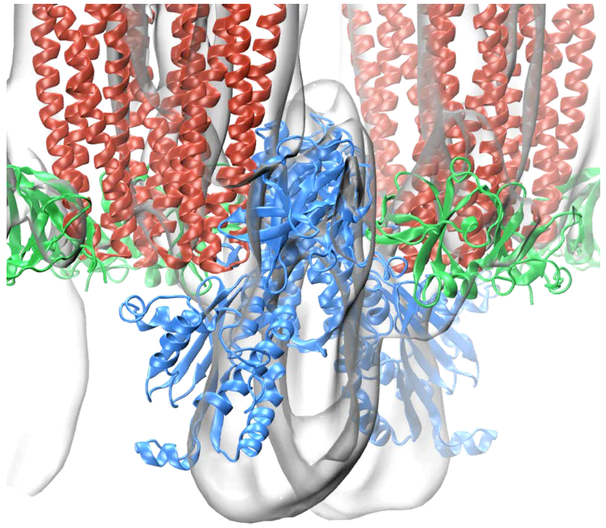
Improving NAMD Performance and Scaling on Heterogeneous Architectures

David J. Hardy and Julio D. C. Maia
NIH Center for Macromolecular Modeling and Bioinformatics
Theoretical and Computational Biophysics Group
Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana-Champaign

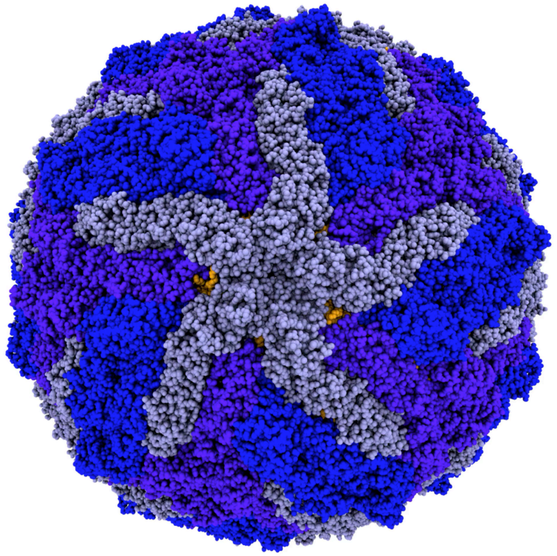
NAMD Scalable Molecular Dynamics

J. Phillips, D. Hardy, J. Maia, et al. *J. Chem. Phys.* 153, 044130 (2020) https://doi.org/10.1063/5.0014475

- Code written in C++/Charm++/CUDA
- Performance scales to hundreds of thousands of CPU cores and tens of thousands of GPUs
  - **Large systems (single copy scaling)**
  - **Enhanced sampling (multiple copy scaling)**
- Runs on laptops up to supercomputers
- Runs on AWS cloud, MS Azure
- TCL/Python script as input file
  - Workflow control
  - Method development at higher level
- Structure preparation and analysis with VMD
  - QwikMD



E. coli chemosensory array

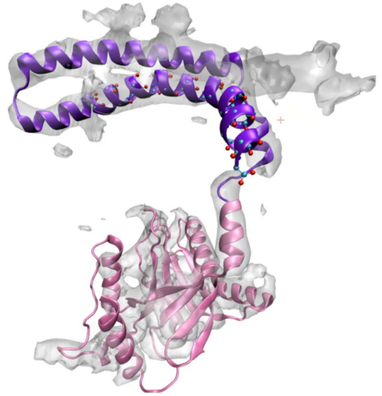


Zika Virus

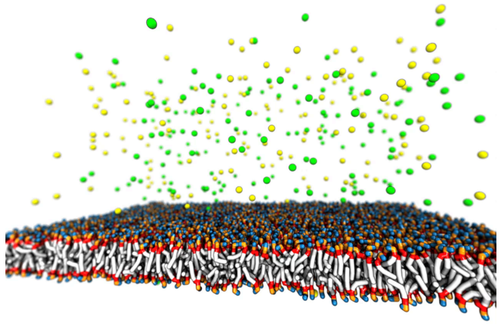
NAMD: http://www.ks.uiuc.edu/Research/namd/

NAMD Highlights

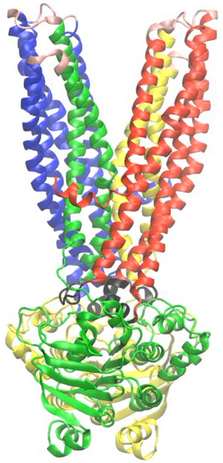
- User defined forces
  - Grid forces
  - Interactive molecular dynamics
  - Steered molecular dynamics
- Accelerated sampling methods
  - Replica exchange
- Collective variables (Colvars)
  - Biased simulation
  - Enhanced sampling
- Alchemical transformations
  - Free energy perturbation (FEP)
  - Thermodynamic integration (TI)
  - Constant-pH molecular dynamics
- Hybrid QM/MM simulation
  - Multiple QM regions



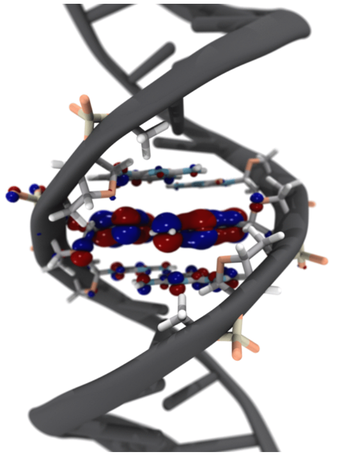
Proteosome (MDFF+IMD)



Membrane vesicle fusion and formation (grid forces)



ABC transporter mechanism (Colvars)



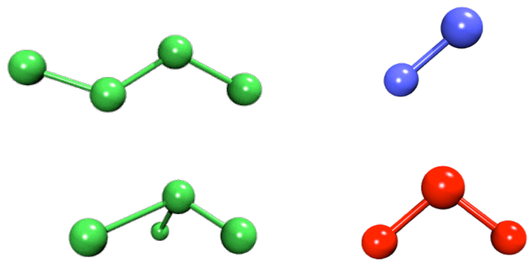
DNA QM/MM simulation

Complete List of NAMD Features: https://www.ks.uiuc.edu/Research/namd/2.14/ug/

Molecular Dynamics Simulation

- Most fundamentally, integrate Newton's equations of motion:

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i = -\vec{\nabla} U(\vec{R})$$ ⟵ integrate for up to billions of time steps



$$U(\vec{R}) = \underbrace{\sum_{bonds} k_i^{bond}(r_i - r_0)^2}_{U_{bond}} + \underbrace{\sum_{angles} k_i^{angle}(\theta_i - \theta_0)^2}_{U_{angle}} +$$

$$\underbrace{\sum_{dihedrals} k_i^{dihe}[1 + \cos(n_i\phi_i + \delta_i)]}_{U_{dihedral}} +$$

most of the computational work ⟶

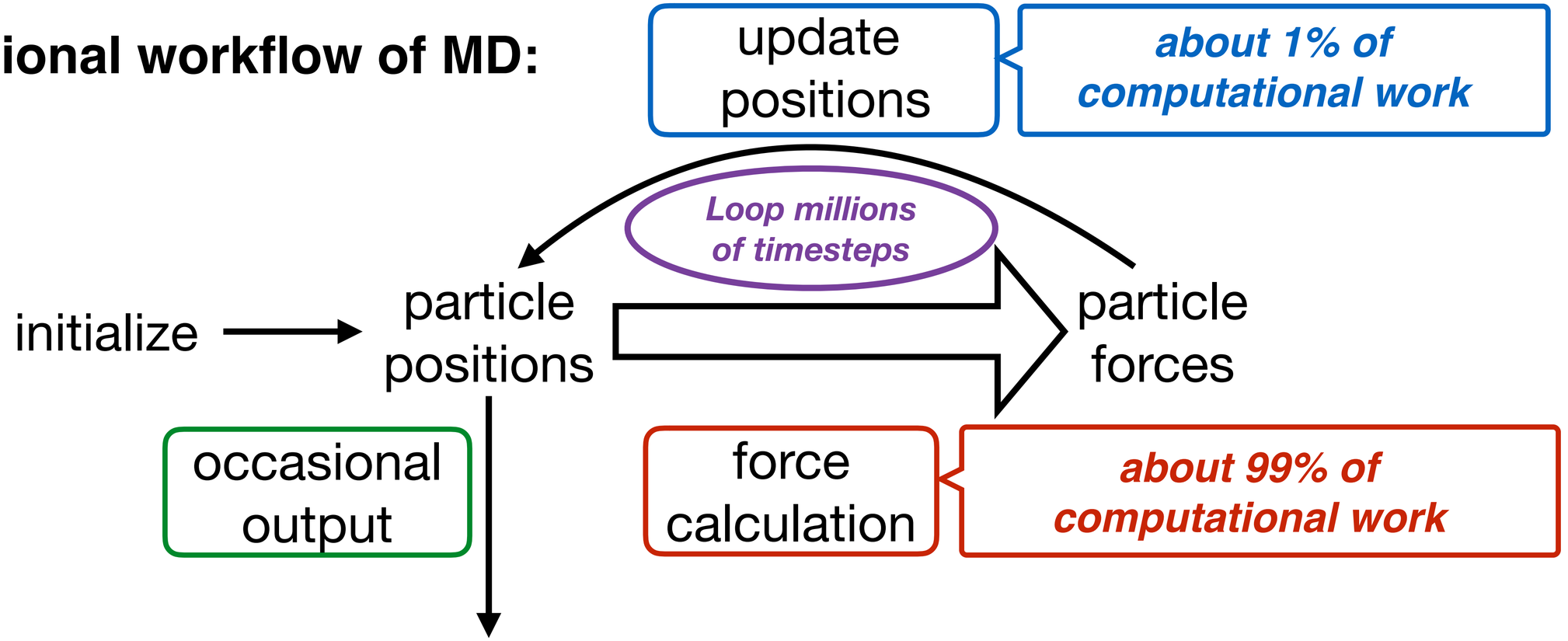
$$\underbrace{\sum_i \sum_{j \neq i} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right]}_{\text{(Lennard-Jones)}} + \underbrace{\sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}}_{U_{nonbond}}$$

