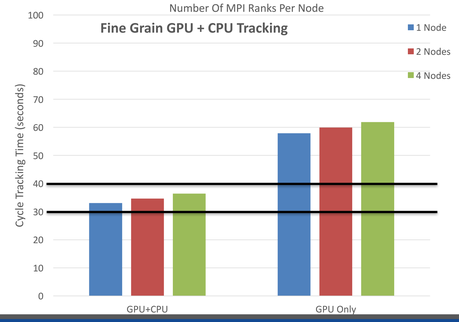






Strong and Weak Scaling for GPUs across Nodes





Future Work

Algorithm Development:

1. Study task scheduling approaches for organizing sourcing, tracking, and communication
2. Compare history-based and event-based approaches within the tracking kernel
3. Determine/Test approaches for dealing with atomics related to output tallies

References

1. Richards, David, et al. *Quicksilver*. No. Quicksilver; 004652WKSTN00. Lawrence Livermore National Laboratory (LLNL), Livermore, CA (United States), 2016.
2. Bleile, R. C., et al. "Investigation of Portable Event-Based Monte Carlo Transport Using the NVIDIA Thrust Library." *Trans. Am. Nucl. Soc* 114 (2016): 941-944.
3. Bleile, R. C., et al. *Algorithmic Improvements for Portable Event-Based Monte Carlo Transport Using the Nvidia Thrust Library*. No. LLNL-CONF-695977. Lawrence Livermore National Laboratory (LLNL), Livermore, CA, 2016.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.
Funding was provided by the LLNL Livermore Graduate Scholar Program.
IM Review and Release number: LLNL-POST-729700

Lawrence Livermore National Laboratory