



NSCI Framework: Software: SCALE-MS – Scalable Adaptive Large Ensembles of Molecular Simulations

Award #:
OAC-1835720

PIs: Kristen Fichtorn (Penn State), Peter Kasson (Virginia),
Michael Shirts (Colorado), Matteo Turrilli (Rutgers)

Sc

Scaleable: Weakly coupled algorithms can make it possible to get near-linear scaling in time-to-solution with 10^4 to 10^6 instances of simulation ensembles, even with large (minutes) latency

A

Adaptive: A framework for simulations to communicate both synchronously and asynchronously, and with this information, make decisions altering and shaping the simulation ensemble

L

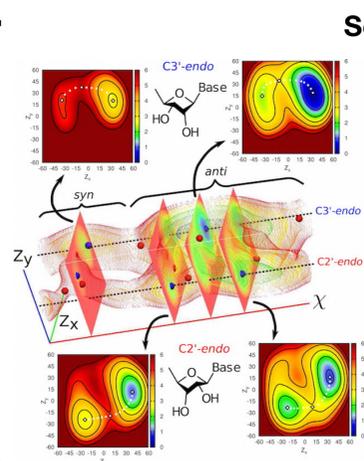
Large: With linear-scaling, weakly-coupled algorithms, it is possible to effectively use leadership computing and cloud computing for molecular problems in chemistry, biology, and materials science

E

Ensembles: Many accelerated simulation methods can be parallelized by running infrequently communicating copies that only exchange partial information

MS

Molecular Simulations: Approach designed for molecular simulation problems that have widely separated timescales, but should be adaptable to multiple-time scale algorithms in other areas



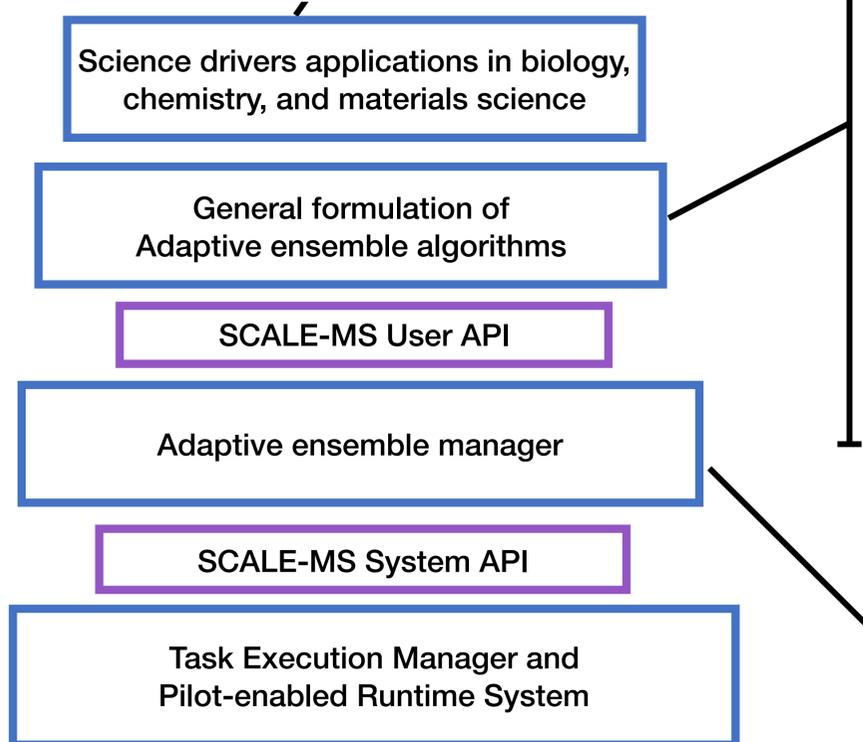
Software kernels:

- GROMACS
- LAMMPS
- OpenMM
- NAMD
- Extensible to others

Types of ensemble members:

- 'Vanilla' molecular dynamics
- Metadynamics instances
- Replica exchange sets
- Umbrella sampling
- Nonequilibrium dynamics

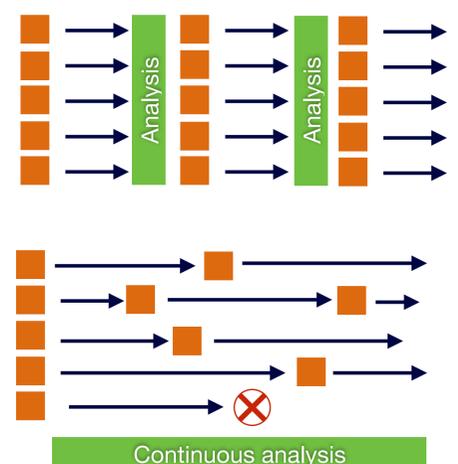
SCALE-MS Software Stack



Example ensemble algorithms

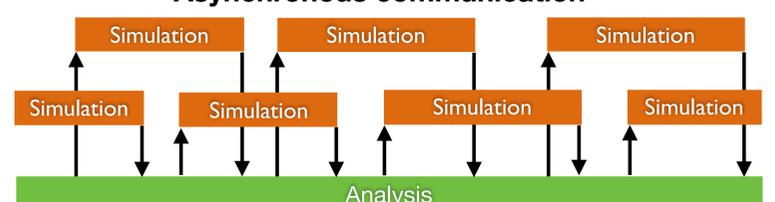
- Umbrella sampling simulations adaptively adjusting biasing potentials
- Communicating sets of replica exchange simulations
- Metadynamics simulations simultaneously exploring large numbers of biasing potentials
- Adaptive Markov State modeling
- Adaptive placement of umbrella biasing potentials
- Forward flux sampling, milestoning, and weighted ensemble sampling over adjustable reaction paths.

Typical ensemble patterns



- Support for **file-based** communication between ensemble members
 - Analysis occurs on file outputs
 - New simulations relaunched with adaptive created inputs
 - Best for **asynchronous** algorithms where data doesn't get 'stale'
 - Easy to plug in new software, adapters require no application changes
- Support for **in-flight** communication between ensemble members
 - Simulations exchange **limited** data, less need for high performance comms
 - Best for **synchronous** algorithms that need only most recent data
 - Handles more rapid communication, but requires application changes

Asynchronous communication



Synchronous communication



A system that allows application-level control of acquired resources via a scheduling overlay and a placeholder job

- Enables the fine-grained "slicing and dicing" of resources
- Provides late-binding of workloads to resources

