Collaborative Research: Elements: SciMem: Enabling High Performance Multi-Scale Simulation on Big Memory Platforms

PI: Dong Li[1], PI: Zhen Li[2], Co-PIs: Yanbao Ma[1] and Liang Shi[1] [CSSI #2104116, #2103967]

[1]*Computer Science and Engineering & Mechanical & Chemistry, University of California-Merced, Merced, CA 95343*

[2]*Department of Mechanical Engineering, Clemson University, Clemson, SC 29634*

Motivation

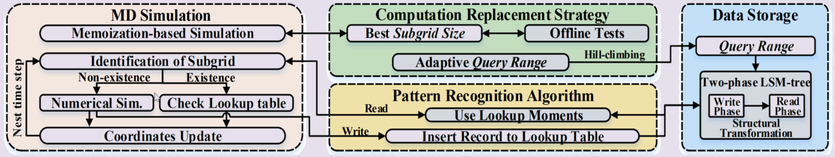
The big memory infrastructure is emerging and has shown great potential to increase scientific computation scales and solve larger numerical problems. However, using big memory architectures for multi-scale simulations is challenging because of limited computing capability in the big memory machines and memory heterogeneity introduced by big memory.

The goal of this project is to create a capability and a software package (named SciMem) that will enable high performance multi-scale simulation on big memory platforms. By designing SciMem as a library, we expect to bring a 10x performance improvement for certain larger-scale multi-scale simulations widely applied in the fields of computational chemistry, fluid physics and material science, i.e., quantum mechanical/molecular mechanical-based molecular dynamics (MD) simulations.

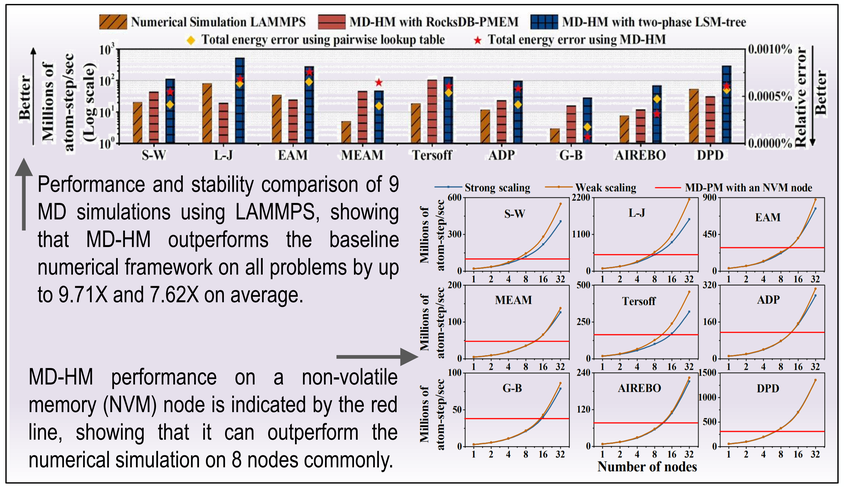
Y1 KEY PROGRESS #1

Computational Challenge in using big-memory platforms to improve the performance of molecular dynamics (MD) simulations by trading memory capacity for computational capability through a lookup table-based memorization technique.

- We developed a memoization-based MD framework customized for big memory systems by using a new two-phase LSM-tree to optimize read/write performance, leading to an average speedup of 7.6X based on the Intel Optane-based big memory system.



MD simulations drive the execution flow and are accelerated by computation replacement strategy, shape matching-based pattern recognition, and heterogeneous memory (HM)-optimized data storage.



Performance and stability comparison of 9 MD simulations using LAMMPS, showing that MD-HM outperforms the baseline numerical framework on all problems by up to 9.71X and 7.62X on average.

