GA

TN

MP

CI

ASTROPHYSICS ON GRAPHICAL PROCESSING UNITS

Allocation: Illinois/230 Knh
PI: Hsi-Yu Schive[1]
Co-PI: Matthew Turk[1]
Collaborators: John ZuHone[2], Nathan Goldbaum[3], Massimo Gaspari[4], Chin-Yu Chen[1], Ui-Han Zhang[5]

[1]University of Illinois at Urbana-Champaign
[2]Harvard-Smithsonian Center for Astrophysics
[3]National Center for Supercomputing Applications
[4]Princeton University
[5]National Taiwan University

EXECUTIVE SUMMARY

GAMER is a GPU-accelerated adaptive mesh refinement (AMR) code for astrophysics. In this project, we have incorporated a rich set of new physics and functionality into the code, including particles, magnetic field, chemistry and radiative processes, star formation, adaptive time-stepping, memory pool for efficient memory reuse, bitwise reproducibility, and optimization on load balancing. To demonstrate the capability of GAMER, we directly compared it with the widely adopted code Enzo [2] on a realistic astrophysical application, i.e., star formation in an isolated disk galaxy. This comparison showed that GAMER not only produces physical results that are very consistent with Enzo but also runs significantly faster, by an order of magnitude. In addition, we demonstrated good weak scaling with a parallel efficiency of 60%–70% and an overall performance of $5x10^{10}$ cells/sec using 4,096 GPUs and 65,536 CPU cores simultaneously for both hydrodynamic and magnetohydrodynamic applications.

RESEARCH CHALLENGE

Efficient utilization of the computing power of heterogeneous CPU/GPU supercomputers like Blue Waters remains extremely challenging for astrophysical AMR codes, especially when targeting realistic astrophysical applications requiring a rich set of physics and when running with a large number of nodes. The main challenge lies in the fact that the performance bottleneck, in general, does not reside in a single computational task. Instead, multiple tasks can be comparably expensive, for example, hydrodynamic solver, Poisson solver, chemistry and radiative processes, particle update, and AMR data structure reconstruction. The improvement in overall performance using GPUs is thus usually very limited if only certain specific tasks are ported to GPUs. On the other hand, not all tasks are GPU-friendly. Moreover, the GPU memory is still much smaller than the CPU's main memory. Therefore, it is necessary to efficiently utilize both CPU and GPU resources and carefully distinguish GPU-friendly and GPU-unfriendly tasks. Last, but not least, load imbalance and MPI communication may also become the performance bottleneck, especially if the computation time is significantly reduced by GPU utilization.

METHODS & CODES

According to the above, in this project we have significantly revised the GAMER code to incorporate a rich set of new physics, optimized the overall performance and parallel scalability, and directly compared it with other AMR codes on realistic astrophysical applications. The major features of GAMER include:

- *Multiphysics*. GAMER supports a rich set of physics, including hydrodynamics, magnetohydrodynamics, self-gravity, dark matter particles, chemistry and radiative processes, and star formation.
- *Adaptive mesh refinement* for resolving a large dynamic range.
- *Hybrid OpenMP/MPI/GPU*. GAMER uses GPUs as PDE solvers and CPUs to manipulate the AMR structure. It adopts OpenMP for intra-node parallelization in CPUs, MPI for inter-node communication, and CUDA as the GPU programming interface.
- *Asynchronous operations*. GAMER utilizes overlapping CPU computation, GPU computation, and CPU-GPU communication to boost the performance further.
- *Hilbert space-filling curve* for load balancing.
- *Memory pool* for efficient reuse of CPU memory.
- *Bitwise reproducibility* for scientific reproducibility.

GAMER uses the publicly available library GRACKLE [1] for the chemistry and radiative processes.

RESULTS & IMPACT

Fig. 1 shows the weak scaling of GAMER for the AMR-enabled hydrodynamic and magnetohydrodynamic (MHD) simulations on Blue Waters using up to 4,096 XK nodes. For the hydrodynamic test, we simulated a three-dimensional Kelvin–Helmholtz instability using the corner transport upwind scheme, piecewise parabolic reconstruction, and Roe's Riemann solver. For the MHD test, we simulated the Arnold–Beltrami–Childress (ABC) flow using the same hydrodynamic scheme but with the constraint transport technique to reinforce the divergence-free constraint on the magnetic field. GAMER achieves a parallel efficiency of 60%–70% and an overall performance of $5x10^{10}$ cells/sec. Note that when using, for example, 4,096 XK nodes, GAMER utilizes