(electrostatics)

NIH National Institutes of Health

Parallelism for MD Simulation Limited to Each Timestep

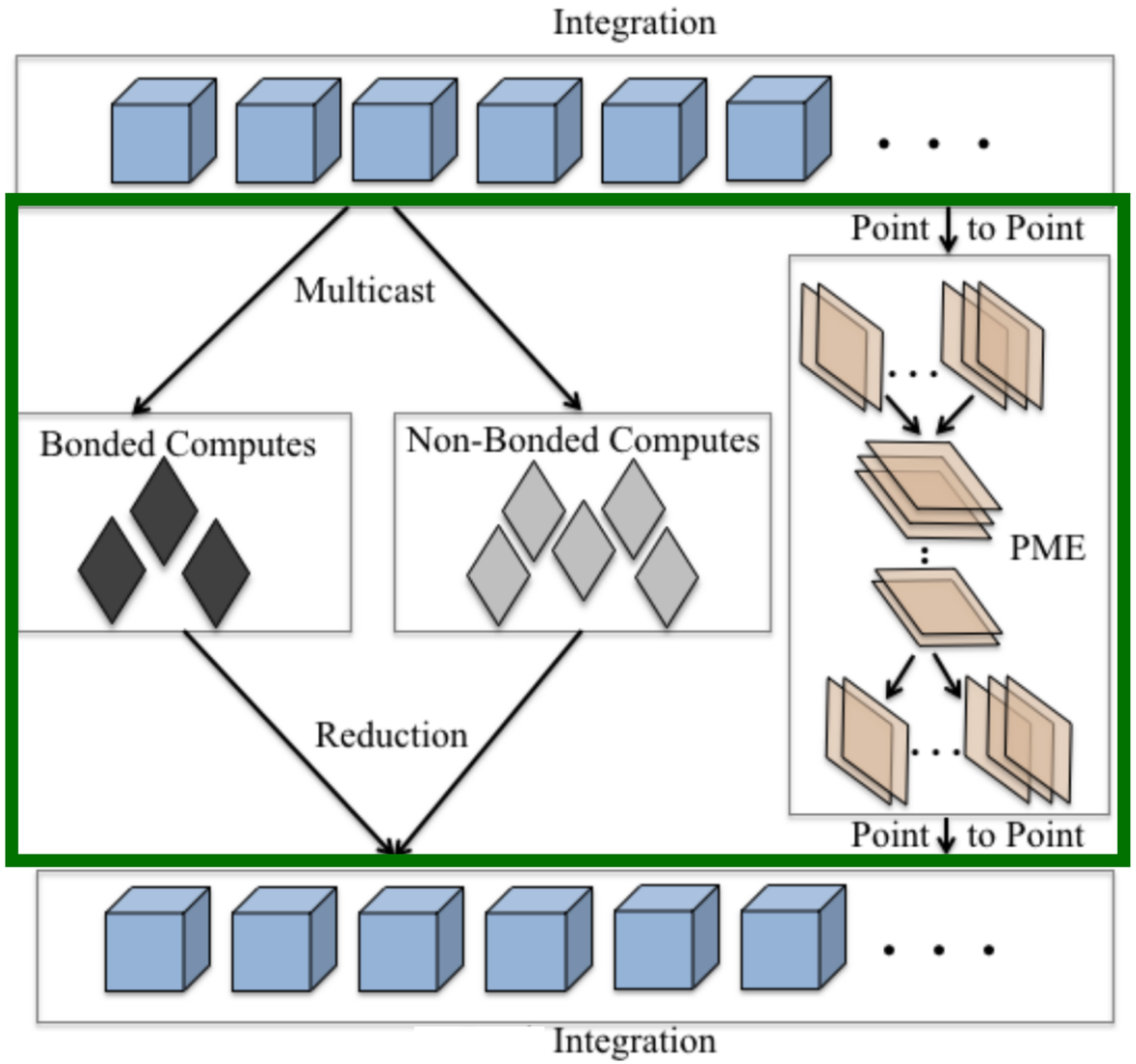
Computational workflow of MD:



reduced quantities (energy, temperature, pressure)

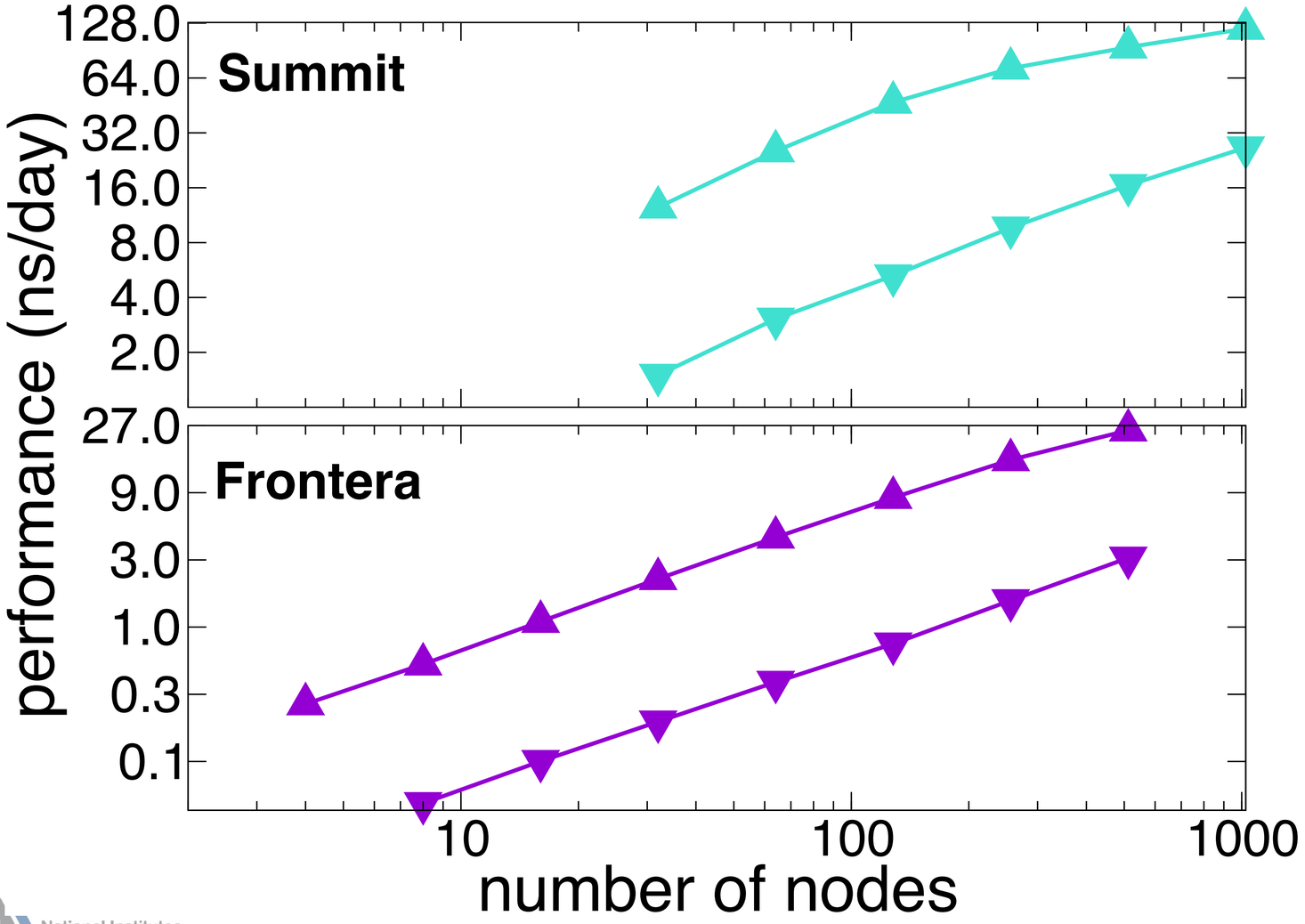
position coordinates (trajectory snapshot)

