LAMMPS Strategic Development Plan

Authors: Aidan Thompson, Steve Plimpton, Paul Crozier, and John Aidun

Date: 12/13/2012

# Purpose

LAMMPS is a classical molecular dynamics (MD) simulation code developed within Center 1400. It is widely used at Sandia, other energy and defense labs, universities, and industry to perform large-scale atomistic simulations of to investigate material behavior. It is designed to run efficiently on large parallel computing platforms. LAMMPS is distributed by Sandia under a GNU Public License. The continued development of LAMMPS is supported directly and indirectly by internal Sandia projects and code contributions from external collaborators. This model of incremental improvement and improvisation has worked quite well, but it has also limited our ability to develop the code according to a long-term strategy. LAMMPS is dominant in many areas of materials science, chemistry, and physics. It is widely used in many other fields such as molecular biophysics and granular mechanics. Because of this success, there is constant demand from users for improved serial performance, efficient scaling on advanced architectures, more interatomic potentials, more pre-, post- and on-the-fly analysis capabilities, etc. Demand far outstrips available development resources. At the same time, Sandia has specific mission areas where LAMMPS can and should play an important role. It is essential that we be strategic in identifying those areas where improvements will have the most impact. This will guide our efforts to seek funding and invest resources in those strategic areas.

The purpose of this document is to identify LAMMPS development areas of greatest strategic importance. It is divided in to three parts. The first part is a survey of the current environment for LAMMPS development. This consists of the Sandia missions areas that drive the continued development of LAMMPS. It also lists currently supported code development activities. The second part is a survey of possible LAMMPS development areas. The final part gives a detailed description of areas where Sandia mission drivers overlap with development needs.

# Part I: Current Development Environment

## Materials Modeling Drivers

While LAMMPS is widely used at Sandia and elsewhere, many of these applications are driven by basic science questions rather than solving specific technological problems.  Sandia’s mission drivers belong primarily to the latter category.  In the past, LAMMPS simulations have tended to play a supporting role in Sandia projects, providing qualitative understanding of physical phenomena. Examples include physical stress relaxation in polymers, diffusive transport in microporous materials, and frictional wear in MEMS devices. In more recent years, LAMMPS simulations have tended to target specific materials and provide quantitative predictions of specific properties and phenomena. A few of the most important area are listed below:

* Radiation damage in semiconductors
* Mechanical response of polycrystalline materials
* Energetic materials
* Dynamic materials response
* Electronic device modeling
* Granular materials

MD fulfills two important materials modeling roles. Firstly, MD functions as a bridge between continuum and *ab initio* atomistic models. In continuum models, material properties are treated as field variables that evolve according to partial differential equations that capture key physical and chemical processes. In ab initio models, for small assemblies of atoms, the energy or free energy of non-core electrons are calculated as a function of the atomic positions. This enables the determination of stable configurations of atoms, as well as classical trajectories of atoms. Continuum models rely on very strong assumptions about types and magnitudes of the important processes. In the absence of suitable experimental data, these models must be validated against more rigorous models. Ab initio models make only a few well-understood approximations, are very accurate and reliable for many types of materials, but are computationally expensive. They are limited to simulations involving 101-103 atoms and 102 -103 configurations, and so cannot be directly compared with continuum calculations, except in highly idealized situations. A single MD model can be used to be calculations directly compared to both ab initio and continuum models. In this way, MD provides the missing link needed to validate continuum models against ab initio calculations.

Secondly, MD simulations can describe atomistic phenomena that emerge on large length- and times-scales, which cannot be captured by either ab initio or continuum models. For example, in MD simulations of energetic materials, dissipative structures such as shear banding can be observed. These structures are too large to be calculated using ab initio methods, and are completely absent from continuum models.

## Computer Science Drivers

MD is computationally demanding and a heavy consumer of computing resources. At the same time, the computational kernel is simple and unambiguous. Reliable benchmarks for speed and accuracy exist. This makes it both worthwhile and relatively straight-forward to apply new algorithms and run on advanced computer hardware. For this reason, all the most obvious ideas for performance improvement have been discovered, and re-discovered. Nonetheless, the emergence of extreme scale parallelism, heterogeneous architectures, and other factors have created new opportunities and challenges that are driving the development of new MD algorithms. These include:

* GP-GPU processor support
* Thread-based parallelism for multicore machines
* Reduced memory bandwidth for short-range forces
* Reduced communication parallel FFTs
* Alternative long-range electrostatic solvers
* Reduced energy consumption per atom-timestep
* Exascale file I/O

All of these areas are suitable candidates for co-design i.e. the identification of optimal combinations of hardware features and MD algorithms. In addition to these hardware-related areas, there are purely software challenges, including:

* Automated interatomic potential development
* Data analysis and visualization
* Software quality assurance
* Uncertainty quantification

## Current Development Support

Sandia LAMMPS-related projects receive quite a lot of support from various sources. However, only a small fraction of this is available for code development.  Here is a list of areas where code development support currently exists:

* Quantum-accurate potentials
* Scalable electrostatics algorithms
* GPU Support
* Reactive dynamics of energetic materials

The appendix lists current and past project support.

