Petascale Molecular Dynamics on Roadrunner

Timothy C. Germann, Kai Kadau, T-1; Sriram Swaminarayan, CCS-2

The scalable parallel short-range molecular dynamics (SPaSM) code was developed in the early 1990s by David Beazley, Peter Lomdahl, and collaborators at LANL to simulate the dynamical behavior of materials under extreme conditions. Over the past decade SPaSM has been successfully used to provide insight into dislocation dynamics in metals [1], shock compression-induced plasticity [2] and phase transformations [3], and hydrodynamic instabilities [4,5]. These phenomena are important in aspects of many applications that range from inertial confinement fusion and nuclear weapons performance to supernovae explosions. A major breakthrough occurred recently when the SPaSM code was successfully used to perform the fastest molecular dynamics simulation to date, reaching ~369 TFlop/s on the full Roadrunner machine.

Molecular dynamics (MD) is a computational approach that embodies the concept of scientific reductionism—that is, the idea that the nature of a complicated process can be understood by understanding the interaction of the parts. In this case, the parts are the individual atoms and MD works by first determining the fundamental interactions between the atoms (from experimental data and theoretical insights into the interatomic force description) and then building a picture of how a macroscopic material should behave by simulating the interactions of millions to billions of atoms. (Similarly, national-level mitigation strategies for pandemic influenza spread can be assessed by simulating the interactions between the ~300 million people within the United States [6].)

The original SPaSM code design was targeted at massively parallel computer architectures, and written in an era when memory and arithmetic operations were the typical performance bottlenecks, while interprocessor communication was relatively cheap. The efficiency of this original SPaSM implementation is demonstrated by its use in earning an IEEE Gordon Bell Performance Prize in 1993 (50 GFlop/s on the CM-5), an IEEE Gordon Bell Price/Performance Prize in 1998 ($15/Mflop/s on the Avalon Beowulf cluster), and an IEEE Gordon Bell finalist in 2005 (50 TFlop/s on BlueGene/L). With the arrival of Roadrunner the original design considerations were wholly obsolete: memory is plentiful and computation on the IBM Cell Broadband Engine (Cell BE) synergistic processing unit (SPU) cores is virtually free, but memory accesses to and from the SPUs are the bottleneck. We have rewritten almost the entire SPaSM code base, particularly its communication infrastructure and data structures, to accommodate the Cell BE processor and allow for better asynchrony between the Opteron and Cell BE processors. The resulting code demonstrates excellent parallel scaling, reaching 369 TFlop/s on the full 17 connected units (CUs) of Roadrunner (see Fig. 1), and was an IEEE Gordon Bell Prize finalist again in 2008.

The key benefit of Roadrunner for atomistic simulations is this remarkable gain in speed, not just size. This gain in speed is because large-scale parallel MD codes, such as SPaSM, have typically been compute-bound rather than memory-bound—while one trillion atoms (in single precision) can fit into memory on either Roadrunner or the 212,992-CPU BlueGene/L platform at LLNL, few phenomena

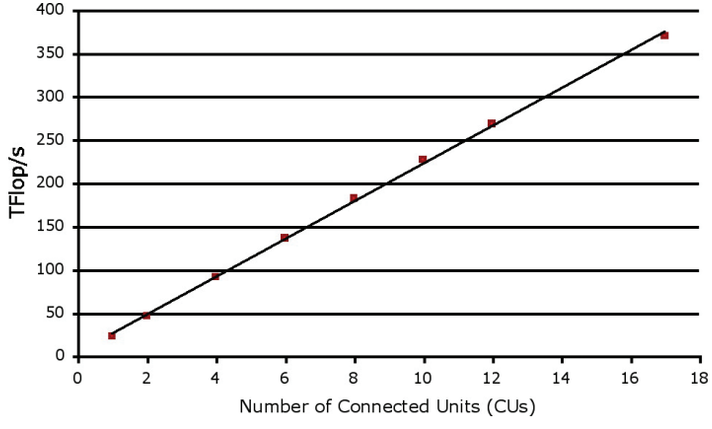

Fig. 1. Weak scaling of SPaSM on Roadrunner for the benchmark Lennard-Jones (LJ) problem, measured over 20 time steps. The symbols are measured data while the solid line is the best linear fit through all the data.