NAMD 2.14 Decomposes Force Terms into Fine-Grained Objects for Scalability

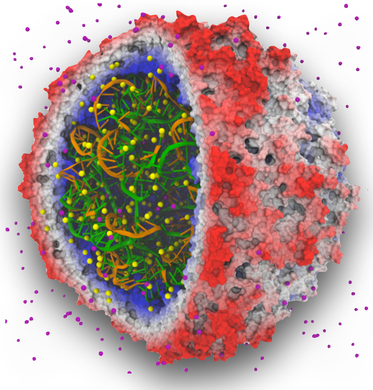


Offload forces to GPU

NAMD 2.14 Excels at Scalable Parallelism on CPUs and GPUs



Replications of the Satellite Tobacco Mosaic Virus (STMV)



▲ = 5x2x2 grid = 21M Atoms

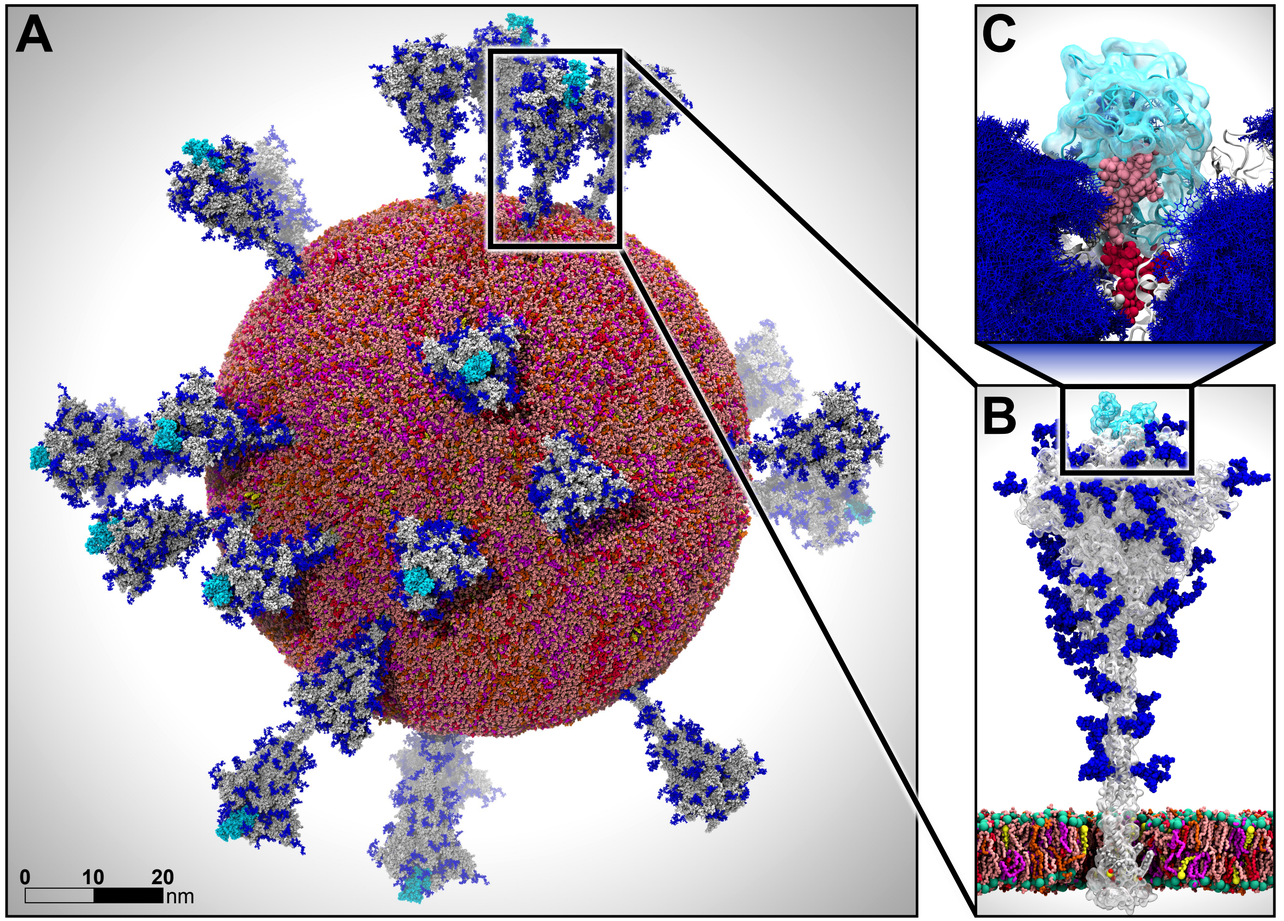
▼ = 7x6x5 grid = 224M Atoms

NIH National Institutes of Health

NAMD 2.14 Simulating SARS-CoV-2 on Summit

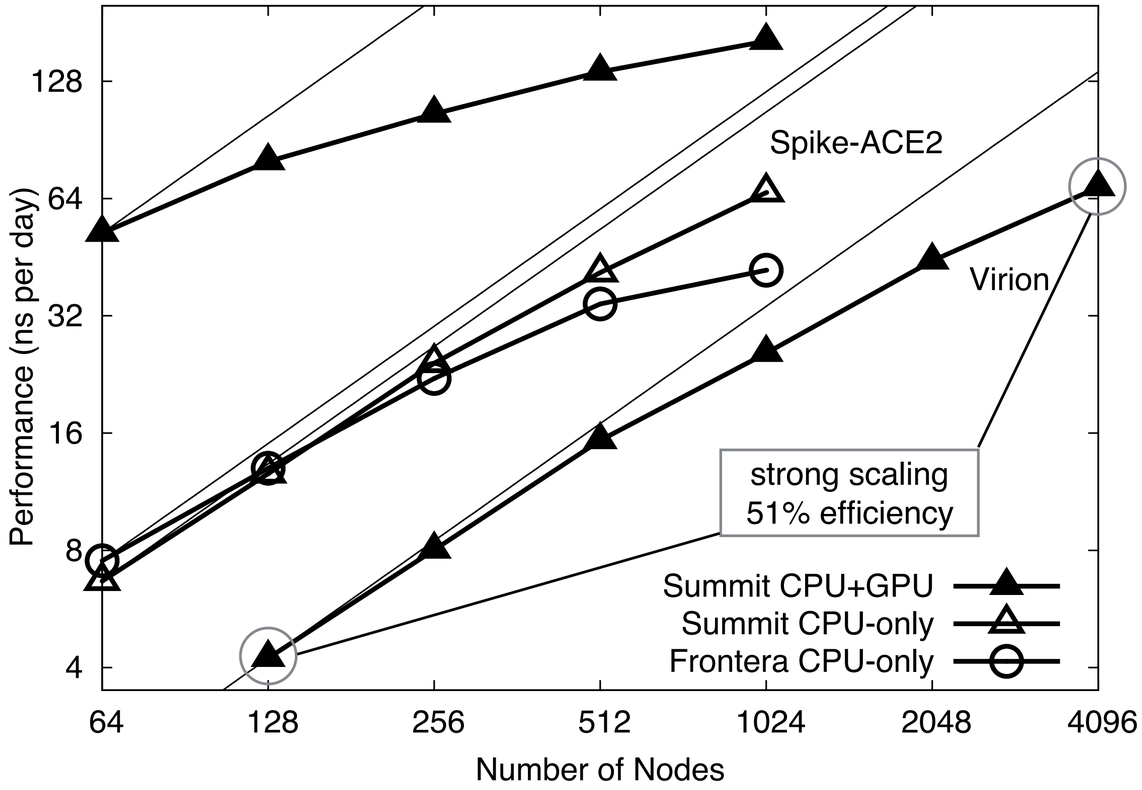
Collaboration with Amaro Lab at UCSD, images rendered by VMD

