SciDAC-EFRC Collaboration: WastePD

David Riegner, Nikolas Antolin, Christian Oberdorfer, Wolfgang Windl

Dept. of Materials Science and Engineering
The Ohio State University, Columbus, OH
Two general forms of nuclear waste in the US:

- Spent Nuclear Fuel (SNF) from commercial power plants
  - Responsibility of DOE Office of Nuclear Energy (NE)
- Legacy waste from fabrication of nuclear weapons
  - Responsibility of DOE Office of Environmental Management (EM)

Current US law:

All nuclear waste will be disposed of in Yucca Mountain, NV.

Current US situation:

YM defunded, everything is in limbo.
Most exists in underground tanks in Hanford, WA and Savannah River, SC.

Current form is liquid, sludge, and precipitated solids.

Most will be mixed with borosilicate glass and cast into SS containers:

- has been ongoing at SR for 20 y
- has not started at Hanford because of design issues with vit plant

Some radionuclides cannot be processed through glass melter and must be stabilized in ceramic or metal hosts.

Final waste forms should be stable for very long periods of time, >$10^5$ y.

Containers are metallic: steel underground tanks, SS dry storage casks, CRA canisters for final disposal.
Center for Performance and Design of Nuclear Waste Forms and Containers

Goals:

• Develop fundamental understanding of degradation mechanisms of waste forms and containers
• Develop new materials with improved properties
Common Method: Atomic-Scale Modeling

Tasks:

- **Structural stability and energetics**
- **Surface processes** (e.g. oxidation, desorption)
- **Kinetic processes** (e.g. deformation, diffusion)

"**Alloy**": Random mixture of different atom types

"**Phase separation**": Energy is lower when atoms separate

Alloy is stable if its energy is lower than the sum of the energies of the separated phases
Tasks:

- Structural stability and energetics
- Surface processes (e.g. oxidation, desorption)
- Kinetic processes (e.g. deformation, diffusion)
Tasks:

- Structural stability and energetics
- Surface processes (e.g. oxidation, desorption)
- Externally-controlled processes (deformation, field evaporation)
Molecular Dynamics

• One way to simulate that: “Molecular Dynamics”

  • “Potential”: Come up with a function (or method) that describes energy of system as function of atomic positions

  • Simulation: Evolve position of atoms (e.g. Newton’s law)

  • Analysis: Visualize / evaluate what happened

\[ \mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t) \Delta t + \frac{1}{2} \mathbf{a}(t) \Delta t^2 \]

\[ \mathbf{a}(t + \Delta t) = -\frac{1}{m} \nabla V[\mathbf{r}(t + \Delta t)] \]

\[ \mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \frac{1}{2} [\mathbf{a}(t) + \mathbf{a}(t + \Delta t)] \Delta t \]

Potentials are functions fitted to e.g. experimental data
Our Wish List to SciDAC (8/22/16)

• **High computational demands in many calculations**
  • Efficient compilation and use of DFT and MD codes
  • GPU support for VASP, LAMMPS (running at OSC, speedup 4-5 times) and other codes? E.g., VASP website has list of tasks that are not yet parallelized for efficient GPU execution.
  • Other architectures, or alternative execution/platform (e.g. KOKKOS, CUDA)?

• **Empirical potential fitting**
  • RAMPAGE package (atomistics.osu.edu)
  • Improve fitting reliability and speed, assess quality

• **Visualization and analysis of complex atomistic simulation results**
  • Amorphous and complex systems
  • Dynamic processes (infrequent events)
  • Data analysis – find specific atomic arrangements or infrequent events

• **Uncertainty quantification**
  • Reproducibility of results
    • *E.g. between DFT codes, different levels of atomistic simulations*
  • Empirical potential development
  • Fast, efficient and reliable MC (e.g. simulated annealing) and KMC
  • Error propagation in multiscale modeling
    • *E.g. DFT numbers as parameters in rate-theory modeling or KMC*
  • Reduced-model generation
    • *Which reactions *really* need to be included in a rate model to solve a given problem?*
### SciDAC Collaborations

<table>
<thead>
<tr>
<th>Generation of new empirical potentials to expand MD capabilities</th>
<th>Efficient analysis and visualization of multi-component MD data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>QUEST:</strong> Uncertainty quantification of EAM potentials generated with *RAMPAGE*</td>
<td><strong>SDAV:</strong> Real-time analysis of metallic glass structures using ADIOS integration</td>
</tr>
<tr>
<td><strong>SUPER:</strong> Parallelization and performance enhancement of *RAMPAGE*</td>
<td><strong>SDAV:</strong> Indexing and Visualization of data using FastBit and VisIt</td>
</tr>
</tbody>
</table>

