GCLab

A4MD

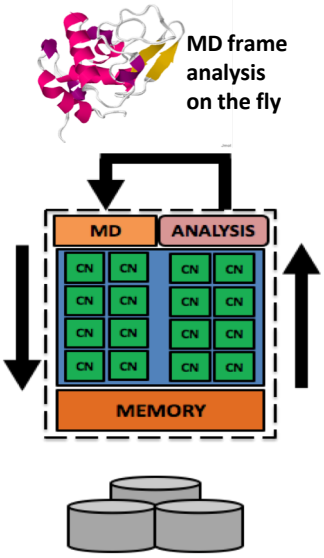
Michela Taufer[1], Trilce Estrada[2], Ewa Deelman[3], Rafael Ferreira da Silva[3], Michel Cuendet[4], Harel Weinstein[4]
[1]U. Tennessee Knoxville, [2]U. New Mexico, [3]U. Southern California, [4]Weill Cornell Med. College of Cornell U.

Project Description and Impacts

Overview:

This project targets one of the most common simulations at petascale: molecular dynamics (MD) simulations studying the (classical) time evolution of a molecular system at atomic resolution.

Next-generation computing systems will have dramatically higher performance than current systems. But I/O bandwidth and parallel file system capacity will not grow at the same rate. Burst buffers and I/O staging will not be able to address the analysis challenges in MD simulations.

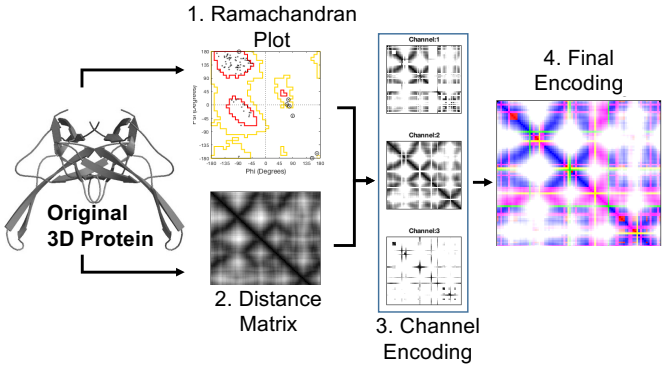


Our goal is developing new comprehensive workflows for MD simulations in which HPC meets data analytics.

Data Representation for Trajectory Analysis

Analysis Approach:

We represent macromolecules (e.g., from MD frames) in a way that exposes their structure (i.e., secondary and tertiary characteristics) and can be processed efficiently through state-of-the-art machine learning techniques such as convolutional neural networks.



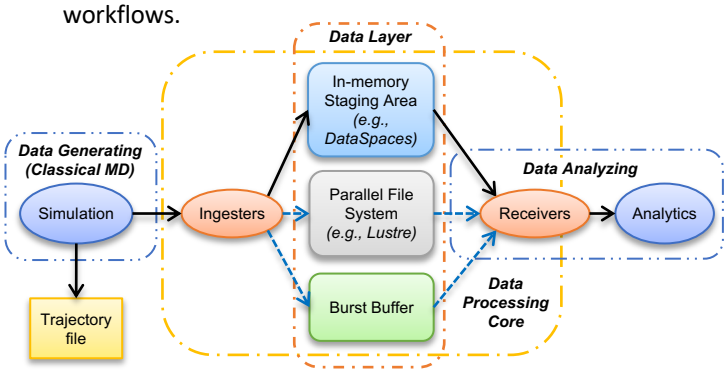
By harnessing data from ensembles of MD trajectories in situ, we capture rare events and state transformations in MD trajectories at runtime.

Simulation and Analytics Integration

In Situ Approach:

Our approach transforms the centralized nature of canonical MD analysis into a distributed one performed in situ.

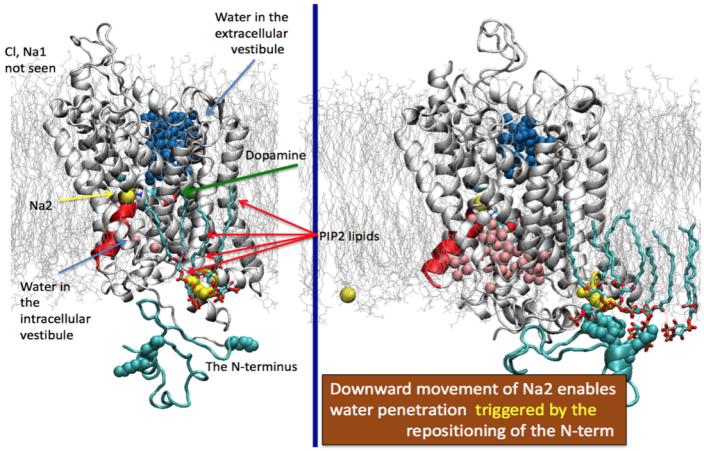
- Designed to support a broad range of MD codes, and enables on-the-fly tuning of MD workflows (i.e., stop, start, and fork MD jobs).
- Analyzes data as they are generated, save to disk only what is really needed for future analysis, and annotate MD outputs to drive the next steps in increasingly complex MD workflows.



Application to Relevant MD Systems

Dynamic information extraction:

Use our approach to extract and annotate quantitative information about the dynamics of complex molecular machines in the membrane as the MD simulations evolve. Specific example: ***The Human dopamine transporter (hDAT)***



Spontaneous Inward Opening of the Dopamine Transporter Is Triggered by PIP2-Regulated Dynamics of the N-Terminus
Khelashvili et al. (ACS Chemical Neuroscience 2015 6 (11), 1825-1837DOI: 10.1021/acschemneuro.5b00179)

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Publications:

T. Estrada, J. Benson, H. Carrillo-Cabada, A. Razavi, M. Cuendet, H. Weinstein, E. Deelman, and M. Taufer: **Graphic Encoding of Macromolecules for Efficient High-Throughput Analysis.** BCB 2018: 315-324.
X. Chen, M. Peterson, J. Benson, M. Taufer, and T. Estrada: **KeyBin2: Distributed Clustering for Scalable and In-Situ Analysis.** ICPP 2018: 34:1-34:10.