BLUE WATERS
SUSTAINED PETASCALE COMPUTING

Allocation: Innovation & Exploration /500 Knh
PI: Greg Bauer
National Center for Supercomputing Applications
Biology, Chemistry & Health

IMPROVING NWCHEM SCALABILITY USING THE DATASPACES FRAMEWORK

Research Challenge

- Limited pool of experimental data and the extremely labor-intensive nature of parameter optimization represent the major limiting factors in improving the quality of classical force field parameters, in which these parameters are essential to improve the predictive ability of Molecular Dynamics (MD) simulations.
- The challenge with lack of experimental data in parameter optimizations can be addressed through reliable electronic structure computations at the CCSD(T) level, however a faster and more computationally efficient implementation is needed to optimize the high computational cost of CCSD(T) method.
- Labor-intensive nature of classical force field determination, and the manually driven optimization drives the need to introduce a quality control into the optimization process, in order to be able to reproduce the results, revise and improve the protocol.

Methods & Codes

- Performance optimization of the CCSD(T) method is being developed through offloading large memory arrays from compute nodes to a dedicated data storage employing the DataSpaces data management framework.
- A methodology based on data extracted from a previously optimized version of CCSD(T), used together with scalable tool to optimize Lennard-Jones parameters in the AMBER force field has been developed to generate significant sets of quality parameters for use in subsequent MD simulations.

Results & Impact

- A new approach to parameter optimization for classical force fields that combines a high-level electronic structure calculation method to extract additional previously inaccessible information from experimental data.
- Reengineered optimization procedure that maximizes HPC resources.
- 10x shorter time needed for parameter optimization.
- Transformation of force field optimization from an empirical to a well-structured discipline enables reproducibility of results by the community instead of limiting the accessibility to only a few highly capable teams.

Why Blue Waters

- The fast interconnect and large memory per core of Blue Waters enables CCSD(T) computations of molecular systems encountering a thousand basis functions, which is vital for the success of the developed parameter optimization procedure.
- Availability of large numbers of nodes is essential to address the extremely resource demanding needs of exhaustive exploration of parameter space.