Towards Zettascale Computing on Exascale Platforms

Shantenu Jha
Rutgers University and Brookhaven National Lab.

http://radical.rutgers.edu

Outline

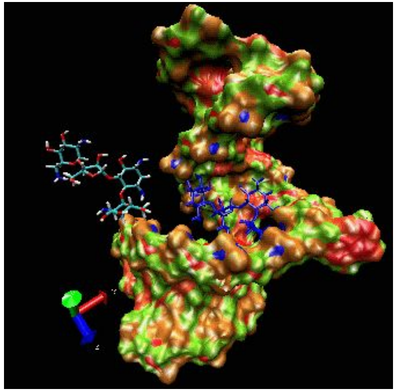
- Ensemble Computational Model Molecular Science Case Study
  - Adaptive Ensemble Execution
  - VAE driven Ensemble Execution
  - DeepDriveMD: DL driven Adaptive Ensembles on Summit
- Learning Everywhere!
  - MLaroundHPC: Classification, Examples
- MLaroundHPC: Opportunities and Challenges
  - System and Software Challenges, “Effective Performance”
  - Reference Architecture, Benchmarks

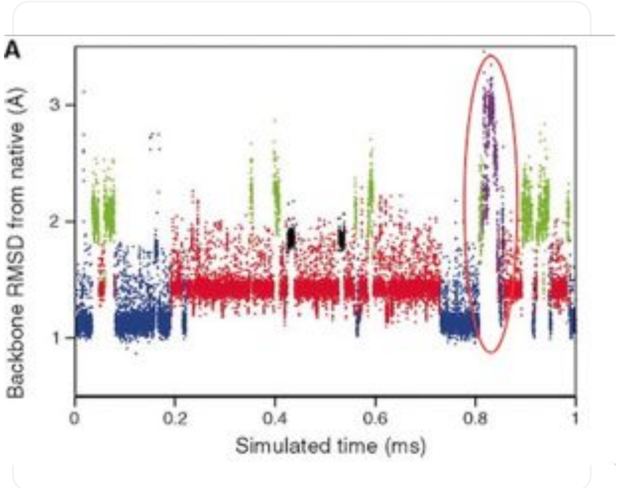
Learning EveryWhere: Motivation

- Current trends towards performance and scale **unsustainable**
  - Super complex scientific alg. & code exposed to architectural churn
- Some scientific domains have refrained and survived post-Dennard!
  - Biomolecular sciences still several orders of magnitude away
- Importance of effective methods critical at extreme scales
  - Methods effective at low scale, but not necessary at greater scales
  - Campaigns have traditionally been “static” and not used intermediate data
- Can ML enhance the **effective performance** of HPC simulations ?
  - Argue ML can enhance HPC simulations by 10^6 (?) if not greater!
  - Enhancement measured by science “achieved”
- **Learning Everywhere:** Control, Substitution and Assimiliation
  - Many system & application and architecture & software challenges

Ensemble Biomolecular Simulations

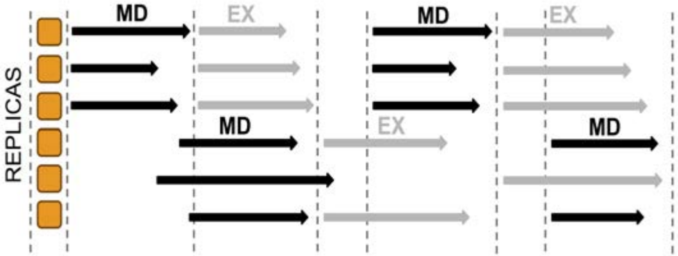
- Molecular Dynamics (MD): Newtons' Laws to integrate atoms over many timesteps
  - Immense success! (Chem, Nobel 2013)
- Single MD simulations not sufficient
  - Time scale vs quantitative accuracy
- Generate ensemble of simulations in parallel as opposed to one realization of process
  - Statistical approach: **O(10^6 - 10^8)** !
- Specialized hardware, e.g., DE Shaw "Anton" valuable, *but can ensemble-based algorithms do better than specialized hardware*? **YES**

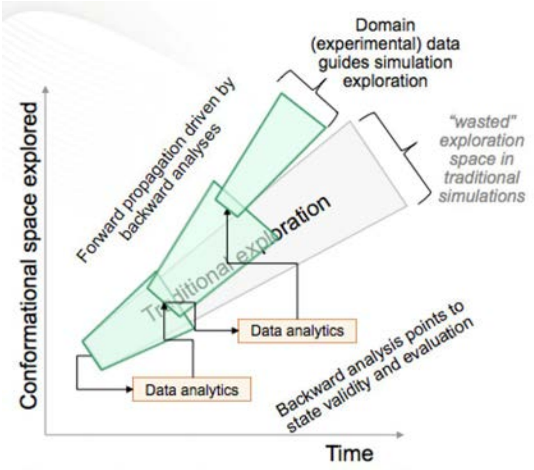




Adaptive Ensemble Algorithms

- Ensemble-based MD simulations significant improvement over single MD simulations
- Ensemble-based methods necessary, but not sufficient !
- Adaptive Ensemble-based Algorithms: **Intermediate data, determines next stages**
- Adaptivity: Better, Faster or Greater Sampling
- Adaptivity: How, What
  - Internal data: Simulation generated data used to determine “optimal” adaptation
  - External data: Experimental or separate computational process.
  - What: Task parameter(s), order, count, ….

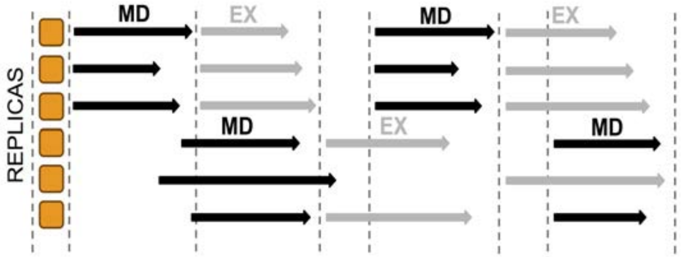


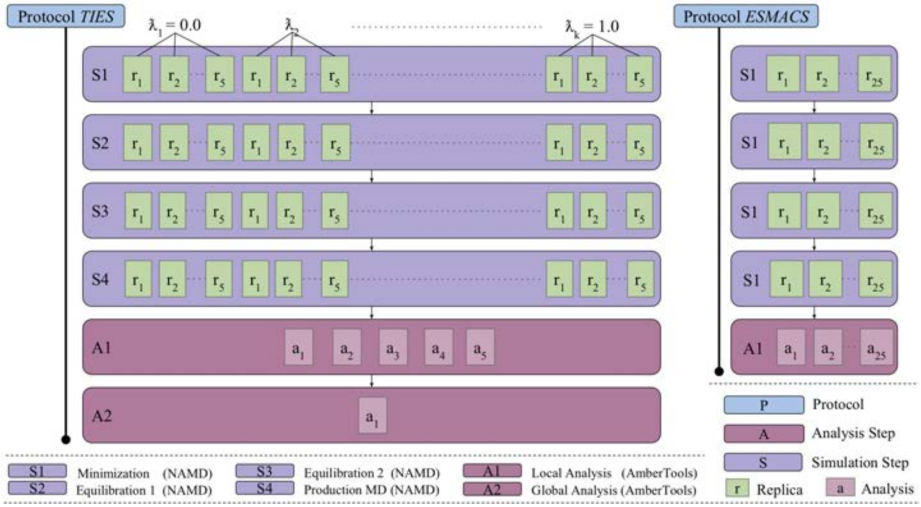


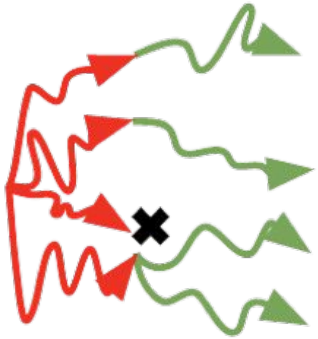
*Chodera, J.D., Noe, F., Curr. Opin. Struct. Biol. (2014)