# Part II: LAMMPS Development Areas

Below is a list of possible LAMMPS development areas, divided into different technical areas.

## Performance

* 1. Single processor
     1. Kernel optimization
     2. Cache-efficiency
     3. Data locality
  2. Parallel
     1. Dynamic load-balancing
     2. Scalable long-range electrostatics
  3. Next-generation hardware
     1. GPU processing
     2. Multicore processing
     3. Co-processor support
     4. Heterogeneous architectures
     5. Multithreading
     6. Co-design
     7. I/O systems
     8. Fault-tolerance

## Physics and Chemistry Models

* 1. Interatomic potentials
     1. Implementations
     2. Interfaces
     3. Development
     4. Automated quantum-accurate potentials
     5. Genetic programing potentials (GP)
     6. Spectral neighbor analysis potentials (GAP, SNAP)
  2. Special-purpose models
     1. Electrons (wave-packet, entropy, TTM)
     2. Reactive
     3. Variable charge
     4. Polarizable (point dipoles)
     5. Long-range dipole-dipole interactions
     6. Coarse-grain
     7. Granular (contact force models, non-spherical particles, etc.)
     8. Peridynamics
     9. Solvation
  3. Coupling to other models
     1. Continuum
     2. Quantum
     3. Mesoscale
     4. Multi-scale

## Simulation Methods

* 1. Accelerated Sampling Methods
     1. Multi-replica methods
     2. Energy minimization
     3. Monte Carlo methods
     4. Ensembles, constraints
     5. Accelerated MD
  2. Monte Carlo Methods
     1. NVT ensemble
     2. NPT ensemble
     3. VT ensemble
     4. Gibbs ensemble
     5. GCMD simulation
     6. Kinetic Monte Carlo
     7. Direct Simulation Monte Carlo
     8. Coupling to other Monte Carlo code(s)
  3. Numerical Algorithms
     1. Linear solvers
     2. Energy minimization

## Set-up and Analysis Tools

* 1. Builders
  2. Diagnostics and post-processing
  3. Format inter-conversion
  4. Python tools
  5. Visualization
  6. Spatial profile sampling
  7. Statistical analysis
  8. Data/workflow management
  9. Uncertainty quantification
  10. Non-standard model support
      1. Shock physics support
      2. Polymer support
      3. Biophysics support
      4. Electronic devices support

## Software Quality

* 1. Software quality assurance
  2. Regression test suite
  3. Verification and unit testing
  4. Model validation

## User support

* 1. Documentation
  2. Website
  3. E-mail list, help desk
  4. Workshops, tutorials
  5. Building and platform support
  6. Bug fixes & feature requests

# Part III: Strategic Development Areas

Based on the above development areas and the needs of Sandia mission drivers, the following areas have been identified as strategically important:

1. **Scalable algorithms for exascale computing:** Currently, simulations involving 106 atoms and 102 processors are routine. LAMMPS performs well in this regime. However, there is a modeling and simulation need simulations on involving up to 109 atoms, in order to directly compare MD with continuum models, and also to identify processes that are only observable on these scales. At the same time, computing resources are moving towards processors counts in excess of 105. It is essential that LAMMPS perform efficiently in this regime. Many of basic problems that arise in this regime have already been addressed in LAMMPS (parallel file I/O, 64-bit counters). Unresolved issues include:
   1. Long-range electrostatics
   2. Charge equilibration
   3. Fault tolerance
   4. Heterogeneous load-balancing
2. **Automated interatomic potential development:** Most potentials in LAMMPS are implemented in a form that is transparent: functional forms are explicitly hand-coded, specific parameter values are read in from a file hand-entered parameter file. With the emergence of machine learning methods (genetic algorithms, genetic programming, Gaussian process, neural networks, KIM), neither the functional forms nor the parameter values can be modified by hand. Instead, these are in turn generated by other software and stored in complex data structures. This presents a number of issues, including
   1. Efficient general algorithms for evaluating machine-learning potentials
   2. Work-flow systems that keep track of which potential is to be used, and was used.
   3. Using LAMMPS to evaluate the large training set for many candidate potential