occur within the picosecond timescale accessible by such MD simulations. The order-N linear scaling down to as few as several thousand atoms per processor allows us to sacrifice system size for longer simulation times, until the communication overhead begins to dominate for smaller system sizes. For this reason, recent extended (day- to week-long) full-system runs on BlueGene/L by us and another group at LLNL have independently chosen the "sweet spot" around 1 billion atoms for 1 ns, the longest time scale accessible before performance degradation sets in on BlueGene/L. (The basic time step in MD simulations must be short enough to resolve atomic vibrations, typically 1–5 femtoseconds, so up to a million time steps may be required to reach ns timescales.) However, longer times are clearly essential for many phenomena—1 ns is barely enough time for the round-trip traversal of a sound or shock wave through a micron-long sample. Similarly, a model of the entire world population of 6.6 billion people can readily fit within many existing supercomputers, but the SPaSM-based EpiCast code has thus far been limited to the US population (300 million people) due to the computational time required (several minutes to simulate a single day of disease progression). The acceleration that is provided by the Cell BE processors in Roadrunner directly addresses this important time scale issue, enabling unprecedented insight into materials, fluid, and infectious disease dynamics.

The materials science community is extremely excited about the discovery opportunities now presenting themselves with the Roadrunner-class petascale computers. Our initial focus will be in understanding how shock waves can create material damage and ejecta (see Fig. 2) and how hydrodynamic instabilities initiate and evolve. Subsequent studies will investigate a major yet-unsolved question, namely how to simulate polycrystal grain effects at the relevant length and time scales. This is important because grain effects are known to play an essential role in the deformation and failure of metallic materials under extreme loading conditions, such as in nuclear weapons, nuclear reactors, bridges, and buildings. A petascale computer like Roadrunner is essential to simulate the grains at the proper scale and duration.

For further information contact Timothy C. Germann at tcg@lanl.gov.

[1] S.J. Zhou et al., *Science* **279,** 1525-1527 (1998).
[2] B.L. Holian, P.S. Lomdahl, *Science* **280,** 2085-2088 (1998).
[3] K. Kadau et al., *Science* **296**, 1681-1684 (2002).
[4] K. Kadau et al., *Proc. Natl. Acad. Sci. (USA)* **101,** 5851-5855 (2004).
[5] K. Kadau et al., *Proc. Natl. Acad. Sci. (USA)* **104,** 7741-7745 (2007).
[6] T. C. Germann et al., *Proc. Natl. Acad. Sci. (USA)* **103,** 5935-5940 (2006).

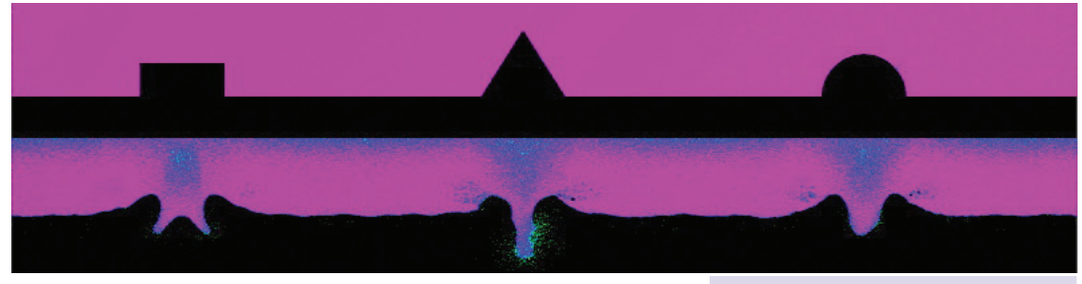

Fig. 2. Demonstration run showing the ejection of material from a shocked copper surface, run on 1 CU. The sample is 3-D, but with a thin cross section into the plane—only a small section of the total sample height is shown, at two different times during the simulation. (Top) Free surface prior to shock arrival. (Bottom) Initial jet formation after a shock wave has reflected from the free surface. Three different surface defects are initially present, with the same volume but different shapes, leading to different jetting patterns that will subsequently break up into ejecta particles with different size and velocity distributions.

Funding Acknowledgments
DOE, NNSA, Advanced Simulation and Computing Physics and Engineering Models Program