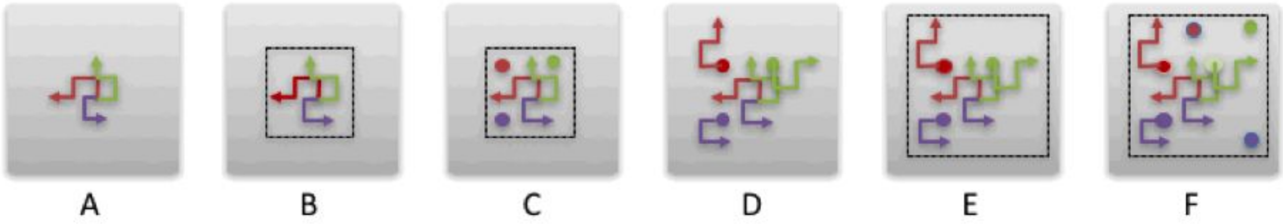
Adaptive Ensemble Algorithms: Variation on a theme

Better, Faster, Greater sampling

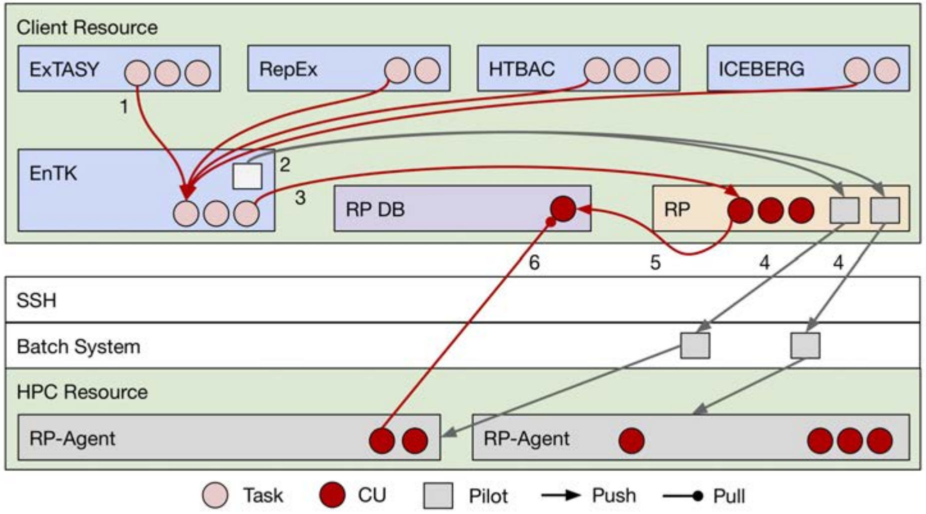


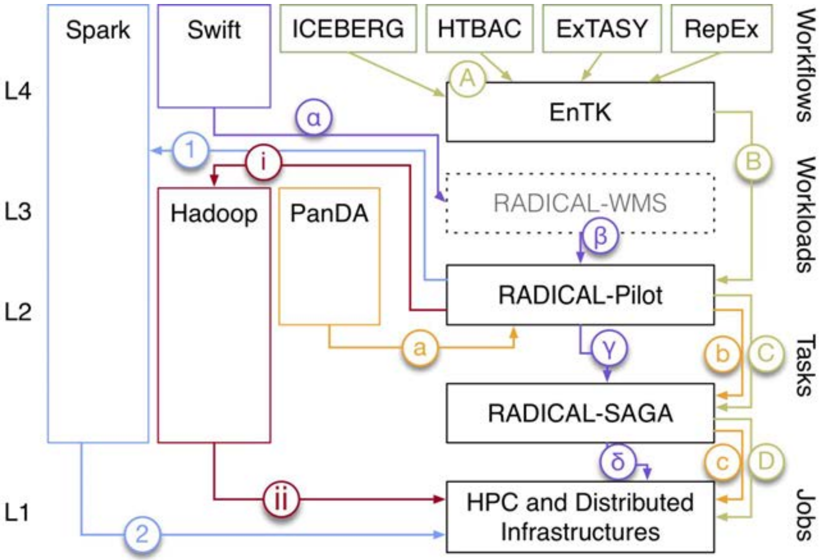






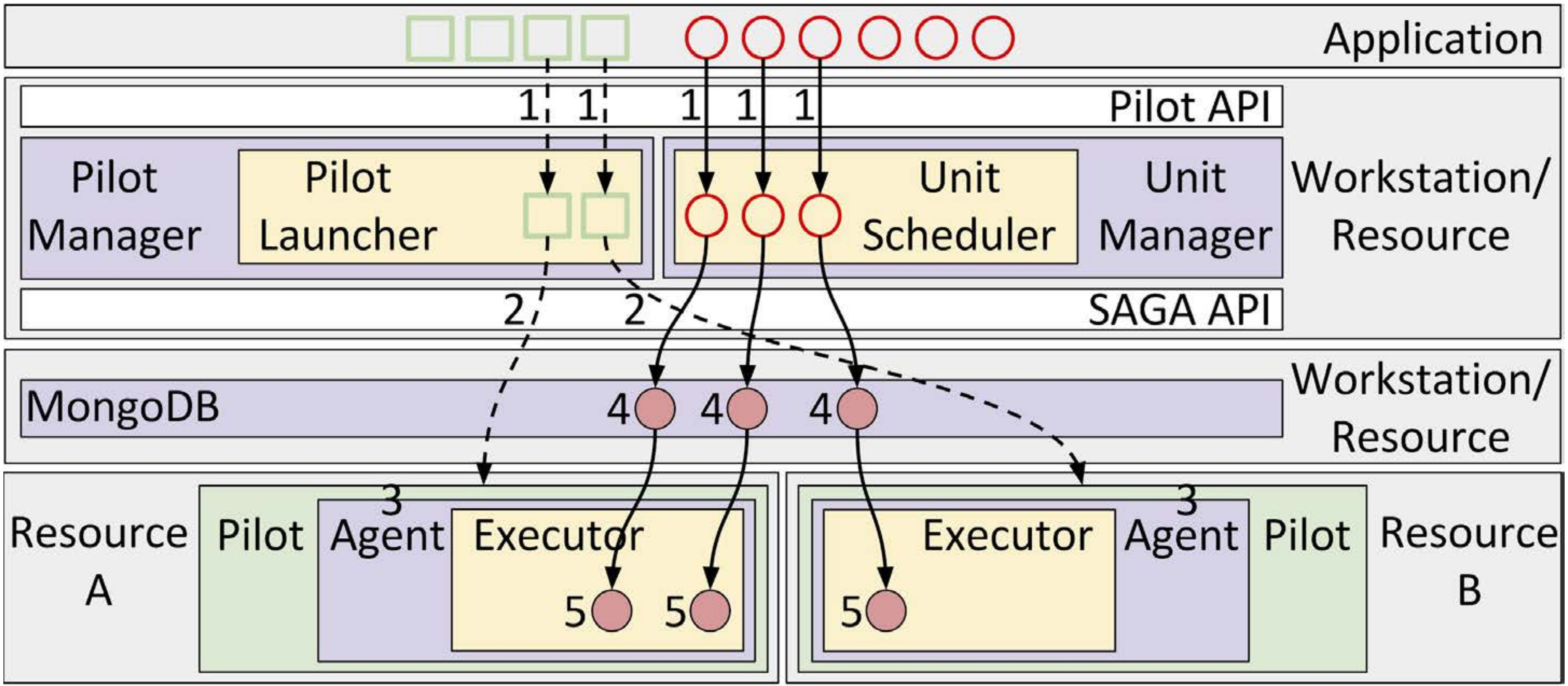
Software Systems Challenge: Specificity with Performance