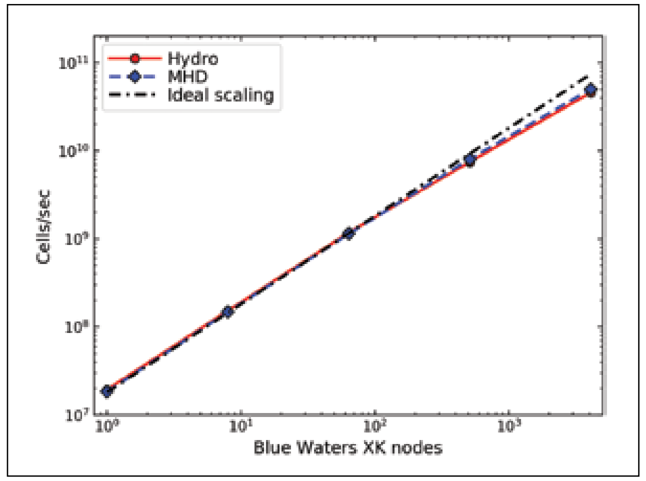

Figure 1: Weak scaling of GAMER for the AMR-enabled hydrodynamic and MHD simulations using up to 4,096 XK nodes. Thanks to the hybrid CPU/GPU parallelization, GAMER can use both 4,096 GPUs and 65,536 CPU cores simultaneously and achieve a parallel efficiency of 60%–70% and an overall performance of $5x10^{10}$ cells/sec.

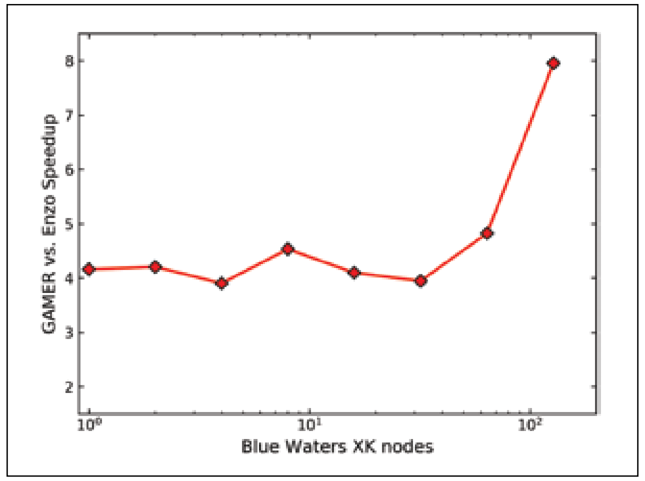

Figure 2: Performance speedup of GAMER over Enzo on the strong scaling of isolated disk galaxy simulations. GAMER outperforms Enzo by factors of four to eight. More importantly, the speedup increases when using more than 32 XK nodes, suggesting that GAMER also exhibits better parallel scalability.

both the 4,096 GPUs and the 65,536 CPU cores thanks to the hybrid OpenMP/MPI/GPU parallelization.

Fig. 2 shows the performance speedup of GAMER over Enzo for strong scaling using 1–128 XK nodes. Note that instead of using a simple test problem for the comparison, we conducted isolated disk galaxy simulations that combined gas, stars, and dark matter and incorporated hydrodynamics, self-gravity, particle dynamics, chemistry, radiative cooling/heating, and star formation. This comparison thus provides very convincing demonstrations of the accuracy and performance of GAMER. First, we found that the physical results obtained by the two codes are in very good agreement. Further, GAMER was measured as being four to eight times faster than Enzo. More importantly, this speedup ratio is roughly a constant when using a smaller number of nodes and increases when using more than 32 nodes, suggesting that GAMER not only runs faster, but also scales better. This significant performance improvement will allow us to study the dynamical evolution of the interstellar medium in unprecedented detail.

We have also compared GAMER with the AMR code FLASH [3] on cluster merger simulations and showed that GAMER is two orders of magnitude faster. Last, but not least, we have demonstrated that, for the uniform-resolution simulations, GAMER can achieve a maximum resolution of $10,240^3$ and $8,192^3$ for hydrodynamics and MHD, respectively, when using 4,096 XK nodes.

WHY BLUE WATERS

Blue Waters provides a unique opportunity for us to test and optimize GAMER on an extreme scale using thousands of GPUs and tens of thousands of CPU cores simultaneously. Moreover, the Blue Waters staff is very efficient in helping solve technical issues such as fine-tuning the hybrid OpenMP/MPI model. Without their help, it would have been very difficult to achieve an optimal throughput.

PUBLICATIONS & DATA SETS

Zhang, U-H., H-Y, Schive, and T. Chiueh, Magnetohydrodynamics with GAMER. *The Astrophysical Journal Supplement Series*, 236 (2018), DOI:10.3847/1538-4365/aac49e.

Schive, H-Y., et al., GAMER-2: a GPU-accelerated adaptive mesh refinement code—accuracy, performance, and scalability. Submitted to *Monthly Notices of the Royal Astronomical Society*, arXiv: 1712.07070 (2017).