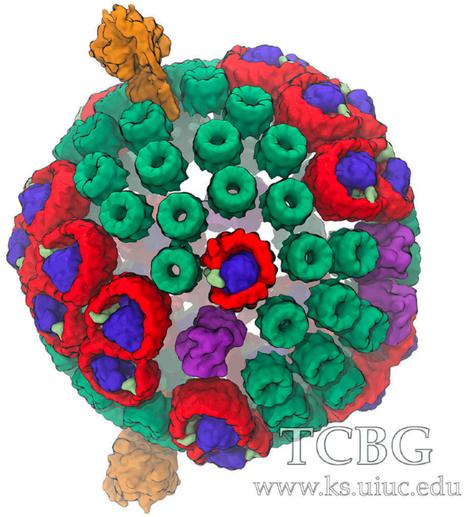
NAMD
Scalable Molecular Dynamics



Scaling Molecular Dynamics Simulations on Aurora with NAMD

David J. Hardy
University Of Illinois at Urbana-Champaign

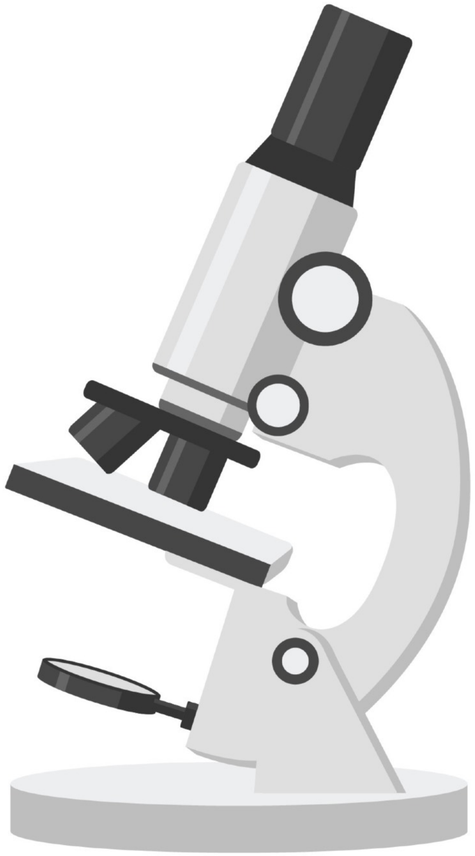
Eric J. Bohm
University Of Illinois at Urbana-Champaign