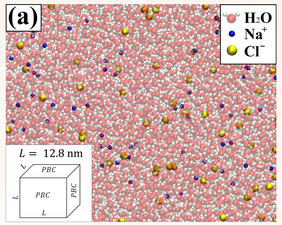
MD-HM performance on a non-volatile memory (NVM) node is indicated by the red line, showing that it can outperform the numerical simulation on 8 nodes commonly.

Y1 KEY PROGRESS #2

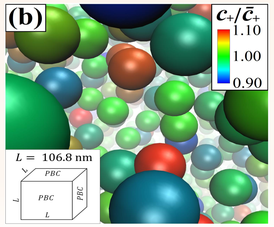
Theoretical Challenge in using big-memory platforms in multiscale simulations: Lack of theoretical guidance to determine how long we should keep the useful computational data in time-dependent multiscale simulations.

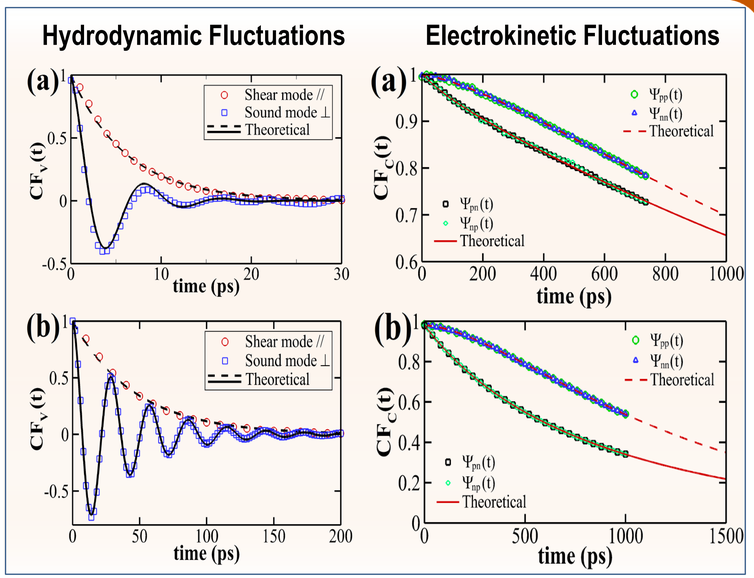
- We have established a theoretic framework based on the Landau-Lifshitz theory for quantifying decorrelation in coupled fluctuating hydrodynamics and electrokinetics, leading to analytical solutions of temporal correlations of hydrodynamic and electrokinetic fluctuations that indicate how long we should store temporal data in multiscale simulations to ensure accuracy.

Atomistic molecular dynamics (MD)



Mesoscale charged dissipative particle dynamics (cDPD)





Summary

- We have designed a runtime framework called MD-HM, which builds a new data structure and eliminates redundant patterns to enable efficient use of heterogeneous memory on big memory systems for scientific computations. Results show that MD-HM yields better performance than traditional numerical simulations performed in LAMMPS and the state-of-the-art key-value store.
- We have presented theoretical closed-form expressions derived from the linear response theory for predicting the decorrelation time for fluctuations of electrolyte bulk solution, which can be used to determine how long we should keep the data in memory for time-dependent multiscale simulations. These theoretical expressions have been validated by numerical experiments of both atomistic MD and mesoscale cDPD simulations at length scales from 10 to 100 nanometers.

References:

- Zhen Xie, Wenqian Dong, Jie Liu, Ivy Peng, Yanbao Ma and **Dong Li**, *MD-HM: memoization-based molecular dynamics simulations on big memory system,* Proceedings of the ACM International Conference on Supercomputing, 2021, Pages 215-226.
- Mingge Deng, Faisal Tushar, Luis Bravo, Anindya Ghoshal, George Karniadakis and **Zhen Li**, *Theory and simulation of electrokinetic fluctuations in electrolyte solutions at the mesoscale.* Journal of Fluid Mechanics, 2022. 942: A29.

PASA LAB @UCM

MuthComp @CLEMSON

Supported by NSF CSSI 2104116 & 2103967