(A) Virion, (B) Spike, (C) Glycan shield conformations

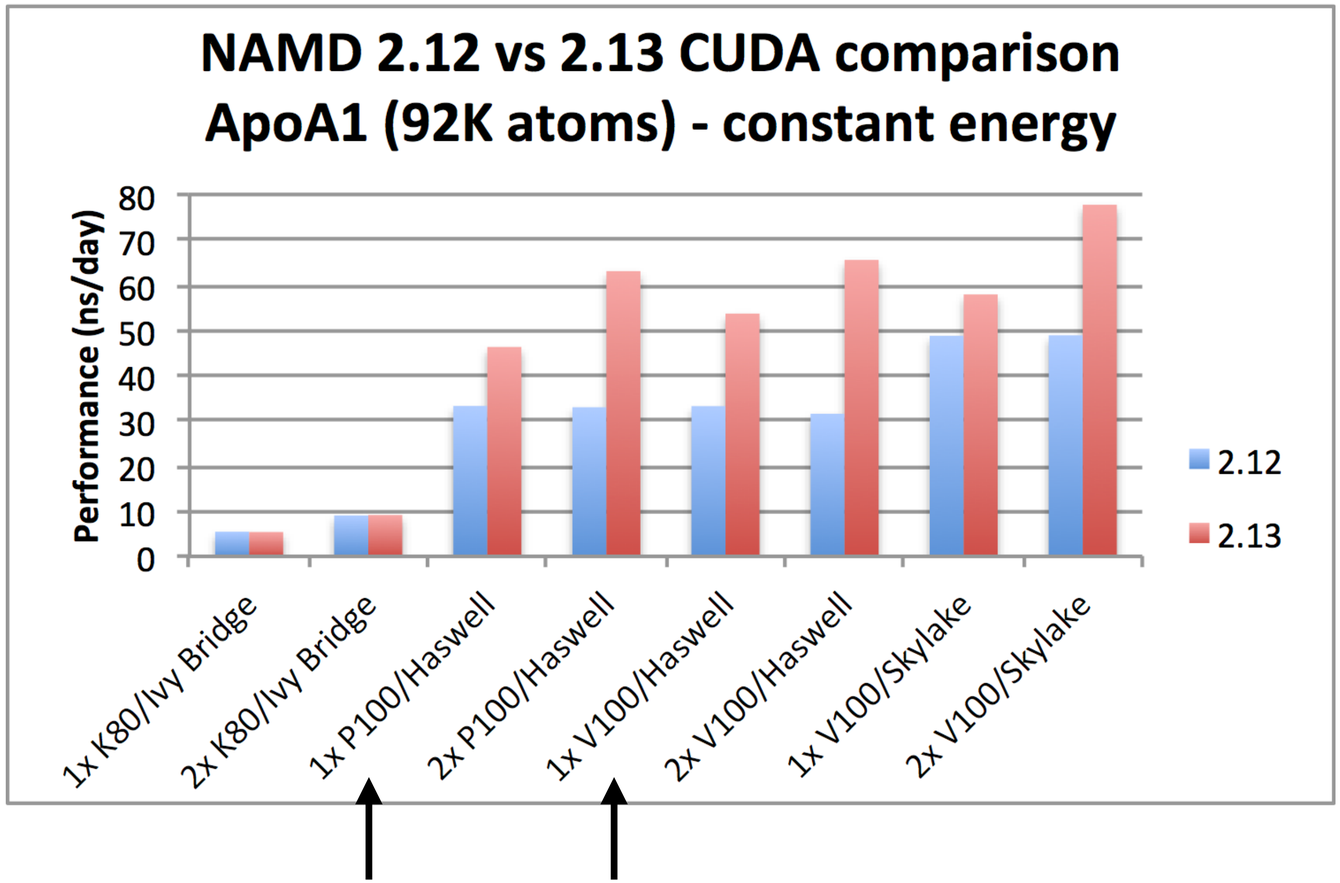


Scaling performance:

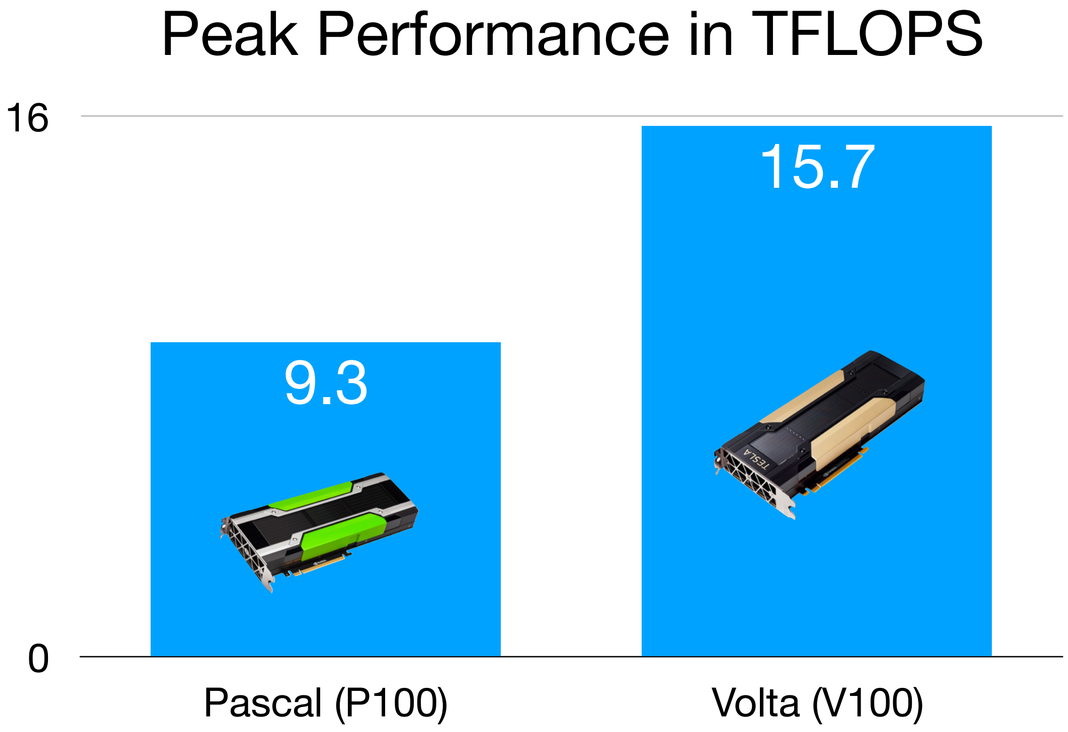
- ~305M atom virion
- ~8.5M atom spike



Benchmarks on Single Nodes and Newer GPUs Reveal Problems

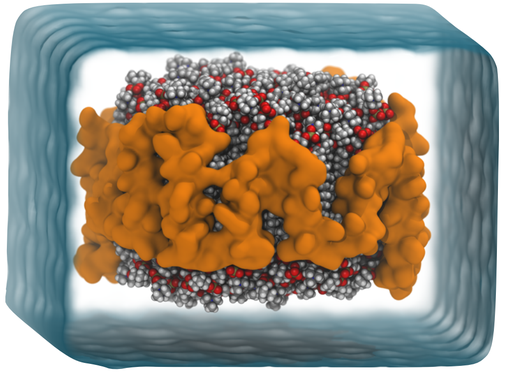


NAMD 2.13 (2018) has ~20% perf improvement from P100 to V100

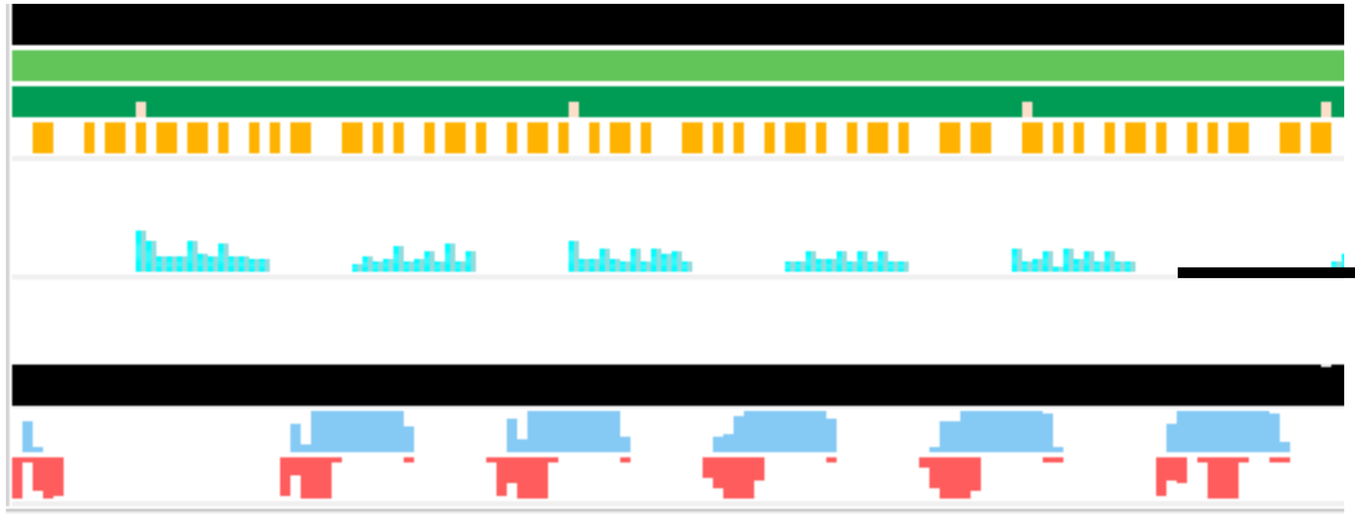


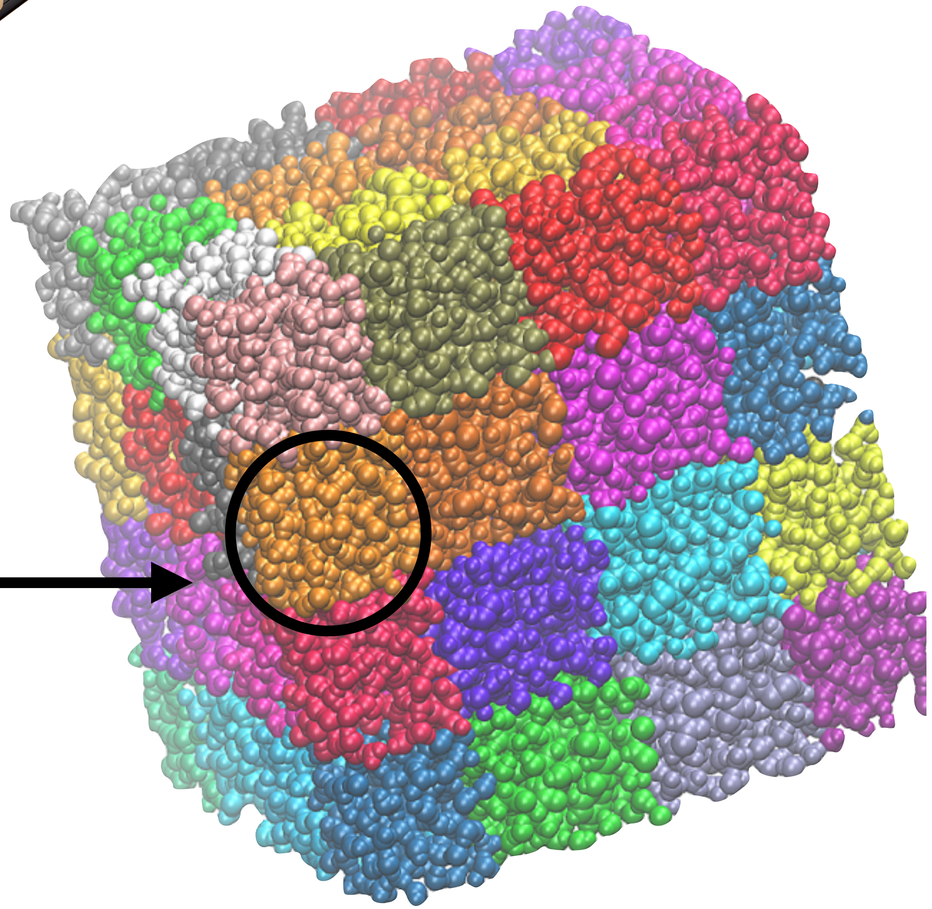
Hardware has ~70% perf improvement!

Profiling on Modern GPUs



Profiling ApoA1, 92k atoms
NAMD 2.13, 16 cores and 1 GPU Volta





Gaps in the blue strip = GPU is idle!

NIH National Institutes of Health

NAMD 2.13 and 2.14 Have Limited GPU Performance

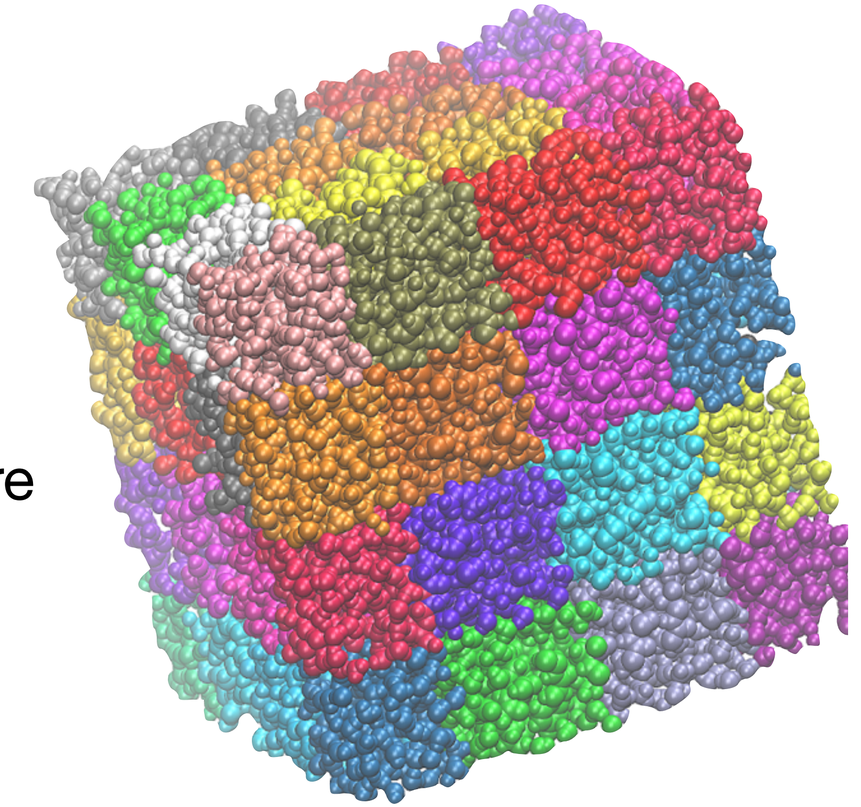
- Offloading force calculation is not enough!
- Overall utilization of modern GPUs is limited
- We want better single GPU performance
  - Majority of MD users run system sizes < 1M atoms on a single GPU
- Must transition from **GPU-offload** approach to **GPU-resident**!

The DGX-2 has 16 V100 GPUs but only 48 CPU cores: We need to do more GPU work with less CPU Power

NAMD 3.0: GPU-Resident NAMD

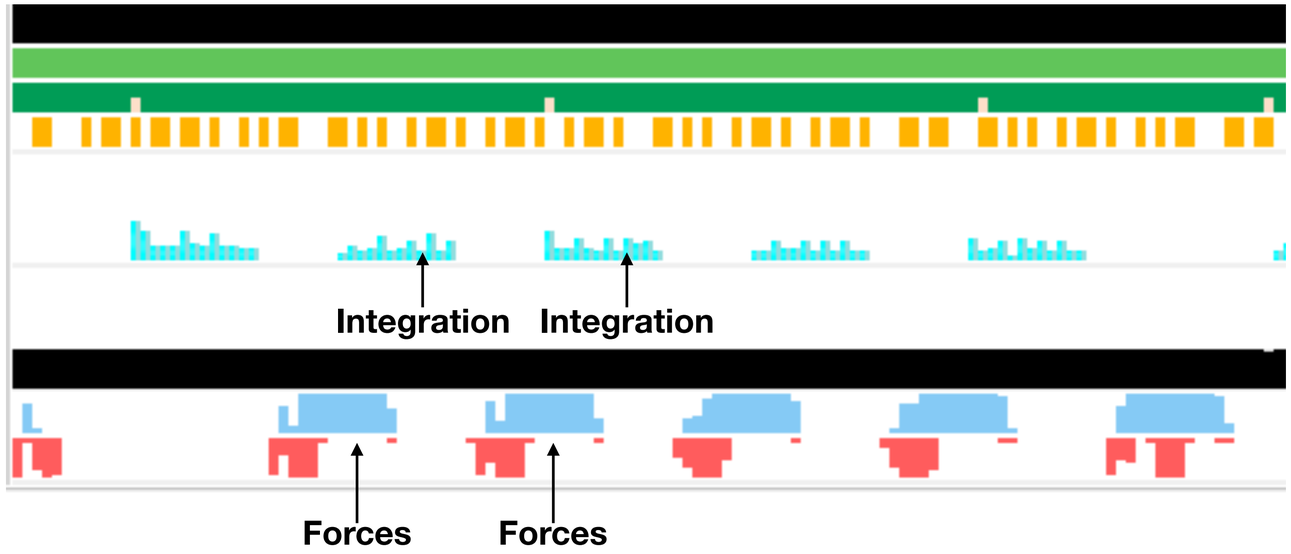
https://www.ks.uiuc.edu/Research/namd/3.0alpha/