Ke Yue
Intel Corporation

Wei Jiang
Argonne National Laboratory

NIH National Institutes of Health

A Computational Microscope



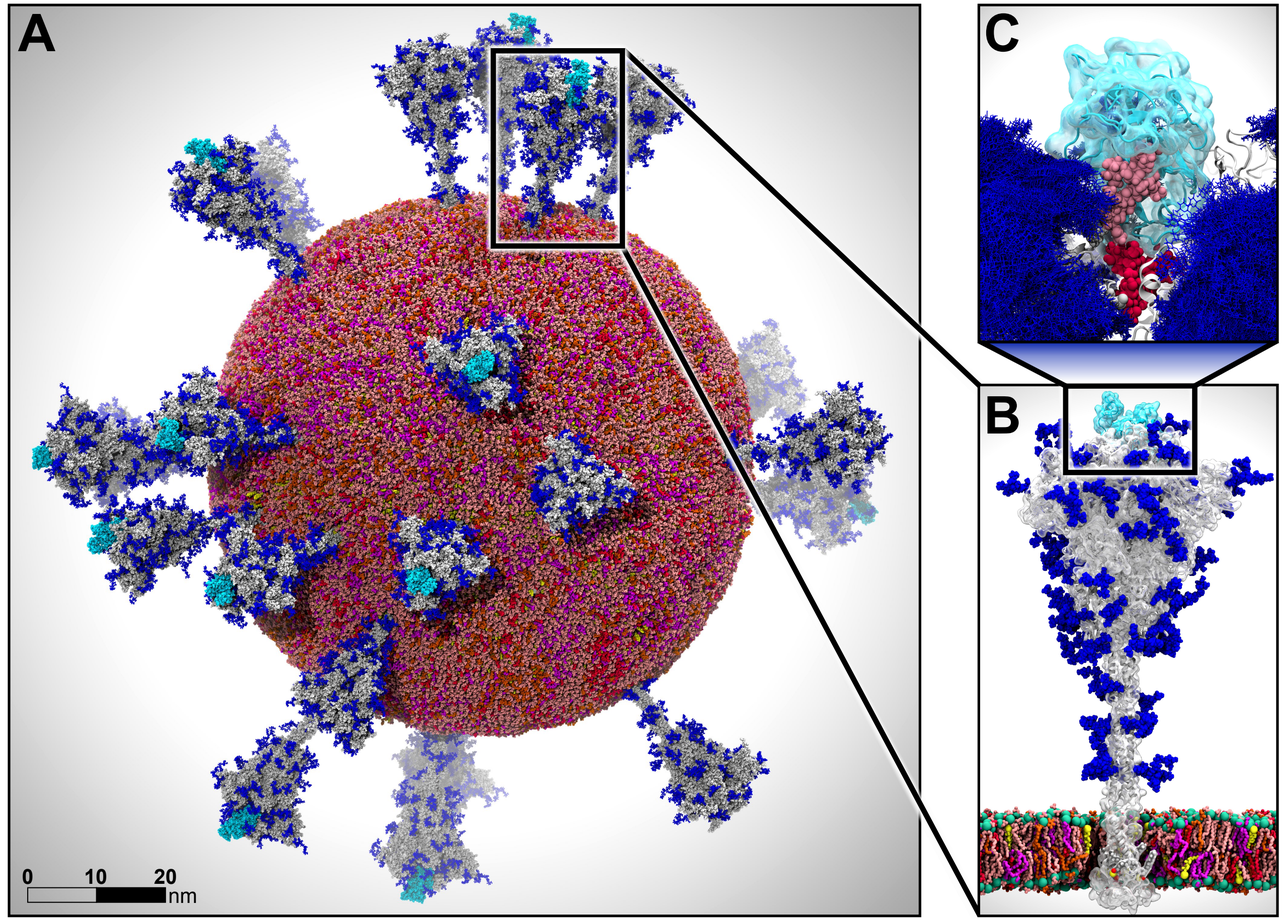
- Molecular dynamics provides biomedical researchers with a theoretical "computational microscope"
- Enables seeing dynamical atomistic details of biomolecules on time and length scales that are not accessible through experimental techniques alone
- Provides scientific insight that can ultimately lead to improvements in human health