1. **Large-scale simulation of advanced electronic devices:** Large-scale MD simulations of electronic materials (memristor, graphene) can help discover new science leading to materials breakthroughs, in partnership with other Centers. This will contribute to the Center 1400 campaign goal to develop computing breakthroughs via collaborations in architectures, algorithms, materials, microelectronics, and physics.”
2. **Large-scale simulation of radiation damage in III-V semiconductors:** Large-scale MD simulation of radiation damage in III-V semiconductor films can have a significant impact on Sandia’s QASPR program. This will contribute to the Center 1400 campaign goal to increase the impact of distinguishing 1400 technologies and expertise on NW problems.
3. **Simulation of Granular Materials** The GRANULAR package in LAMMPS and the LIGGGHTS package built on top of LAMMPS allow simulation of large particulate systems. Granular simulations, also known as Discrete Element Simulations, differ from conventional MD in that the particles have rotational momentum, no just translational momentum. This allows for the possibility of dynamic friction between particles. Granular simulations can capture behaviors of particulate systems that are not achievable using continuum mechanics models.
4. **Monte Carlo Simulation** LAMMPS now contains a simple grand canonical Monte Carlo simulation capability. We may also couple LAMMPS with another Monte Carlo code such as CASSANDRA (Notre Dame U.)
5. **Code Coupling** LAMMPS is designed to be modular, and can be called by others codes as a library. This provides many opportunities to couple with different codes. Conversely, it is relatively simple to create an interface in the form of a pair style to interface with another code that can provide forces and energy of a configuration of atoms.
6. **Partnering Opportunities** Industry (Materials Design, Inc., Global Foundries), academia (Vanderbilt U.), DOE, DoD.
7. **Miscellaneous** Advisory board, privacy policy, ...

# Appendix A: Historical LAMMPS Support

## Current Support

(1) Aidan Thompson (PI): "Automated Algorithms for Quantum-Level Accuracy in Atomistic Simulations", EPS LDRD FY12-FY14.  Use quantum data to generate spectral neighbor analysis potentials (SNAP) for mission-relevant materials. Develop better-scaling algorithms for long-range electrostatics.

(2) Aidan Thompson (PI): “Atomistic Energetic Modeling”, ASC P&EM FY09-current. Apply LAMMPS ReaxFF capability to identify reaction pathways in initiation of energetic materials (chemical species analysis).

(3) Aidan Thompson (PI): “Predictive Modeling of Non-Ideal Explosives”, HSD LDRD FY11-FY13. Apply LAMMPS ReaxFF capability to identify reaction pathways in initiation of non-ideal explosives. (MSST method)

(4) Paul Crozier: “Advanced Nuclear Energy Fuel Cycle Waste Forms”, Office of Nuclear Energy. Build Grand Canonical Monte Carlo (GCMC) support in LAMMPS, including the ability to model molecular gas uptake in microporous materials.

## Core Support

4 LDRDs that contributed to substantial code development in LAMMPS:

(1) Greg Wagner (PI): "Equation-Free Simulation Methods for Multiple Timescale Diffusion Processes in Solids", EPS FY09-FY11.  Use equation-free methods to extend atomistic simulations to the timescales necessary for accurate simulation of diffusion processes. Led to accelerated timescale methods (PRD,TAD,NEB) in LAMMPS.

(2) Paul Crozier (PI); "Substructured Multibody Molecular Dynamics", CIS 2004-2006.  Develop, implement, and optimize substructured, multibody, molecular dynamics (SMMD) for simulation of systems that are currently beyond the capability of all-atom molecular dynamics.

(3) Mark Stevens (PI): "Modeling Biomembranes", BioTech 2003-2005. Coarse-grained models including dipole interactions in LAMMPS.  Let to development of C++ version and open-source release.

(4) John Hamilton (PI): "From Atom-Picoseconds to Centimeter-Years in Simulation and Experiment", Matl Sci 1999-2001.  Development of metal potentials and stress/strain capabilities in LAMMPS.

## Additional LDRD Support

LDRD projects that used LAMMPS as a tool: (Listed person was either the PI or a participant)

(1) Jeremy Templeton: "Atomistic-to-continuum modeling for nano- and micro-fluidics", EPS FY11-13.  Extending the existing AtC method to be appropriate for coupled atomistic/continuum models of fluid transport with the key technological driver being enhanced modeling capabilities for electrical energy storage devices.  As such, we are not only adding mass/momentum/energy coupling with appropriate versions of the Navier-Stokes equations, but also adding in multiphysics capabilities around electron transport and complex molecules.