- Fetches GPU force buffers directly from the force module
- Bypass any CPU-GPU memory transfers - *only call GPU kernels!*
- Convert forces in a *structure-of-arrays* (SOA) data structure using the GPU
- Invoke GPU Integration Tasks Once

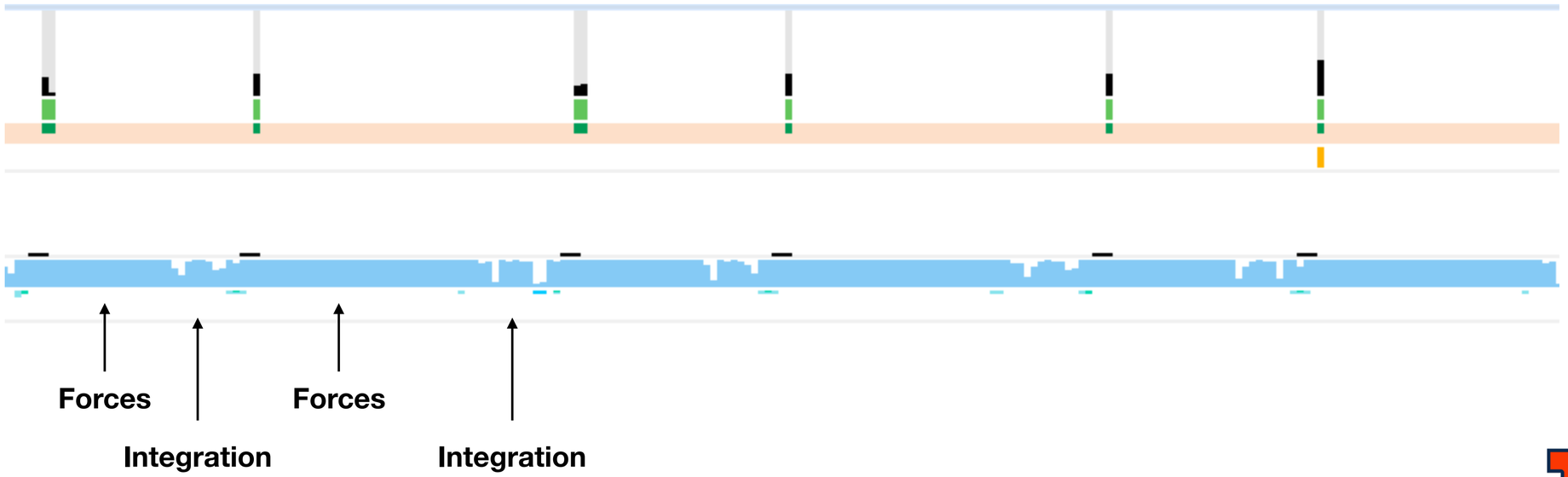


NAMD 3.0 Has Better GPU Utilization

NAMD 2.14
Gaps between GPU tasks



NAMD 3.0
No CPU bottlenecks

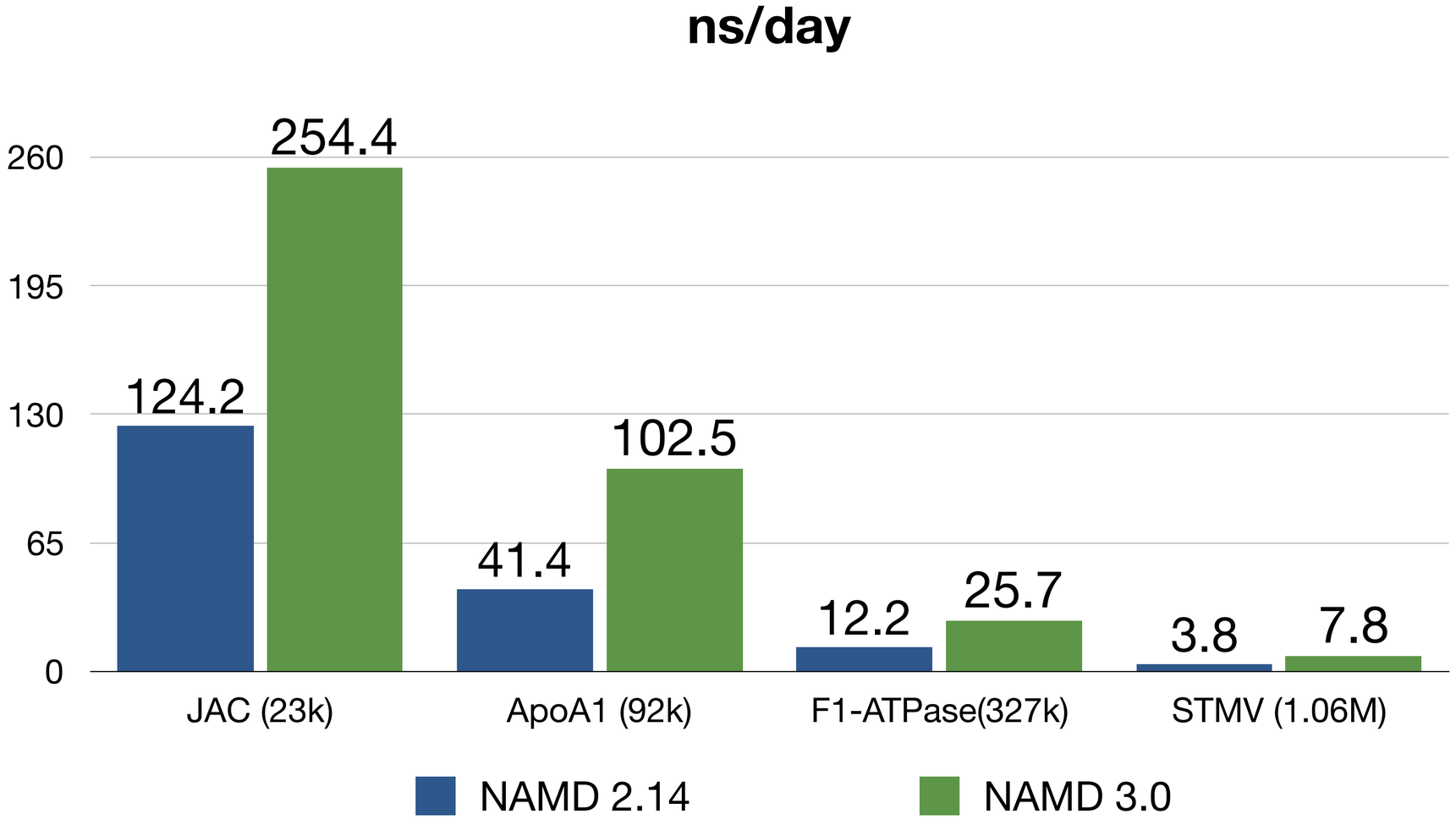


NAMD 3.0: Performance on Different Systems

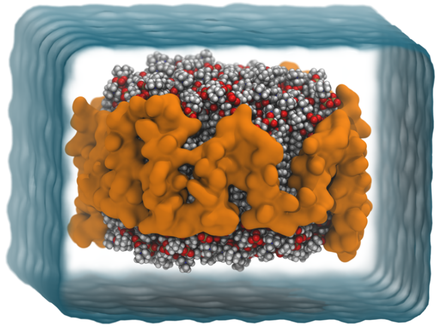
NVE
12A Cutoff
2fs timestep

Intel Xeon E5-2650 V2 w/ 16 physical cores

Nvidia Titan V

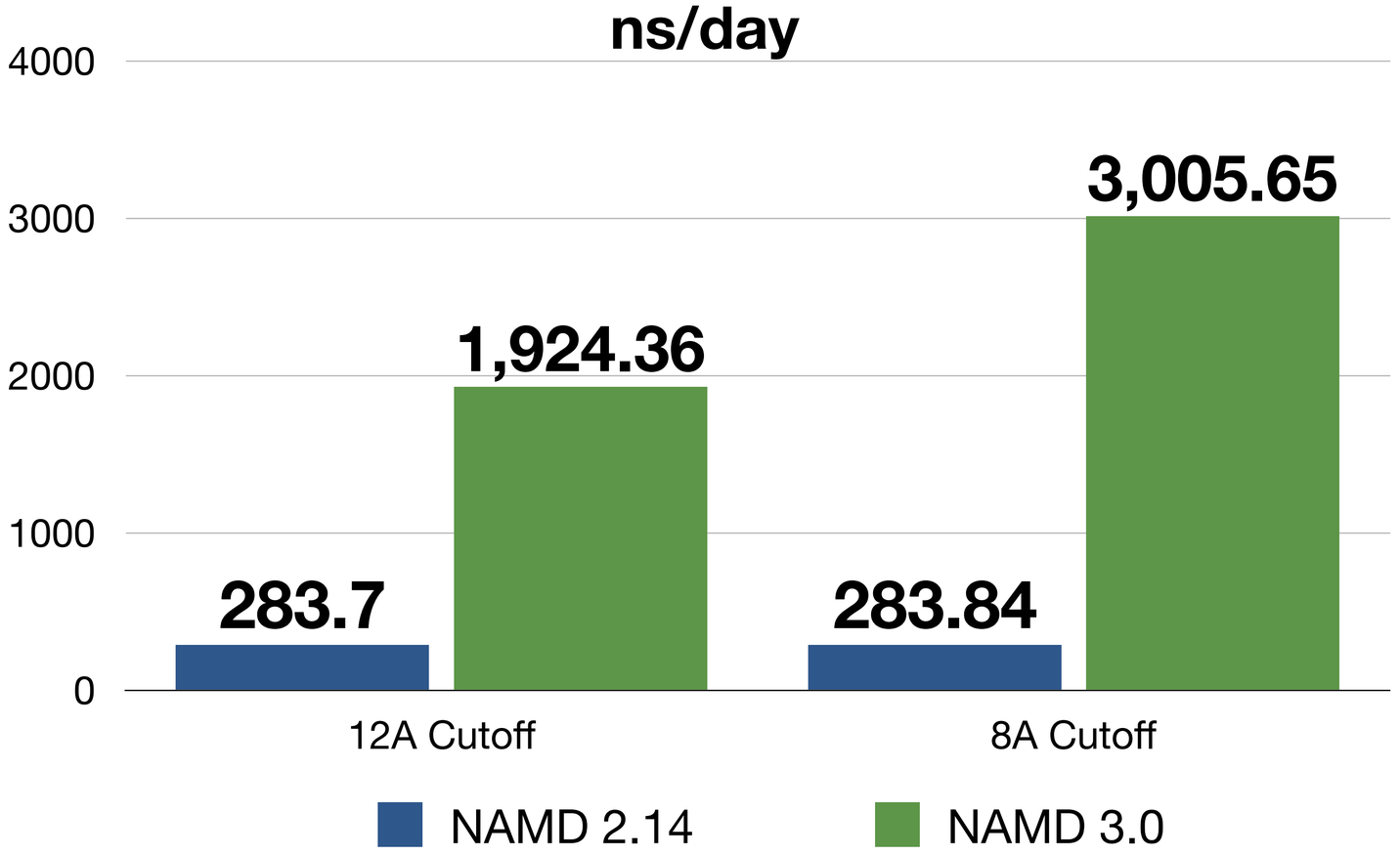


NAMD 3.0: Multi-Copy Performance - Aggregate Throughput With DGX-2



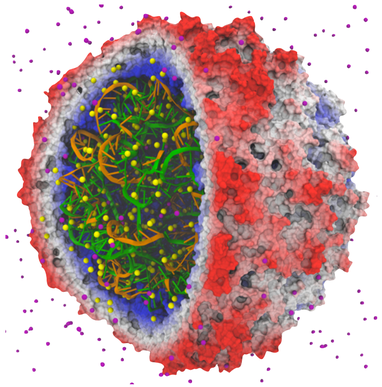
ApoA1
92k atoms

16 Replicas
1 for each NVIDIA V100

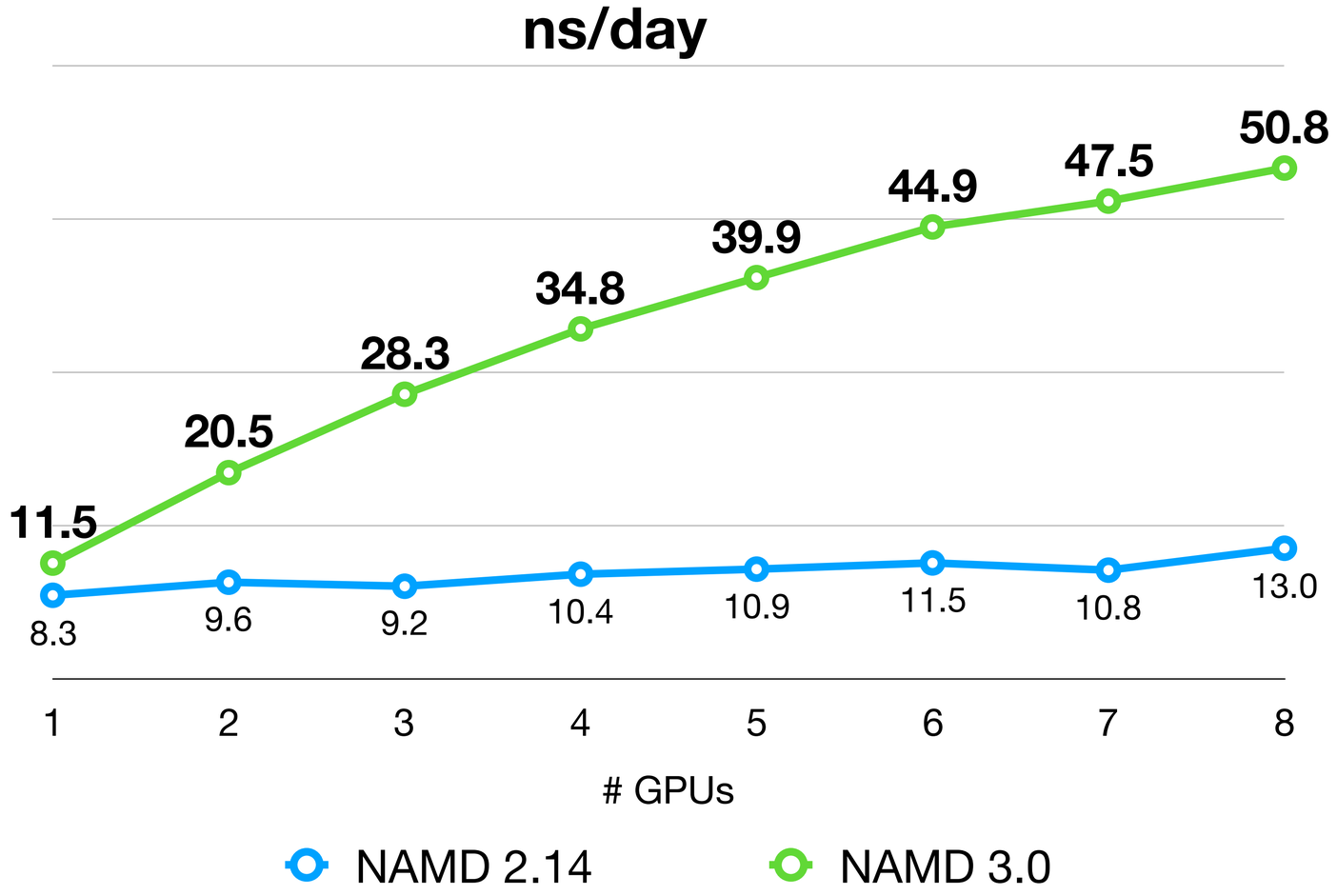


NIH National Institutes of Health

NAMD 3.0: Single trajectory - Multiple GPU Performance



STMV
1.06M atoms
2fs timestep
No PME yet



PME Impedes Scalability

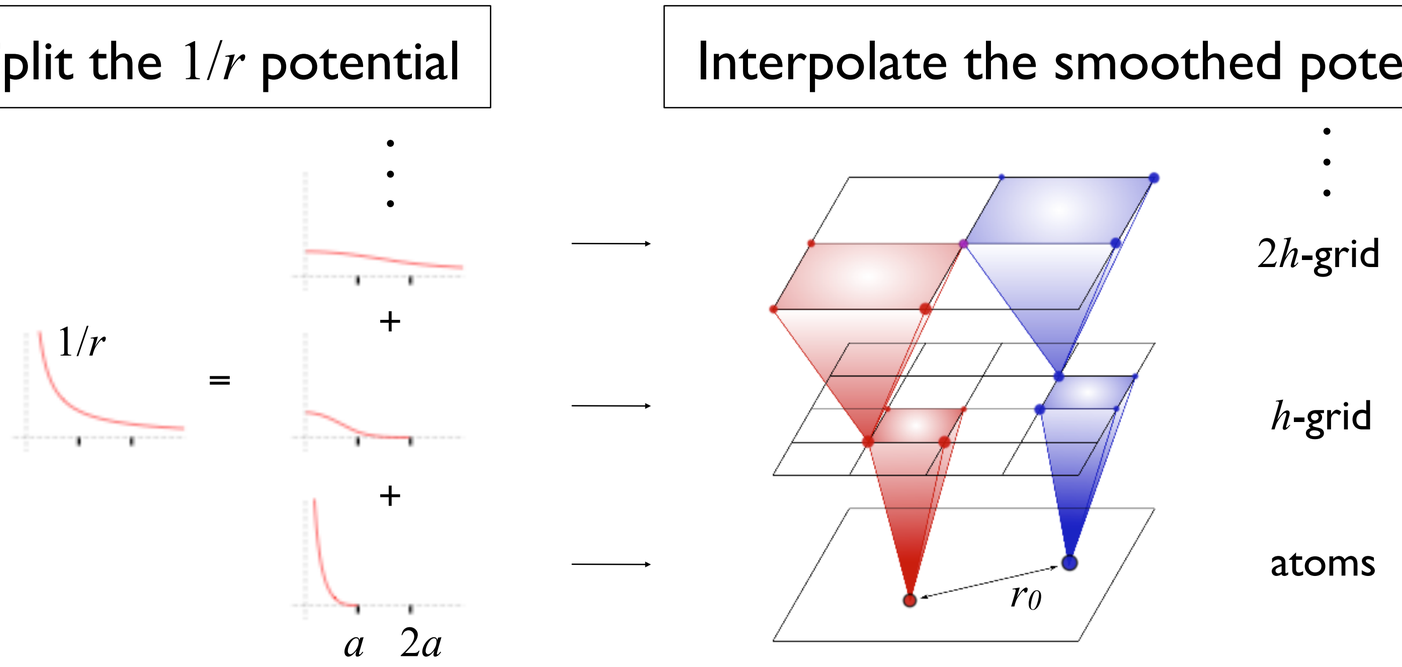
- For multi-node scaling, 3D FFT communication cost grows faster than computation cost
- For single-node multi-GPU scaling:
  - 3D FFTs are too small to parallelize effectively with cuFFT
  - Too much latency introduced with pencil decomposition and cuFFT 1D FFTs
  - Is task-based parallelism best, delegating one GPU for 3D FFTs and reciprocal space calculation?
