New Data-parallel Algorithms Accelerate Cosmology Data Analysis on GPUs

Objectives

Milestone

- Implement application-specific visualization and/or analysis operators needed for in-situ use by LCF science codes
- Use PISTON to take advantage of multi-core and many-core technologies

Target Application

- The Hardware/Hybrid Accelerated Cosmology Code (HACC) simulates the distribution of dark matter in the universe over time
- An important and time-consuming analysis function within this code is finding halos (high density regions) and the centers of those halos

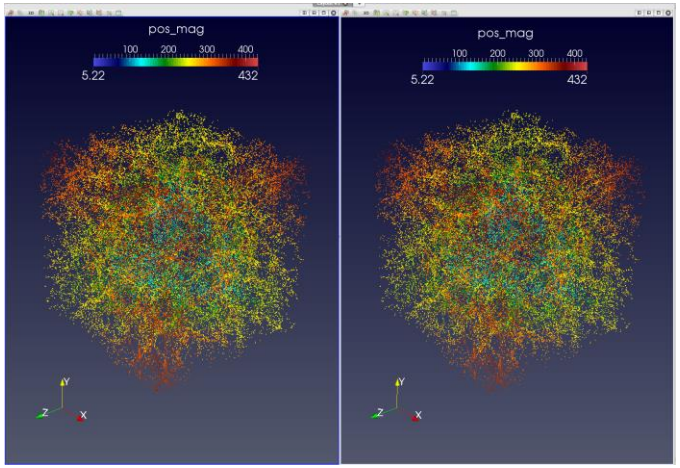
Impact

VTK-m framework

- The PISTON component of VTK-m develops data-parallel algorithms that are portable across many-core architectures for use by LCF codes
- PISTON consists of a library of visualization and analysis algorithms implemented using Thrust, and our extensions to Thrust

Halo and Center Finders

- Data-parallel algorithms for halo and center finding implemented using VTK-m (PISTON) allow the code to take advantage of parallelism on accelerators such as GPUs
- Can be used for post-processing or in-situ, with in-situ integration directly into HACC or via the CosmoTools library



Visual comparison of halos computed by the original HACC algorithms (left) and the PISTON algorithms (right). The results are equivalent, but are computed much more quickly on the GPU using PISTON.

Accomplishments

Performance Improvements

- On Moonlight with 1024^3 particles on 128 nodes with 16 processes per node, PISTON on GPUs was **4.9x** faster for halo + most bound particle center finding
- On Titan with 1024^3 particles on 32 nodes with 1 process per node, PISTON on GPUs was **11x** faster for halo + most bound particle center finding
- **Portability of PISTON** allowed us to also run our algorithms on an Intel Xeon Phi
- Implemented grid-based most bound particle center finder using a Poisson solver that performs fewer total computations than standard $O(n^2)$ algorithm

Science Impact

- These performance improvements **allowed halo analysis to be performed** on a very large **8192^3** particle data set across 16,384 nodes on Titan **for which analysis using the existing CPU algorithms was not feasible**

Publications

- Submitted to SC14: "Utilizing Many-Core Accelerators for Halo and Center Finding within a Cosmology Simulation" Christopher Sewell, Li-ta Lo, Katrin Heitmann, Salman Habib, and James Ahrens