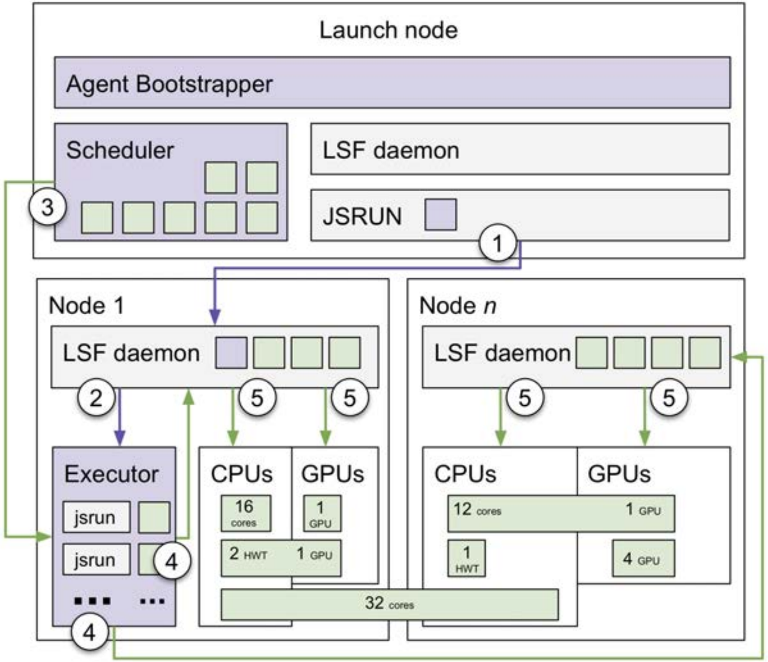


Middleware Building Blocks for Workflow Systems https://arxiv.org/abs/1903.10057

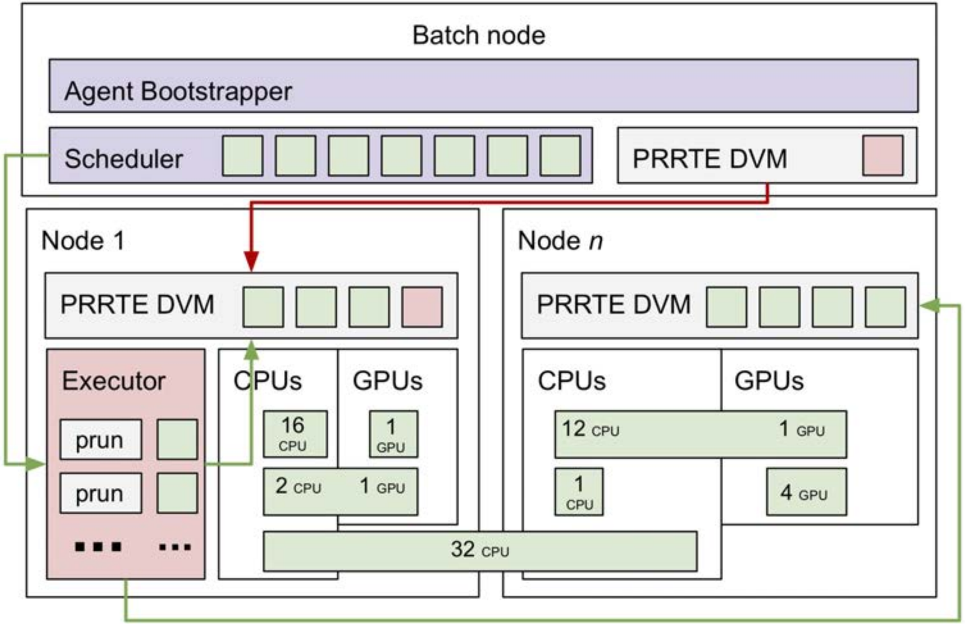
RADICAL-Pilot: Execution Model



Summit

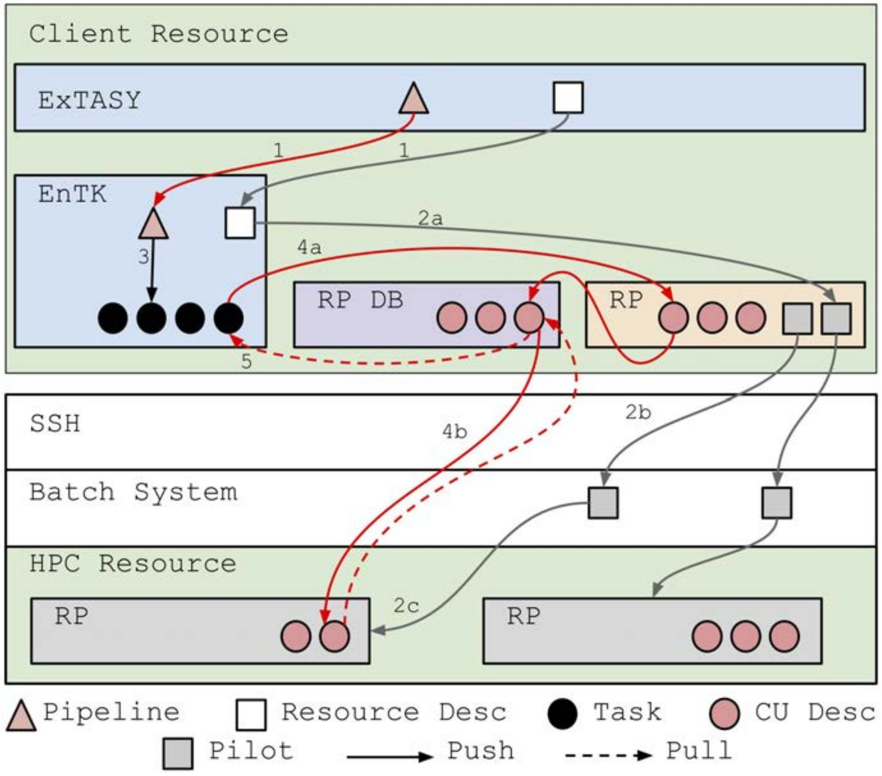


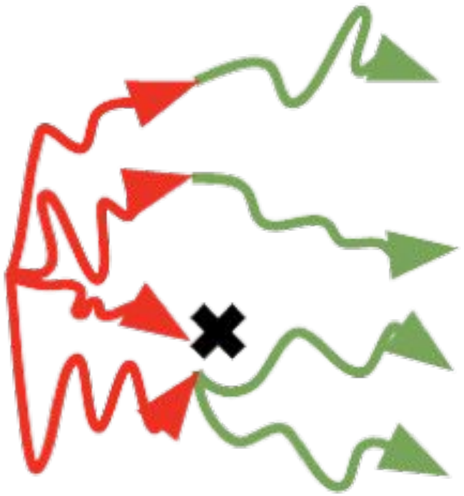
Launch node
Agent Bootstrapper
Scheduler
LSF daemon
JSRUN
1
3
Node 1
LSF daemon
2
5
5
Executor
jsrun
jsrun
4
4
CPUs
GPUs
16 cores
1 GPU
2 HWT
1 GPU
32 cores