  - Requires gathering all grid data to that one GPU and being careful to not overload it with other work
- Why not use a better scaling algorithm, such as MSM?

Multilevel Summation Method (MSM)

D. Hardy, et al. *J. Chem. Theory Comput.* 11(2), 766-779 (2015) https://doi.org/10.1021/ct5009075

D. Hardy, et al. *J. Chem. Phys.* 144, 114112 (2016) https://doi.org/10.1063/1.4943868

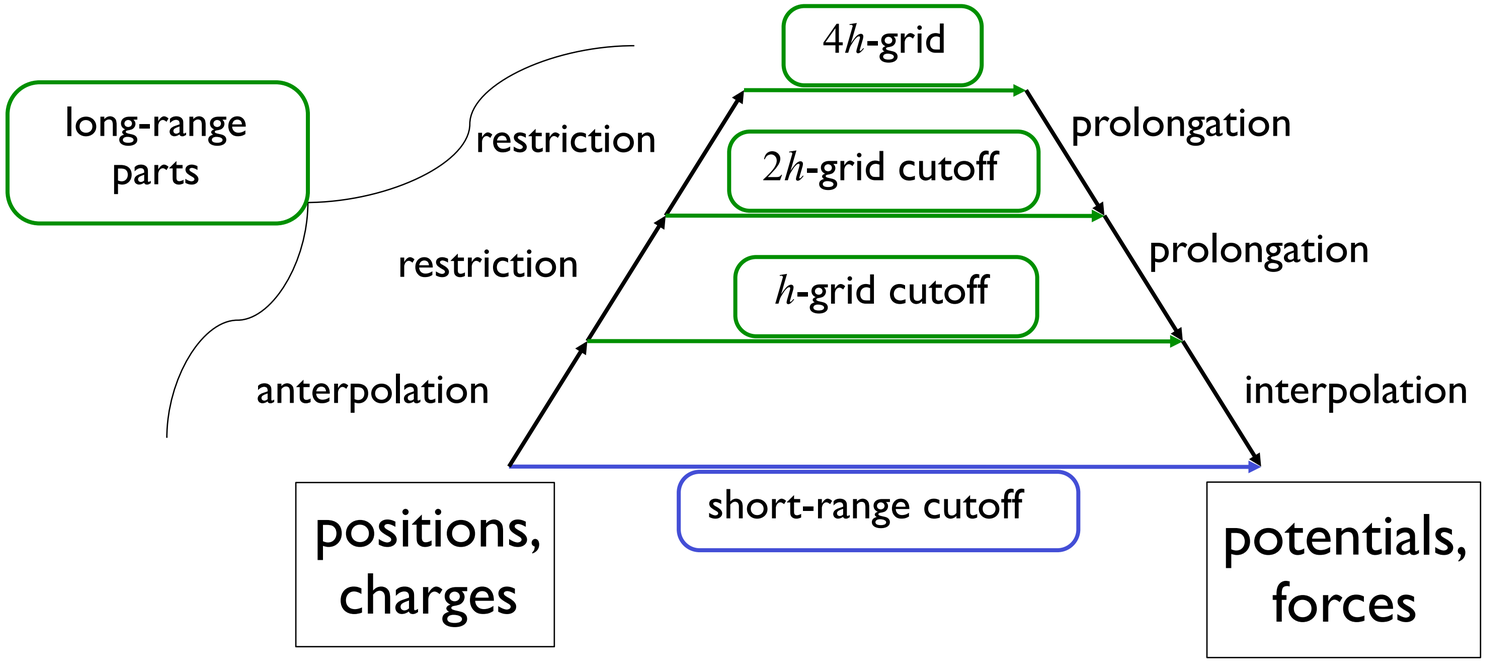
- Split the $1/r$ potential into a short-range cutoff part plus smoothed parts that are successively more slowly varying. All but the top level potential are cut off.
- Smoothed potentials are interpolated from successively coarser grids.
- Finest grid spacing h and smallest cutoff distance a are doubled at each successive level.



MSM Calculation is $O(N)$

force = **exact short-range part** + **interpolated long-range part**

Computational Steps



grid cutoff ⇔ 3D convolution
anterpolation ⇔ PME charge spreading
interpolation ⇔ PME force interpolation

Periodic MSM: Replaces PME

- Previous implementation was fine for non-periodic boundaries but insufficient for periodic boundary conditions
  - Lower accuracy than PME, requires system to be neutrally charged
- New development for MSM:
  - Interpolation with periodic B-spline basis functions gives same PME accuracy