(2) Carl Hayden: "Theoretical and Experimental Characterization of the Electric Double Layer", NTM FY11-13.  This project combines high fidelity experiments with quantum and MD calculations to improve our fundamental understanding of the electric double layer.  MD simulations of the experimental configurations will be performed to both validation our computer models as well as providing additional insights into the physics present in the experiments.

(3) Amalie Frischknecht: Effects of Morphology on Ion Transport in Ionomers for Energy Storage, EPS FY10-FY12.  We're studying morphology and transport properties of ionomers, polymers with a small number of ionic groups.  We're doing both coarse-grained and atomistic simulations using LAMMPS.

(4) Reese Jones: "Improving the contact resistance of nanoscale contacts", NTM 2010-2012.  Using LAMMPS to look at the resistivity of metal-semiconductor interfaces.

(5) Gary Grest: "Drying/Self-Assembly of Nanoparticle Suspensions" FY10 (late start)

(6) Stewart Griffiths: "Density Functional Theory Models of Ion Transport in Nanochannels", ERN FY10-12.  Develop the capability to simulate ion transport in nanochannels with both accuracy and fidelity, a classical density functional theory model is being developed.  Validation of the DFT model is based on comparison with MD calculations that explicitly resolve the atomistic dynamics.

(7) Jeremy Templeton: "Enhanced Molecular Dynamics for Simulation of the Solid-Electrolyte Interphase Layer", EPS 2010 (late start).  The project developed a solver for inhomogeneous (i.e. non-periodic) electric fields driven by atomic charges and applied boundary conditions.  We used the AtC framework to implement this and tested it on several problems related to electric double layers.

(8) Aidan Thompson: "MD Simulation of Ramp Loading", SEE FY09-FY11. Shock physics modeling, Hugoniostat method, new potentials for Be, Al. (SHOCK package)

(9) Jon Zimmerman: "Enhanced Molecular Dynamics for Simulating Porous Interphase Layers in Batteries", EPS FY09.  Develop a coupled atomistic-continuum framework for simulations of charge transport processes in materials.

(10) Jon Zimmerman: "Modeling Ramp Compression Experiments using Large-Scale Molecular Dynamics Simulation", SEE FY09-FY11.  Develop new interatomic potentials, ramp MD simulation approaches, and atomistic-to-continuum metrics to study ramp compression of materials.

(11) Mike Chandross: "Simulations of Nanomanufacturing Processes", 2008-2011.  Looking at the processes related to defect creating in large-scale nanomanufacting process like Step-Flash Imprint Lithography.  Also looking at nanoparticle drying.

(12) Paul Crozier: "Physics of Intense, High Energy Radiation Effects", SEE 2008-2010.  Explore the effects of short pulses of high dose-rate ionizing radiation on materials such as insulators and semiconductors. Use physics-based simulations linked to experiments done at Sandia and other facilities to explore the phenomena.

(13) Reese Jones: "Enhancing molecular dynamics with electron transport", EPS 2008-2010.  Coupled LAMMPS to PDEs of charge transport to do device-sized simulations with atomic detail.

(14) Andy Armstrong: "Impact of Defects on the Electrical Transport, Optical Properties and Failure Mechanisms of GaN Nanowires", NTM 2008-2010.  Used LAMMPS to explore the thermal conductivity of nanodevices with & without defects. With student at Berkeley I also developed Green-Kubo & Evans methods for transport coefficient estimation in LAMMPS

(15) Jack Houston: "Studies of the Viscoelastic Properties of Water Confined between Surfaces of Specified Chemical Nature", FY08-10.

(16) Jon Zimmerman: "Development of Advanced Continuum Models that Incorporate Nanomechanical Deformation into Engineering Analysis", EPS FY07-FY08.  Use atomistic simulations to develop a consistent theory to model both the evolution of disclinations and their kinetics.

(17) Tim Boyle: "Interfacial Property Control of Elastomeric Nanocomposites", FY07-09.  Studied the efffect of coating density and functionalization on the interactions between nanoparticles.

(18) Xiaowang Zhou: "Atomically engineering Cu/Ta interfaces", EPS 2007 (late start).  Study the formation of surface roughness, grain structure, misfit dislocation, and amorphization during vapor deposition of Cu/Ta films.