Node n
LSF daemon
5
5
CPUs
GPUs
12 cores
1 GPU
1 HWT
4 GPU

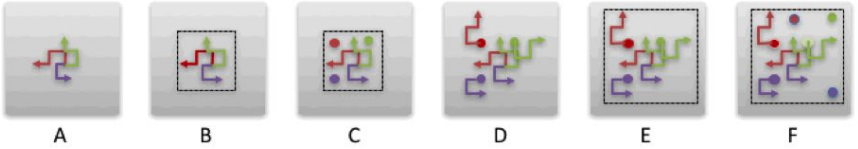


Batch node
Agent Bootstrapper
Scheduler
PRRTE DVM
Node 1
PRRTE DVM
Executor
prun
prun
CPUs
GPUs
16 CPU
1 GPU
2 CPU
1 GPU
32 CPU
Node n
PRRTE DVM
CPUs
GPUs
12 CPU
1 GPU
1 CPU
4 GPU

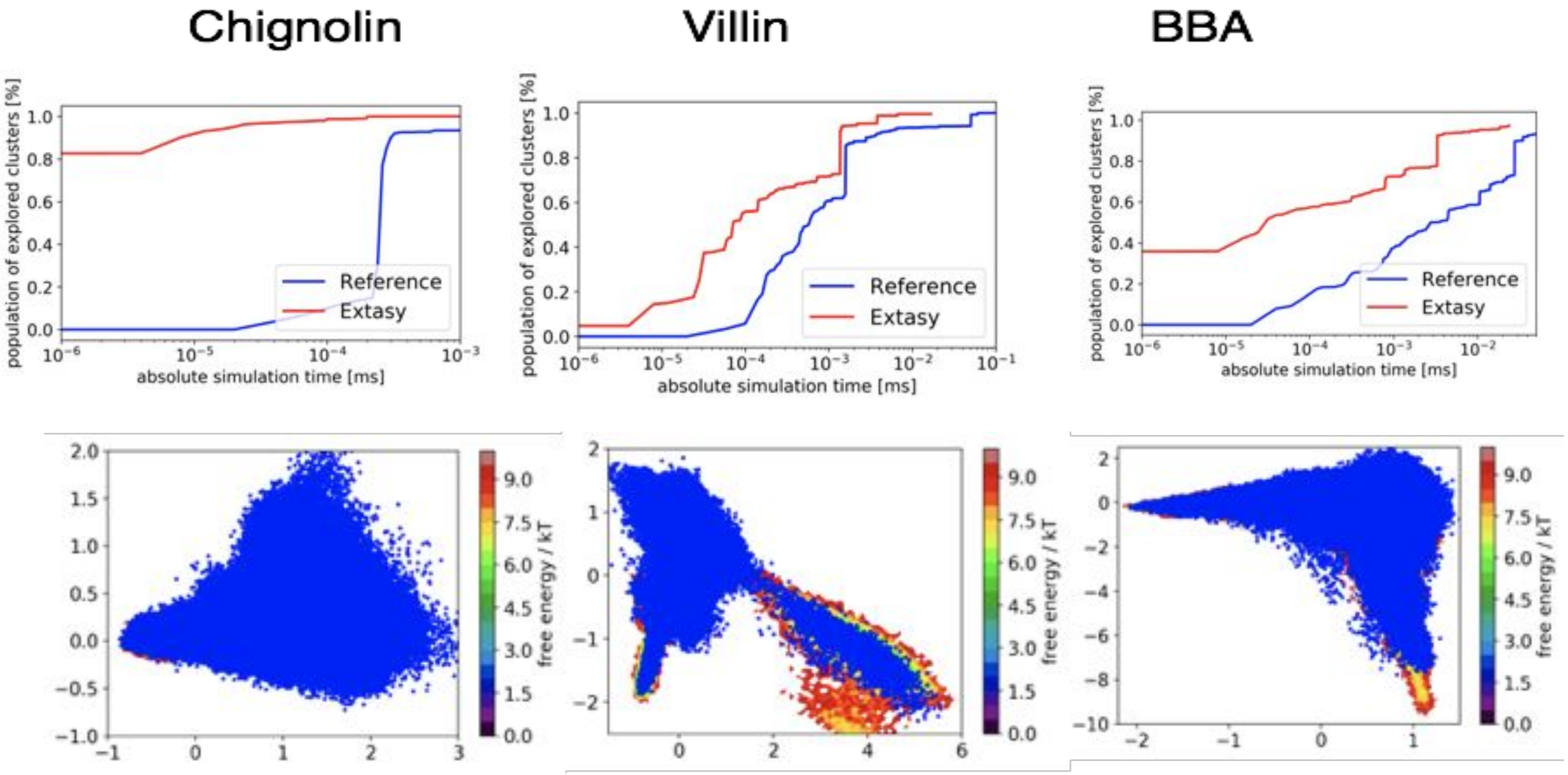
ExTASY: Domain Specific Workflow System





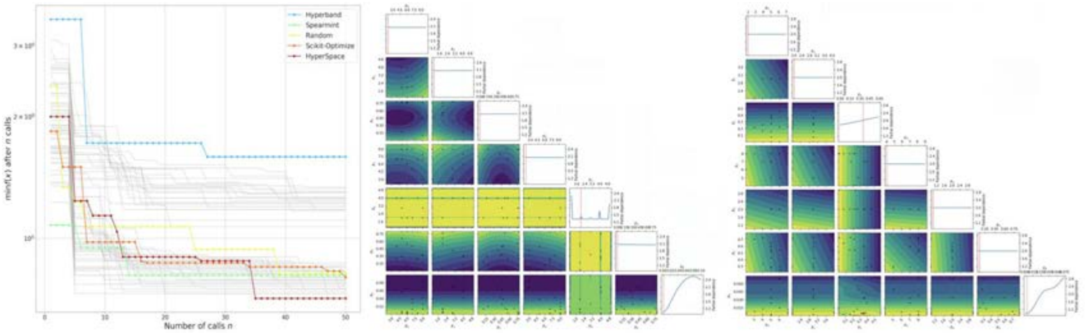


Adaptive Ensemble MD (MLaroundHPC)