  - Handle infinite 1/r tail as reciprocal space calculation of top level grid
  - Number of grid levels can be terminated long before reaching a single point; use it to bound size of FFT
  - Communication is nearest neighbor up the tree to the top grid level

Extending NAMD 3.0 to multiple nodes

- Reintroducing Charm++ communication
  - Fast GPU integration calls the force kernels directly
  - Unused Sequencer user-level threads are put to sleep
  - Awaken threads for atom migration between patches and coordinate output
- Will GPU direct messaging be the best alternative?
  - Charm++ support is being developed

Additional Challenges for NAMD

- Feature-complete GPU-resident version
  - NAMD 3.0 for now supports just a subset of features
- Incorporating Colvars (collective variables) force biasing
  - Poses a significant performance penalty without reimplementing parts of Colvars on GPU
- Introducing support for other GPU vendors
  - AMD HIP port of NAMD 2.14, still working on 3.0
  - Intel DPC++ port of non-bonded CUDA kernels

Intel GPUs

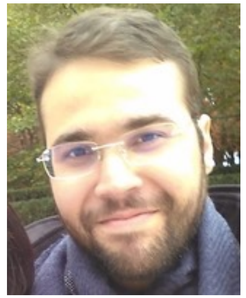
AMD GPUs

Acknowledgments

- NAMD development is funded by NIH P41-GM104601

- NAMD team:

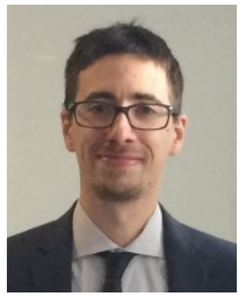

David Hardy


Julio Maia


John Stone


Jim Phillips


Mohammad Soroush Barhaghi


Mariano Spivak


Wei Jiang


Rafael Bernardi


Ronak Buch


Jaemin Choi