- **WastePD** was interested in expanding Molecular Dynamics (MD) capabilities
  1. Lack of usable multi-component potentials
  2. Analysis of huge data sets where only a small subset of the data is relevant

- **SciDAC pilot projects directly address these issues**
  - **QUEST** and **SUPER** focused on *generation and evaluation* of new potentials
  - **SDAV** focused on *acquisition and analysis* of relevant data from large systems

- **SciDAC had the capabilities to directly enhance WastePD’s ability to simulate and analyze data for multi-component alloys**
SciDAC Collaborations – The Experience

- Fantastic opportunity – we have significantly improved tools and exciting new capabilities; would do it again!
  - Three postdocs know their way around basic MPI now
  - Understand first basics of improving code efficiency (execution and memory use)
  - Basic understanding of computational UQ
  - New visualization and analysis capabilities
- Due to short duration, focused on “low-hanging fruit”, which should be a valuable initial approach for majority of materials modeling groups
- Appreciate dedication, patience, “customer-focus”, responsiveness, and collaborative spirit of SciDAC researchers
- Without dedicated man power on our end, the experience was exhausting
  - Multiple weekly videoconferences and continuous e-mail discussions, joint coding and data generation work
  - Was a good idea to *work with several SciDACs on one topic (potentials)*
  - 3 postdocs involved on WastePD end, one taking ~70% of effort
Example Collaboration Topic: Multi-Component Potentials

- Mixing different metals together ("alloying") makes stronger, more corrosion resistant materials
- Hot candidate for corrosion resistant: "High-entropy alloys" (~5 different elements)
- To model them, we need interaction potential functions between all species
- Many researchers spend lots of time to get potentials for one element right, but very few inter-species potentials

Our approach: Reuse A-A and B-B interactions from literature, fit A-B interactions
Multi-Component Potentials

\[ U_{EAM} = \frac{1}{2} \sum_{i \neq j} \Phi_{\alpha \beta}(r_{ij}) + F_\alpha \left( \sum_{i \neq j} \rho_\beta(r_{ij}) \right) \{ r_{ij} < r_{\text{cut}} \} \]

- Binary EAM potential requires 9 total functions (right)
- Six are “known” from the elemental potentials
- The remaining 3 functions (\( \Phi_{AB}, \rho_{AB}, \rho_{BA} \)) must be fit
- Differing methods of fitting trade expense for accuracy

<table>
<thead>
<tr>
<th>Term</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F_A )</td>
<td>Pure A Potential</td>
</tr>
<tr>
<td>( F_B )</td>
<td>Pure B Potential</td>
</tr>
<tr>
<td>( \rho_{AA} )</td>
<td>Pure A Potential</td>
</tr>
<tr>
<td>( \rho_{AB} )</td>
<td>Fitting</td>
</tr>
<tr>
<td>( \rho_{BB} )</td>
<td>Pure B Potential</td>
</tr>
<tr>
<td>( \rho_{BA} )</td>
<td>Fitting</td>
</tr>
<tr>
<td>( \Phi_{AA} )</td>
<td>Pure A Potential</td>
</tr>
<tr>
<td>( \Phi_{AB} )</td>
<td>Fitting</td>
</tr>
<tr>
<td>( \Phi_{BB} )</td>
<td>Pure B Potential</td>
</tr>
</tbody>
</table>

“Holistic” Fitting: Consider many compounds and fit to as many as possible (includes elements)

Johnson Alloy Model:
Weighted average of elemental pair terms by their electron densities

\[ \phi_{\alpha \beta}(r) = \frac{1}{2} \left( \frac{\rho_\alpha(r)}{\rho_\beta(r)} \right) \phi_{\alpha \alpha}(r) + \frac{\rho_\beta(r)}{\rho_\alpha(r)} \phi_{\beta \beta}(r) \]
Potential Development with RAMPAGE

Existence (or creation) of suitable multi-component potentials is the largest barrier to widespread deployment of MD simulations for multi-component problems

RAMPAGE: Rapid Alloy Method for the Production of Accurate, General EAM Potentials

Goals:

1. Utilize available elemental potentials from literature
   • Avoid re-fitting from scratch
2. Calculate strategic “high-value” structures and compounds with DFT using VASP
3. Generate compatible binary potentials
4. Assemble binaries into higher-order multi-component potentials. No additional computation

Overview:

Use LAMMPS as a calculator to optimize new potentials with specific target data. Thus, any property accessible to MD becomes a possible parameter for new multi-component potential optimization.
WastePD was interested in expanding Molecular Dynamics (MD) capabilities  
1. Lack of usable multi-component potentials  
2. Analysis of huge data sets where only a small subset of the data is relevant  