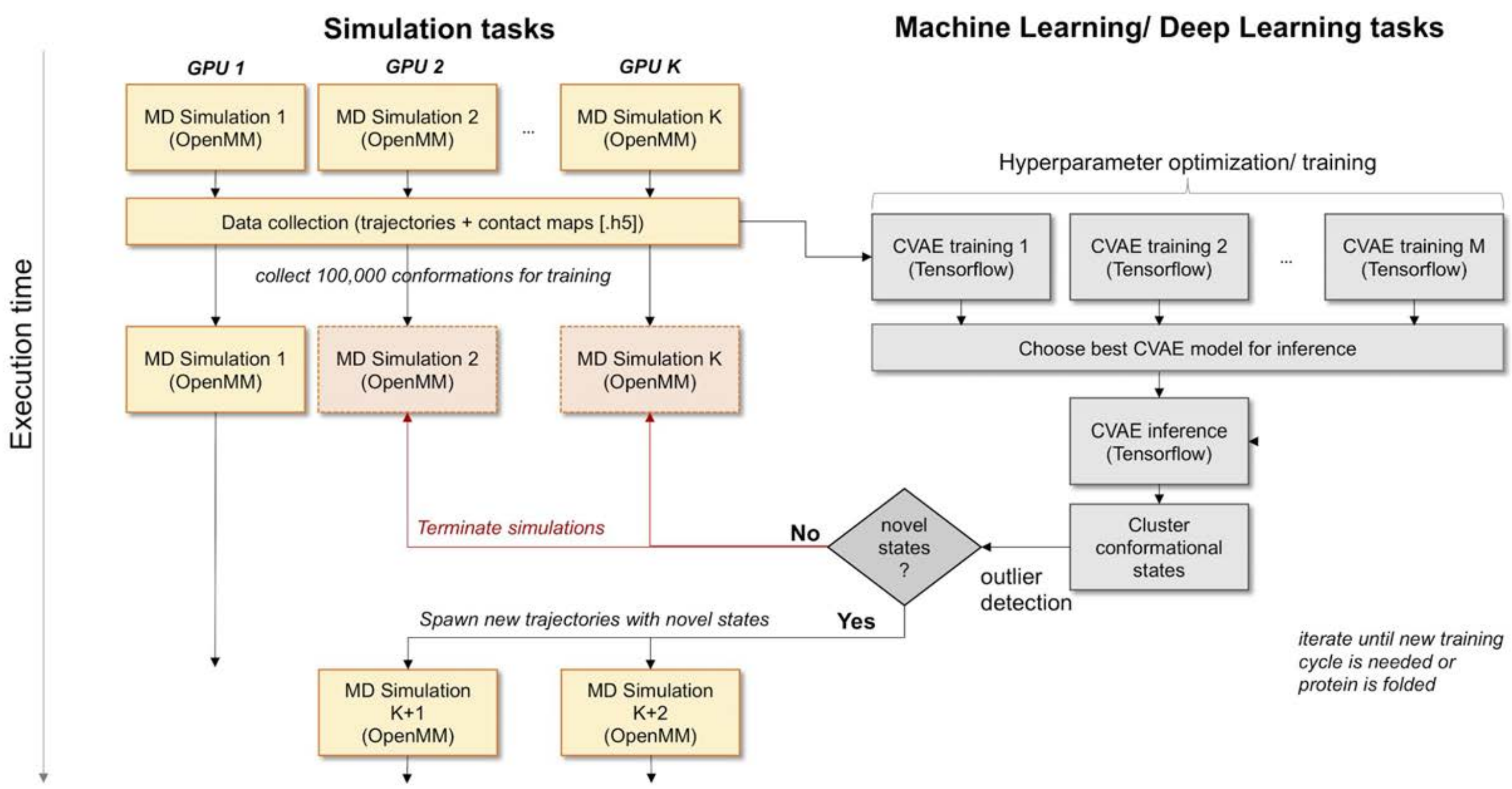
Deep Clustering of Protein Folding (MLafterHPC)

- Using DL to build low dimensional representations of states from simulation trajectories.
  - CVAE can transfer learned features to reveal novel states across simulations
  - Deep clustering of protein folding simulations using CVAE and Bayesian
  - HPC Challenge: DL approaches to achieve near real-time training & prediction!

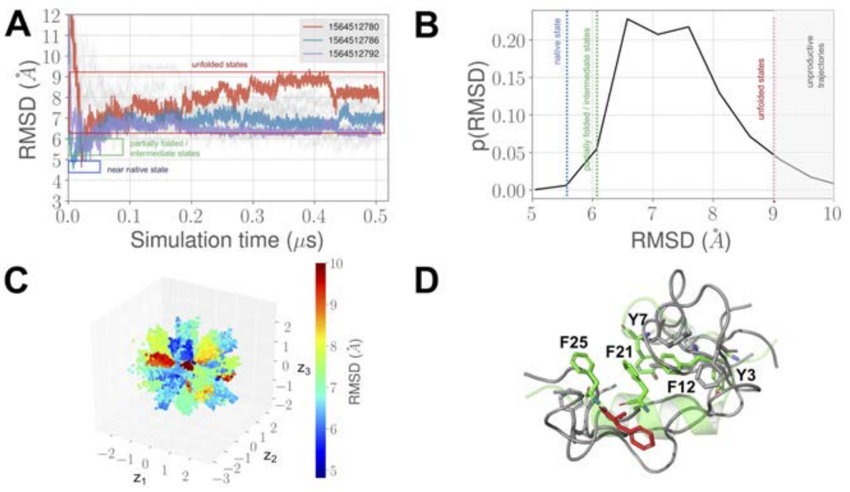


Deep clustering of protein folding simulations, Debsindhu Bhowmik et al, https://doi.org/10.1101/339879

CVAE driven Ensemble-MD



CVAE driven Ensemble-MD



System	Total no. simulations	Total simulation time (μs)	(Shortest*, Longest) simulations (μs)	Iterations	Min. RMSD (Å)
Fs-peptide	31	54.198	1.01, 3.4	90	1.6
BBA (FSD-EY)	45	18.562	0.517, 0.873	100	4.44

Table 1. Summary statistics of simulations. *Only considering the simulations that pass the initial threshold.