NAMD Simulating SARS-CoV-2 on TACC Frontera and OLCF Summit

Collaboration with Amaro Lab at UCSD, images rendered by VMD

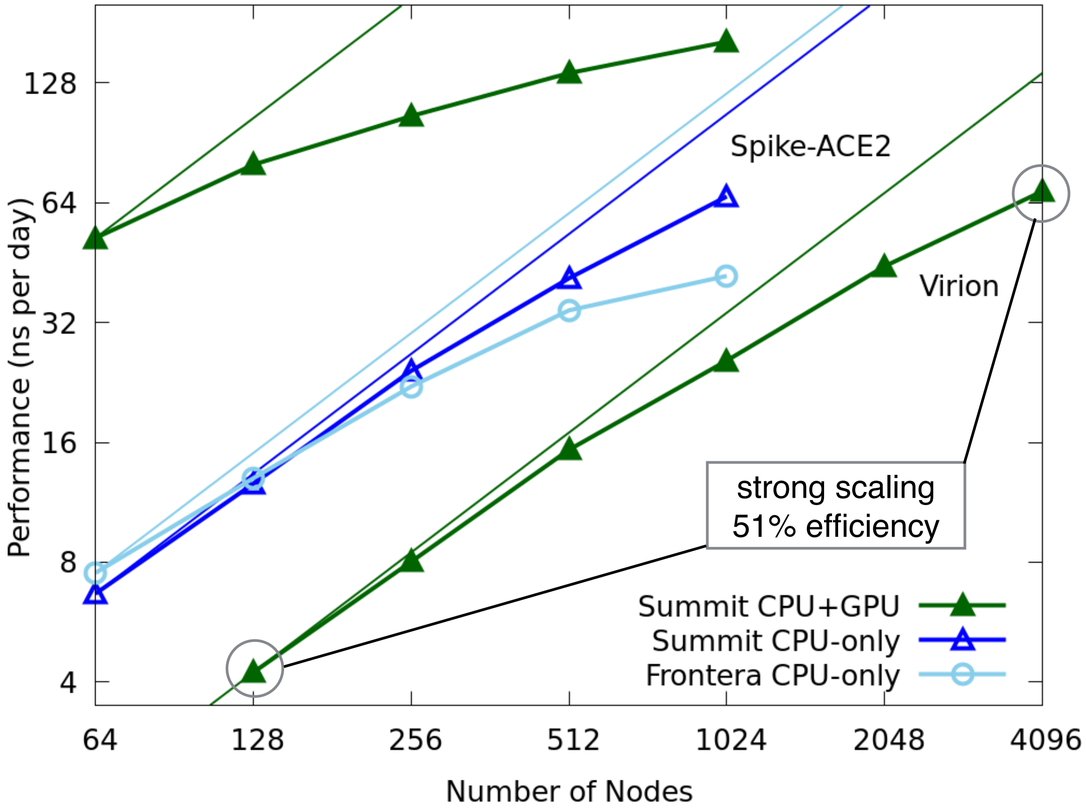
Winner of Gordon Bell Special Prize at SC20, project involved overall 1.13 Zettaflops of NAMD simulation

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



Scaling performance:

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



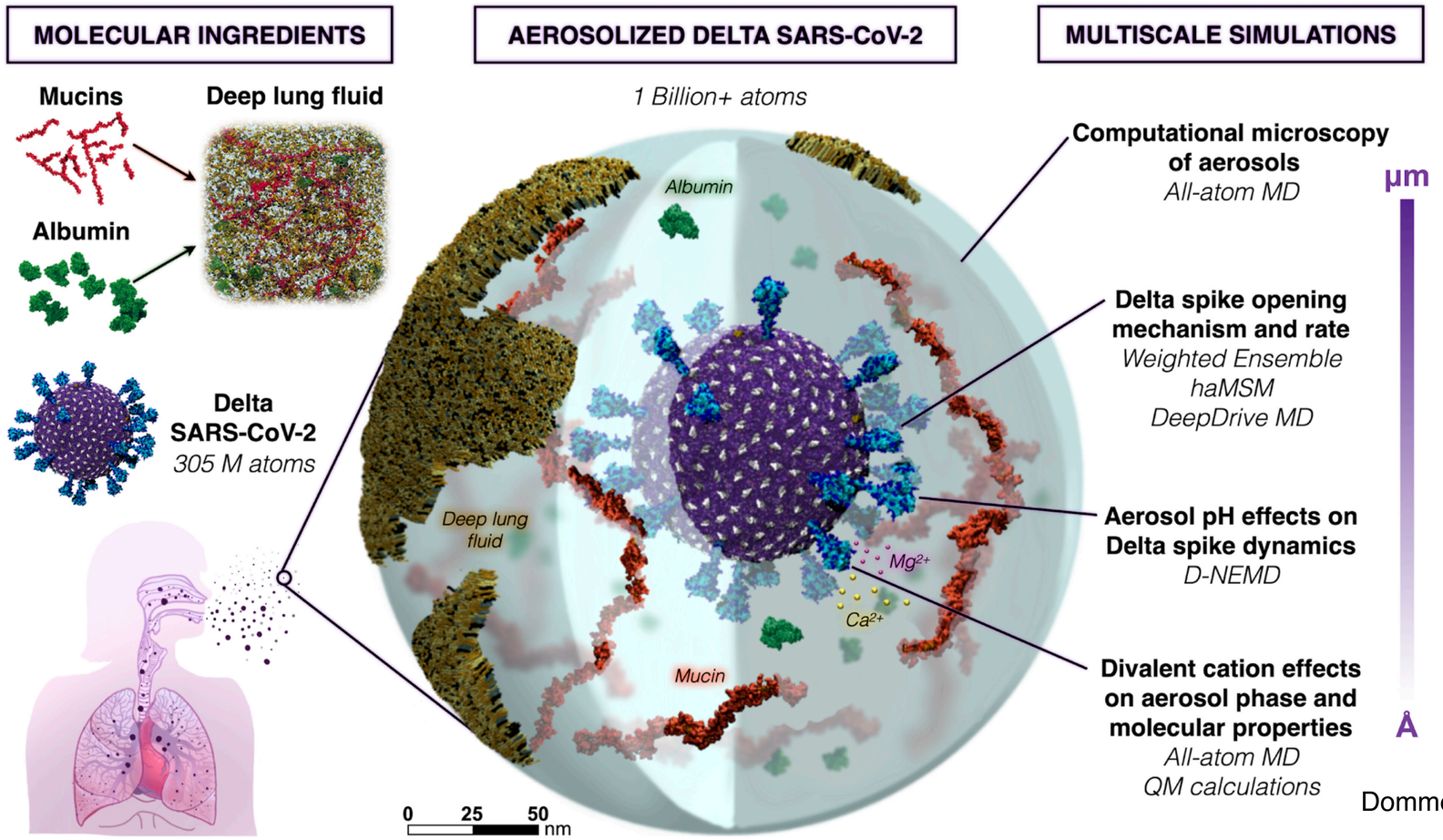
Casalino, et al. *bioRxiv* (2020) https://doi.org/10.1101/2020.11.19.390187