SciDAC pilot projects directly address these issues  
- QUEST and SUPER focused on generation and evaluation of new potentials  
- SDAV focused on acquisition and analysis of relevant data from large systems  

SciDAC had the capabilities to directly enhance WastePD’s ability to simulate and analyze data for multi-component alloys
• **Topic areas:**
  – forward uncertainty propagation
  – reduced stochastic representations
  – inverse problems
  – experimental design & model validation
  – fault tolerance

• **Application:**
  – Evaluation of specific fitting parameters on simulated properties

• **Institute Director:** Habib N. Najm (hnnajm@sandia.gov; SNL)
**Objectives**
- Estimate uncertain parameters in empirical potential fitting
- Identify errors relative to DFT fitting data
- Compare and select among potentials
- Explore potential frameworks for multilevel multifidelity UQ in Kinetic Monte Carlo

**Impact**
- Facilitates new best-practices for potential development
- Identifies optimal potentials and important parameters
- Enables design of multicomponent alloys
  - Leap towards multi-component alloy design with more parameters

**Accomplishments**
- Coupled Dakota, UQTk, and LAMMPS, with Cu-Ni and Cu-Zr material models from WastePD
- Identified **specific parametric correlations** not previously explored
- Identified new **dependences between parameters and simulated properties**
- Parallel Transitional MCMC employed for speed and robust inference
- Inferred uncertain parameters; Bayesian setting
More detail on Accomplishments

- Constructed framework for direct parameter control
  - Two systems: Cu-Ni and Cu-Zr
- Parameter control illuminates model construction failures, refines optimization techniques
- Key parameters: \((r, D, \alpha)\) each have physical meaning and controls a particular property
  - \(r\), bond length - lattice parameter
  - \(D\), bond energy - mixing enthalpy
  - \(\alpha\), curvature of energy well - bulk modulus
- QUEST determined that \(D\) is a primary controller of Cu-Ni bulk modulus, \(\alpha\) is actually secondary!
- Identified model instabilities: quantifies bounding range for automation of parameter selection
More detail on Accomplishments

- In addition to key parameters: \((r, D, \alpha)\), tuning parameters \((S_a, S_b)\) are used for flexibility
  - Numerical effect of \(S_a, S_b\) not previously characterized
  - WastePD had questioned the possibility of \(S_a\) and \(S_b\) introducing multiple viable solutions/minima
  - **QUEST identified bi-modal behavior in \(S_a\) and \(S_b\)**
  - **Illustrates complications in multi-parameter setups**
    - Understanding and characterizing non-physical tuning parameters is paramount to long-term potential development goals
    - Multi-component alloys have \(S_i\) parameter for each element!
    - Initial results (prev. slide) indicate \(S_a, S_b\) only minimally impact properties in Cu-Ni, is this true in other systems?
SUPER: Institute for Sustained Performance, Energy and Resilience

- **Topic areas:**
  - Performance engineering including modeling and auto-tuning
  - Code resilience
  - Parallelization and optimization
  - Energy efficiency (scalability)

- **Application:**
  - Streamline, parallelize, optimize and scale RAMPAGE code

- **Institute Director:** Lenny Oliker (loliker@lbl.gov; LBNL)
**Objectives**

- Improve performance, scalability and flexibility of RAMPAGE, automated EAM potential generator
- Improve **portability** and target DOE computing center platforms
- Improve RAMPAGE **software engineering**
- Implement both portability and distribution

**Impact**

- Enhance quality of the resulting potentials
  - Increased complexity and/or number of trials
- **Decreased software maintenance and development effort; improved reproducibility**

**Accomplishments**

- Improved RAMPAGE performance and scalability
  - Collected baseline performance results for Cu-Ni test problem on OLCF Eos system
  - Modified RAMPAGE’s parallelism approach to support simultaneously running on more hardware
  - Demonstrated reduction in Cu-Ni run time from 1hr 47min to ~15min on 8 nodes of OLCF Eos
  - Discovered performance trade-off for CG minimization parameters between LAMMPS versions.
    - Motivates ongoing work with multiple processes for each LAMMPS run
- Introduced use of source code repository
- Reduced number/complexity of required code packages
SDAV: Scalable Data Management, Analysis and Visualization

• Topic areas:
  – **Data Management**: capture data used in science codes. Efficiently move, index, and compresses data, enable query of scientific datasets
  – **Data Analysis** – application-driven techniques for performing in situ data analysis, filtering, and reduction to optimize I/O and prepare post-processing
  – **Data Visualization** – visualization techniques that support identifying and understanding features in multi-scale datasets

• Application:
  – Streamline workflow for complicated multi-component data sets. This requires improvements in (i) data generation and (ii) data analysis