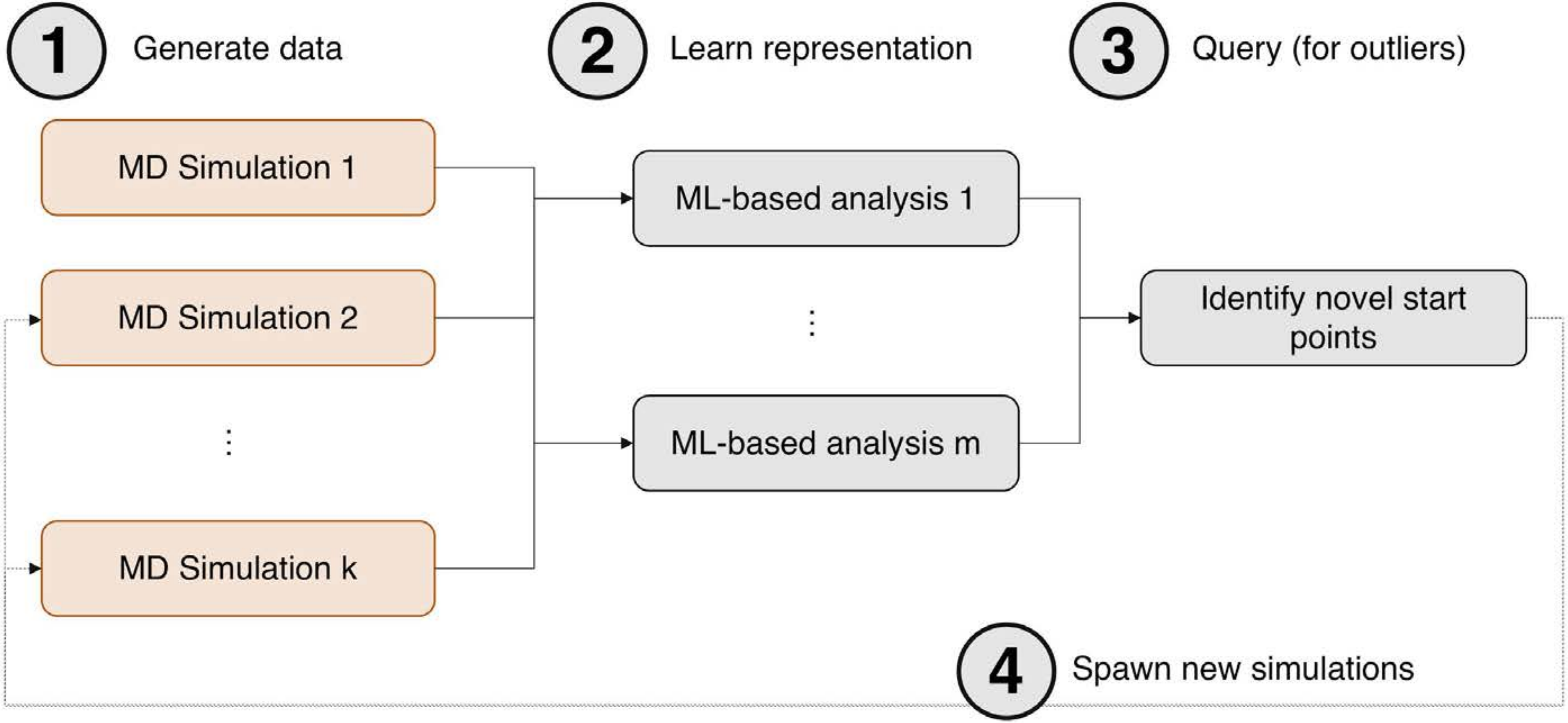
For BBA: 20X improvement over ExTASY!!

System	DL training (100 epochs; minutes)	Time per epoch (seconds)	Inference time (ms/frame)	MD simulations (ns per minute)
Fs-peptide	7	5	5.13	1.25
BBA	11	7	1.27	1.20

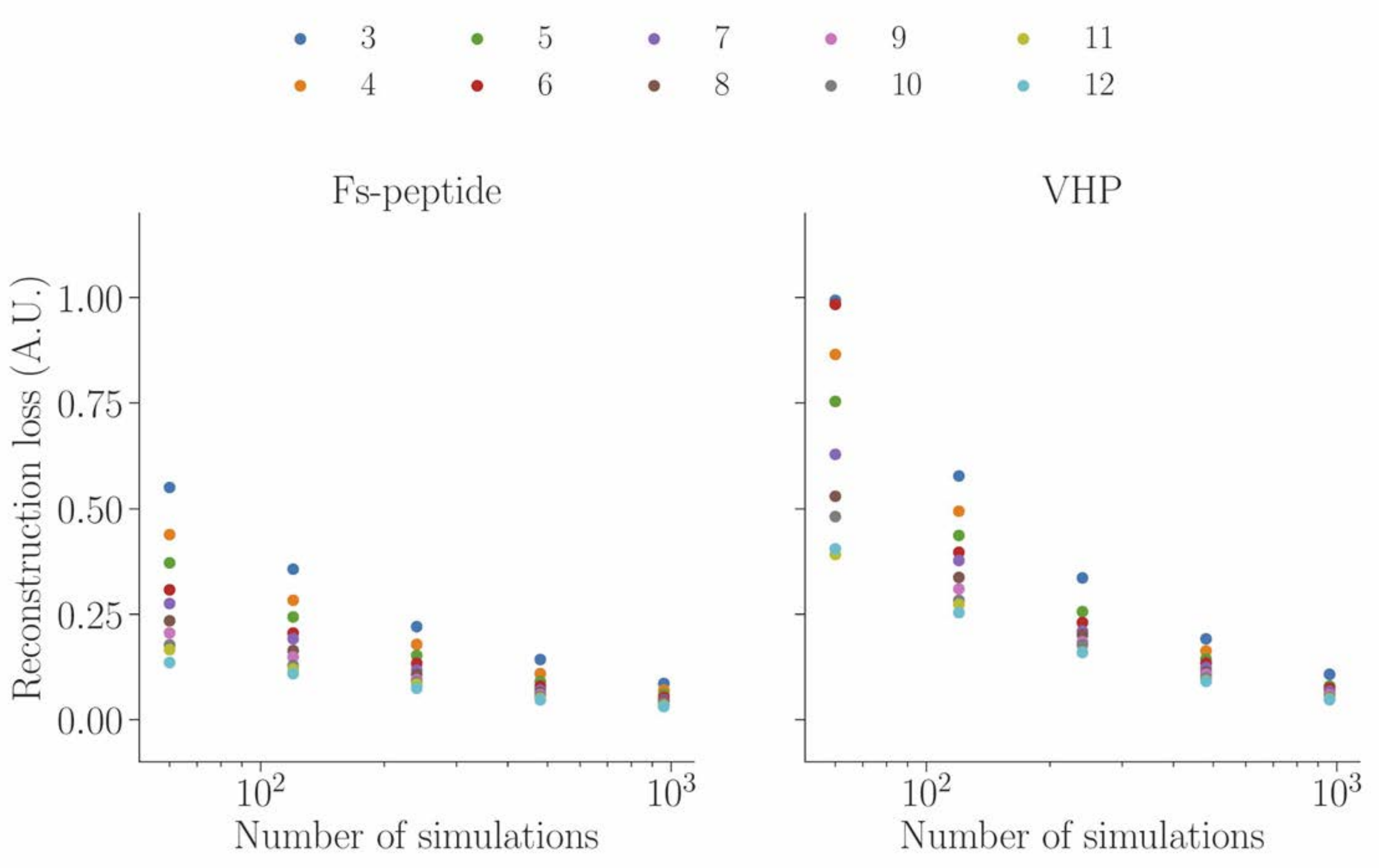
Table 2. Summary statistics of time taken by the individual components of our workflow: (1) train and infer from the CVAE for each system, and (2) running the MD simulation.