(19) Xiaowang Zhou: "Effect of nanoscale patterned interfacial roughness on interfacial toughness", EPS 2007.  Study the interfacial toughness as a function of patterned, nanoscale roughness of devices, and LAMMPS is used to derive cohesive zone law for fracture simulations.

(20) Stephen Foiles: "Grain growth in nanocrystalline metals", NTM 2007-2009 (given an LDRD award).  LAMMPS used to study grain boundary mobility and grain growth in nanocrystalline metals, with the goal of improving the thermal stability.

(21) Randy Schunk: "Rheology and drying of nanoparticle suspensions", EPS/NTM 2006-2008.  Initial coarse graining of atomistic potentials and addition of SRD solvent hydrodynamics.

(22) Jon Zimmerman: "Effect of Nanoscale Patterned Interfacial Roughness on Interfacial Toughness", EPS FY05-FY07.  Perform atomistic simulations to identify possible modifications to the interfacial separation models currently used in nanoscale, finite element fracture analyses.

(23) Jon Zimmerman: "Fundamental Enabling Issues in Nanotechnology: Stress at the Atomic Scale", EPS FY05-FY07.  Combine experiment and modeling approaches to study stress evolution during thin film growth, as relevant to films used in nano- and microscale technology.

(24) Paul Crozier: "Exploiting Interfacial Water Properties for Desalination and Purification Applications", S&T Strategic Objectives / Energy and Infrastructure Assurance 2005-2007.  Identify those properties of interfacial water whose better scientific understanding could lead to new desalination/purification technology opportunities; define the new experimental, theoretical, and simulation approaches to predictive understanding of these key properties; and provide a plan that links research successes in these areas to follow-on technology opportunities.

(25) Amalie Frischknecht: "Carbon Nanotube Sorting via DNA-Directed Self-Assembly", NTM FY05-FY07.  We did some atomistic simulations of single DNA bases adsorbing on a carbon nanotube in water, using LAMMPS and the CHARMM force field.  Various features added to the code right before I started this project were really helpful (e.g. being able to apply constant pressure with the nanotube fixed in the box).

(26) Paul Crozier: "DNA-Based Intelligent Microsensors for Genetically Modified Organisms (GMO)", Biotechnology / CIS 2005-2007.  Develop differentiating capabilities for designing DNA-based biosensor systems that detect mutations in target oligonucleotides, a necessity for GMO detection.

(27) Paul Crozier: "Molecular Dynamics of Membrane Proteins", S&T Strategic Objectives 2004.  Understanding the dynamics of the membrane protein rhodopsin in order to explore implications for other membrane proteins and cellular signaling processes. Study of rhodopsin (rho) as a model light-activated G-protein coupled receptor (GPCR).

(28) Gary Grest: "Elucidating the Mysteries of Wetting", FY04-06 (winner 2006 Sandia LDRD Award for Excellence).

(29) John Curro: "Solution Behavior of PEO", Senior/Junior FY04.  I started the implementation of TIP4P in LAMMPS during this project.  I also did some MD simulations of melt PEO, which we compared to PRISM and scattering (Curro and Frischknecht. The structure of poly(ethylene oxide) liquids: comparison of integral equation theory with molecular dynamics simulations and neutron scattering.

(30) Jon Zimmerman: "Atomistic Modeling of Nanowires, Small-scale Fatigue Damage in Cast Magnesium, and Materials for MEMS", Strategic Partnerships (PECASE) FY02-FY06.  Use experiments and simulation to understand materials characteristics in relation to mechanical properties at the nano and microscales in MEMS.

(31) Jon Zimmerman: "A Robust, Coupled Approach for Atomistic-Continuum Simulation", Engr Sci FY02-FY04.  Develop tools and metrics for coupled atomistic-simulation under quasistatic and finite temperature loading conditions.

(32) Paul Crozier: "Interfacial Bioscience Grand Challenge (IBIG)", Grand Challenge LDRD 2001-2003.  Develop unique experimental and computational capabilities for the understanding of membrane protein structure, membrane protein-protein and protein-lipid interactions, and dynamics of membrane protein structural changes related to toxin action, pathogen-cell interactions, immune system activation, and cell signaling.

(32) Jon Zimmerman: "Crack Nucleation and Growth: Combined Validated Atomistic and Continuum Modeling", Engr Sci FY99-FY01.  Use combined experimental and computational approaches to examine defect nucleation from stress concentrations leading to material failure.

(34) Mike Chandross: "Separation of Xylenes Using Zeolites", FY99-FY01.  We used the GCMC feature (from Ladera) to simulate diffusion of butane and isobutane (mimicking xylenes) in a zeolite.