• Institute Director: ??
• **SDAV Project I:**
  – On-the-fly data analysis
Enabling analysis and validation using an in situ processing framework for the Center for Performance and Design of WastePD
Pls: Scott Klasky, Wolfgang Windl

Objectives

- Integrate ADIOS with LAMMPS simulation code and Voro++
  - Minimizes I/O overheads and facilitate on-the-fly data in simulations of multi-component metallic glass structures
- Implemented ADIOS features as custom code inside LAMMPS framework

Impact

- Reduced I/O costs, allowing higher structural and time resolution to be achieved in simulation data outputs
- Better Voro++ integration in LAMMPS facilitated the creation of new data types via a custom output type that did not previously exist

Accomplishments

- Integration of ADIOS with the LAMMPS framework.
- Ability to save tens of thousands of time steps from LAMMPS runs. Significant speed-up compared with previous implementation that saved only tens of time steps in the original setup.
- Integration of ADIOS with Voro++ via LAMMPS
- Ability to compute Voronoi volumes on tens of thousands of time steps. Major speed up compared with tens or hundreds of time steps in the original setup.

Neighbor data for a metallic glass analyzed using ADIOS integrated Voro++. Local neighbor environments in metallic glasses vary wildly.
Detecting Locally Correlated Events in Metallic Glass Simulation
SciDAC Institute-WastePD EFRC
John Wu (LBNL) and Wolfgang Windl (OSU)

Objectives
Develop and deploy algorithms for detecting changes to the local atomic environment in multi-component metallic glass materials

Impact
Successfully detected and visualized locally correlated events in the evolution of metallic glass structures, crucial to the understanding of metallic glass synthesis and design

Accomplishments
- Connected the simulation code and analysis code with ADIOS, to provide in situ data collection and facilitate efficient analysis of large systems over long simulation times
- The figure shows the number of atoms with the same neighbors evolving as a function of simulation time. Such analysis was not previously feasible.
- As time progress (from front to back), the system transitions from liquid to glass. Liquid atoms retain fewer neighbors, and the system shows a notable shift toward neighbor stability. Further analysis of neighbor environments may enable deliberate glass design.
WastePD-ADIOS Integration

- Simulations of characterization techniques are currently in development that could benefit greatly from real-time analysis
  - APT simulations involve large systems and time-evolving structures.
  - System studied in the pilot project was simplified than WastePD’s long-term goals.
- Major challenges in computational alloy design are time- and length-scales which drive up space requirements
  - Modern modeling techniques demand coupling of multiple codes, but for large data sets I/O makes coupling very (often prohibitively) slow.
  - I/O cost scales with problem size and larger problems are more scientifically relevant.
  - Ability to analyze simulation data near real-time without having to wait until full simulation is done allows the potential to observe patterns as they evolve and reduce data storage requirements.
- WastePD is currently developing new modeling methods and addressing I/O requirements and bottlenecks would benefit tremendously from continued collaborative support.
• **SDAV Project II:**

  – Data visualization & analysis (VisIt, visit.llnl.gov)
Objective
Deploy interactive query-based techniques that incorporate accelerated bitmap indexing to allow WastePD scientists to explore and understand infrequent events such as clustering or migration in large metallic glass simulation cells.

Impact
One of the challenges being undertaken by the WastePD team is the accurate predictive design of metallic glass performance over time, requiring understanding of metallic glass structure and stability over long timescales.

Accomplishments
- Deployed an interface that allows WastePD data to utilize accelerated bitmap indexing (FastBit) with the VisIt toolkit.
- Deployed tools to analyze atom neighborhood data.

Far Left – image from VisIt’s Cumulative Query tool showing histograms of atom neighborhoods and differences that is used for a range based queries (find all atoms who’s neighborhood size did not change).

Near Left – Results of the query showing the number of atoms that remained in the neighborhood (the neighborhood size may remain the same but be composed of different atoms).
Detecting and Understanding Hopping Atoms in Metallic Glass Simulations

SciDAC SDAV Institute-Waste PD EFRC
Allen Sanderson Univ. of Utah

Objective

Deploy interactive query-based techniques that incorporate accelerated bitmap indexing to allow WastePD scientists to explore and understand infrequent events such as clustering or migration in large computational cells in metallic glass simulations.

Accomplishments

- Deployed new parallelized algorithms in VisIt for handling large-scale indexed data.
- Performing range based queries.
- Constructing particle paths.

Far Left – A sub-selection selecting atoms who’s neighborhood size remained the same but all of the neighboring atoms changed between time steps.

Near Left – the paths of four atom selected above (the large jumps are due to periodic boundaries).