Collaboration with AR (ANL): see arxiv

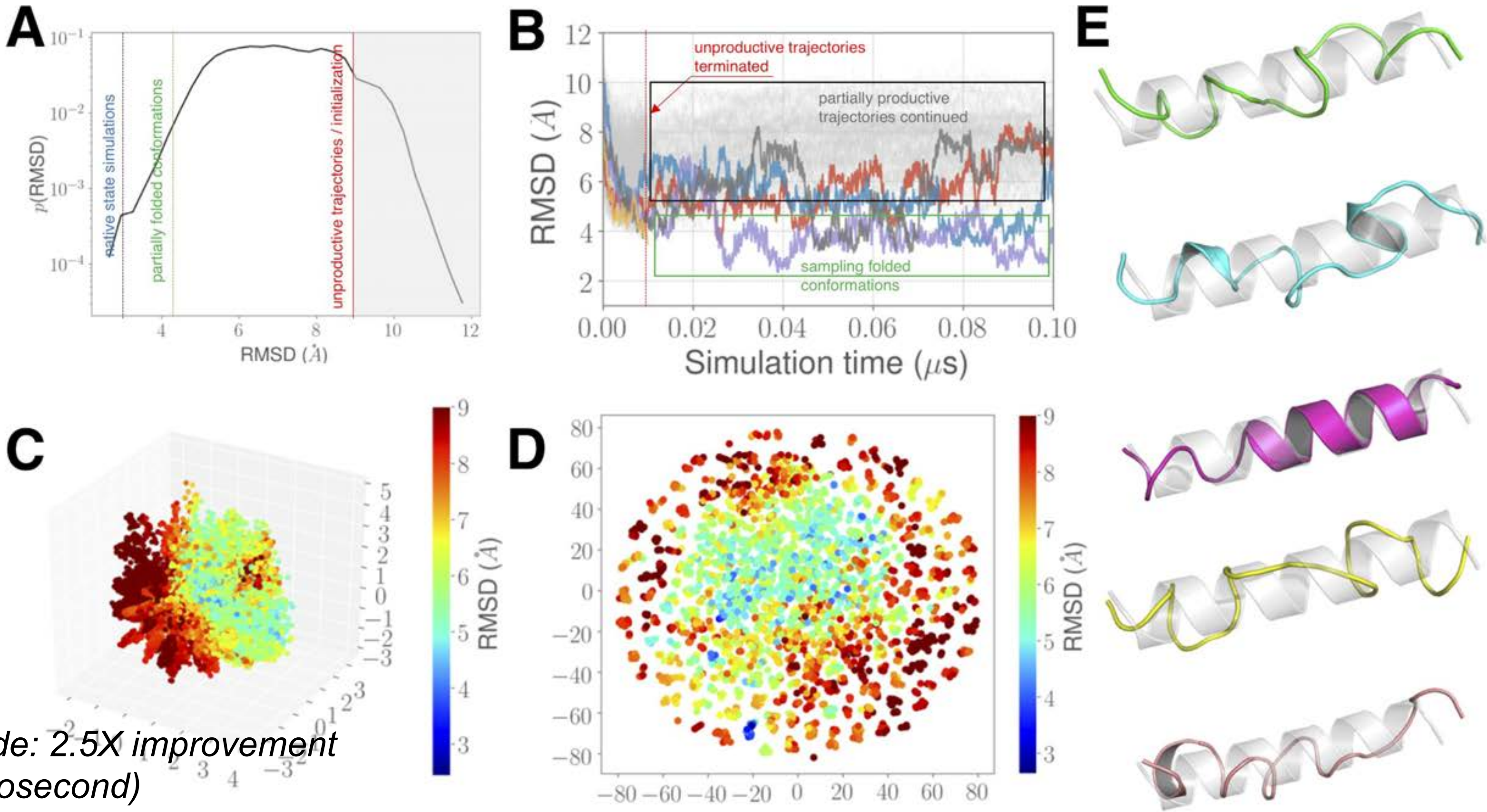
DeepDriveMD: DL driven Adpative MD



Importance of Adaptive Execution



DeepDriveMD: DL driven Adpative MD



For FS-peptide: 2.5X improvement (6 vs 14 microsecond)

Learning EveryWhere: Classification

- **HPCforML**: Using HPC to execute and enhance ML performance, or using HPC simulations to train ML algorithms (theory guided machine learning), which are then used to understand experimental data or simulations.
- **MLforHPC**: Using ML to enhance HPC applications and systems; Big Data comes from the computation
- *Context: Computational Science Effective consumer of* ***HPCforML****; innovative producers of* ***MLforHPC***

MLforHPC: Classification

MLforHPC: Using ML to enhance HPC applications and systems

- **MLaroundHPC:** Using ML to learn from simulations and produce learned surrogates for the simulations or parts of simulations.
- **MLControl:** Using HPC simulations in control of experiments and in objective driven computational campaigns. Simulation surrogates allow real-time predictions.
- **MLAutoTuning:** Using ML to configure (autotune) ML or HPC simulations.
- **MLafterHPC:** ML analyzing results of HPC as in trajectory analysis and structure identification in biomolecular simulations
- Focus on first two arguably most important, rewarding and difficult

MLaroundHPC: Examples

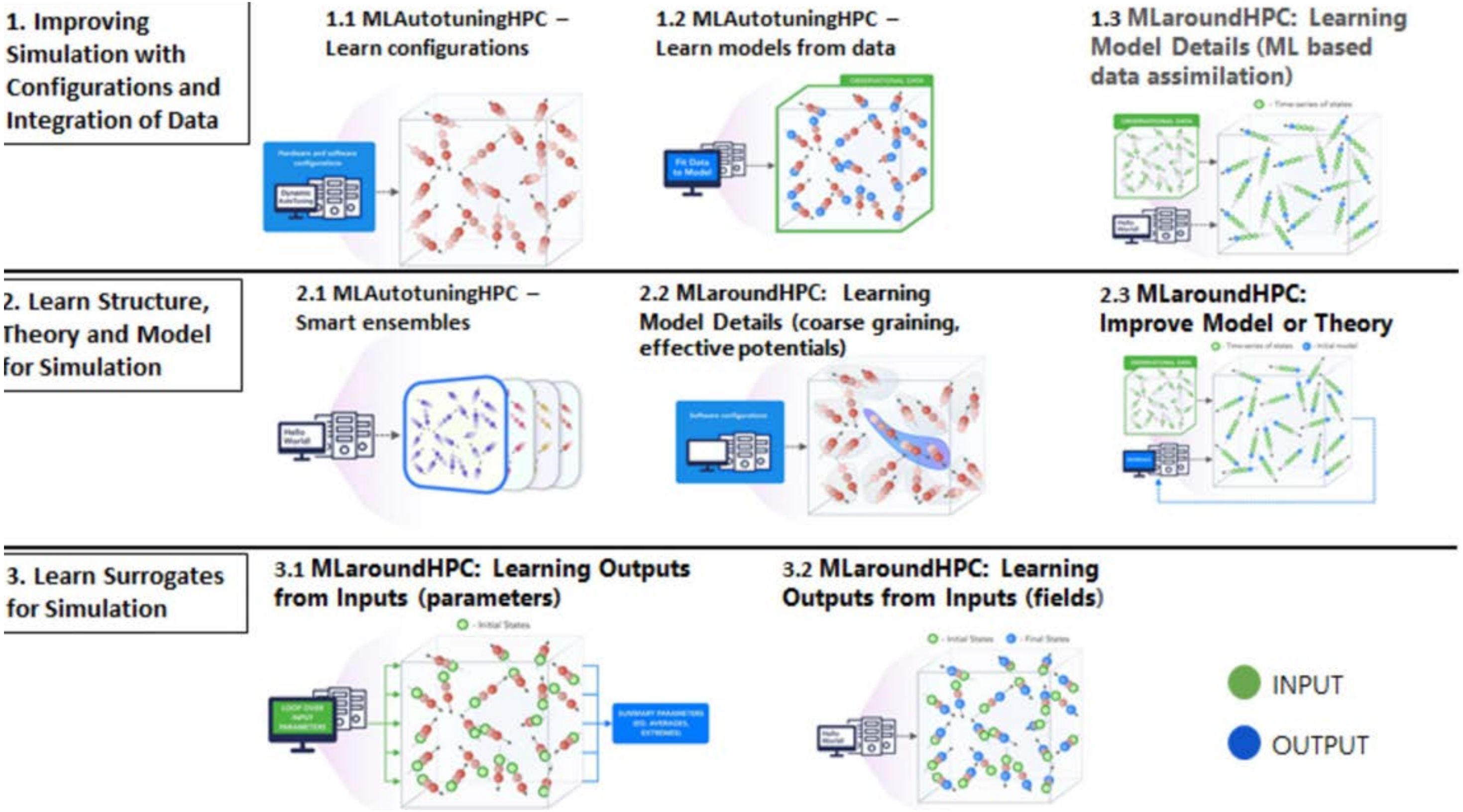
- **MLaroundHPC: Learning Outputs from Inputs**:
  - Simulations performed to directly train an AI system, rather than AI system being added to learn a simulation (includes **SimulationTrainedML)**
- **MLaroundHPC: Learning Simulation Behavior**
  - ML learns behaviour replacing detailed computations by ML **surrogates.**
- **MLaroundHPC: Faster and Accurate PDE Solutions**
  - High-dimensional non-linear PDEs such as diffusion equation using Deep Galerkin Method
- **MLaroundHPC: New Approach to Multi-scale**
  - Effective potential is analytic, quasi-empirical or quasi-phenomological potential that combines multiple effects into a single potential.
  - **...**

MLaroundHPC: Functional Drivers

Three primary functional drivers of ML driving HPC

- **Improving Simulations:**
  - Broad range of possibilities, from learning to configure and select simulations; improving models using simulations
- **Learning Structure, Theory and Model for Simulations:**
  - Improve the model or theory or underlying principles
  - As simulations proceed, model incrementally improved
- **Learn to make surrogates:**
  - Learn the function representing the output of a simulation to determine either the parameters or the effective fields

MLaroundHPC: Functional Drivers



MLaroundHPC: Modes and Examples

Three modes and mechanisms of integrating ML with HPC

- **Substitution:** Create a surrogate as a substitute of an essential element of an expensive calculation. Surrogate can be used for a multi-scale (coarse-grained) modeling or substituting the expensive part of a calculation.
  - Train NN using DFT QM calculation, use NN for ab-initio MD
  - Substitute FF using NN derived potential without loss of accuracy
- **Assimilation:** External data integrated into physics-based models, which are then assimilated into traditional simulations, e.g., improving FF.
  - Approximating sub-grid processes for cloud parameterization. Training a NN to learn from a multi-scale model all atmospheric sub-grid processes
- **Control and Adaptive Execution:** (Ensembles) of simulation are controlled towards important and interesting parts of phase space.
  - Sometimes specifics ensembles
  - Sometimes the entire campaign is governed by an objective (ODED)

MLaroundHPC: Open Questions

Algorithms Benchmarks and Methods

- **Crossover point:** Understand crossover points at which learning and prediction based approaches are better than traditional HPC?
  - Crossover as function of scale, algorithmic complexity, model and data
- **Canonical problems:** Which learning methods work, why and when?
  - Towards benchmarks and proxy applications
- **Measuring effective performance:** Integrated performance of learning & HPC
  - Comparing performance against vanilla approaches
  - Interplay between raw performance, learning performance and effective performance influence the design of applications and systems ?

MLaroundHPC: Open Questions

System Software and Runtime Systems

- **Runtime flexibility of heterogeneous tasks:** Concurrent & integrated execution of simulations (S) and learning (B) over wide range of scenarios
  - Simulations used to generate training data vs. ML only inference phase
  - Learning after simulations (MLafterHPC) vs learning intertwined with simulations (**MLaroundHPC; surrogates**)
- **Single runtime to support concurrent execution of ML and HPC** given diverse coupling between S, L and Experiment (E):
  - Control, data volumes / rates and latency
  - All classes of MLforHPC and granularity of coupling?

MLaroundHPC: Open Questions

Hardware and Platform Issues

- **Fraction of time is spent in ML vs HPC as a function of problem size:**
  - Ratio of L:S small: class supercomputer platform linked to separate learning system? Ratio of L:S large: Tightly integrated systems?
  - How do provide a balanced systems across application types?
- **Role and importance of heterogeneous accelerators**
  - 3 levels of heterogeneity: CPU, GPU and ML-accelerators
  - RNN for time series vs CNN

Reference Architecture: Scaling Considerations

- Strong Scaling Considerations:
  - Strong Scaling of individual L: Enabling L to achieve near real-time training and prediction to control or steer S
    - Build low dimensional representation of states from trajectory analysis
  - Strong Scaling of Integrated L + S: Enabling simulation-trained models to determine where to sample in space
    - RL driven approach to go through large chemical space efficiently
- Weak Scaling Considerations:
  - Weak Scaling of L: Many learning models concurrently
    - Ensemble learning; multiple surrogates, may the best surrogate win
  - Weak Scaling of Integrated L + S: Multiple instantiations of L and S
    - Model-based design of experiments (MBDOE); objective driven experiments and learning effective potentials

Reference Architecture: Resource Management Considerations

- **Resource Management Considerations:**
  - Must consider streaming data so as to include experimental and observational data
  - Must support the learning on the edge, cloudlet or cloud / HPC
  - General Properties of applications
    - Adaptive: Task graph and plan will change based upon intermediate results and data availability
    - Dynamic: Resource availability and performance is time dependent
    - Heterogeneous workflows: Multiple distinct components (E, L and S), and different instances of each component
  - Resource management and system software challenges are similar to adaptive + streaming workflow!

Open Issues and Challenges

- Which learning methods are most effective?
- New algorithmic approaches based upon “effective learning” ?
- Is there a general multi-scale approach using surrogates (MLaroundHPC) ?
- Advances in Uncertainty Quantification
- What are appropriate system frameworks to implement interaction between E, S and L components?
  - Single reference architecture for all 4 categories?
- Runtime system challenges for balanced execution of real & surrogate models?
  - Workload management, resource management and scheduling
  - Strong and weak scaling challenges
- Application / scenario agnostic definition of Effective Performance
- ….

Summary

- **Algorithmic and methodological advances are needed**
  - Current performance tightly coupled to hardware unsustainable
- ML enhance the **effective performance** of HPC simulations
  - 20x over best ensemble based approach!
  - DL driven MD on Summit: 2.5x (FS-peptide)
- **Learning Everywhere**
  - Classification and Examples
  - Open Issues and challenges

Thank You!

Learning Everywhere: G Fox
ExTASY: Cecilia Clementi Clementi
DeepDriveMD: Arvind / ANL
RADICAL Cybertools: http://radical.rutgers.edu
CANDLE-INSPIRE: Rick Stevens, Peter Coveney, John Chodera