NIH National Institutes of Health

NAMD Simulating Aerosolized SARS-CoV-2 on OLCF Summit

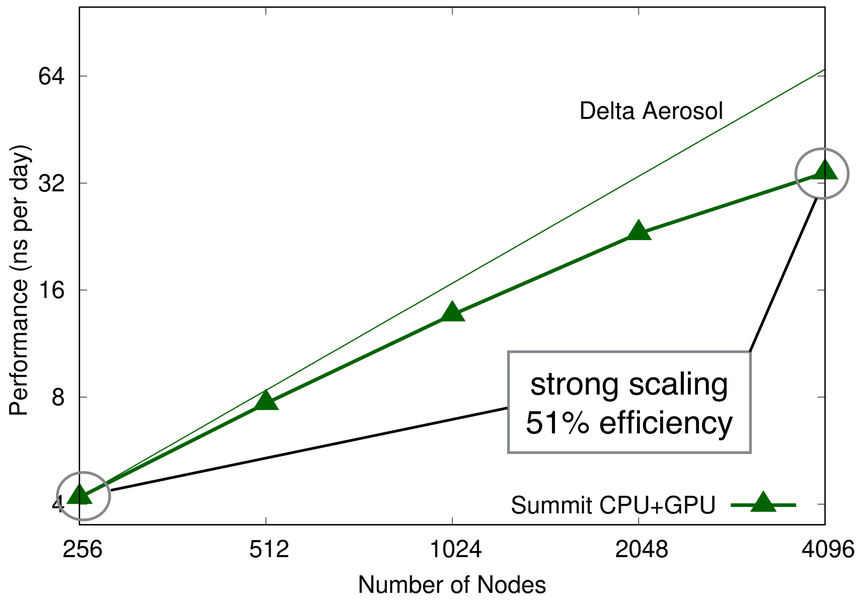
Collaboration with Amaro Lab at UCSD, images rendered by VMD

Finalist for Gordon Bell Special Prize at SC21



Scaling performance:

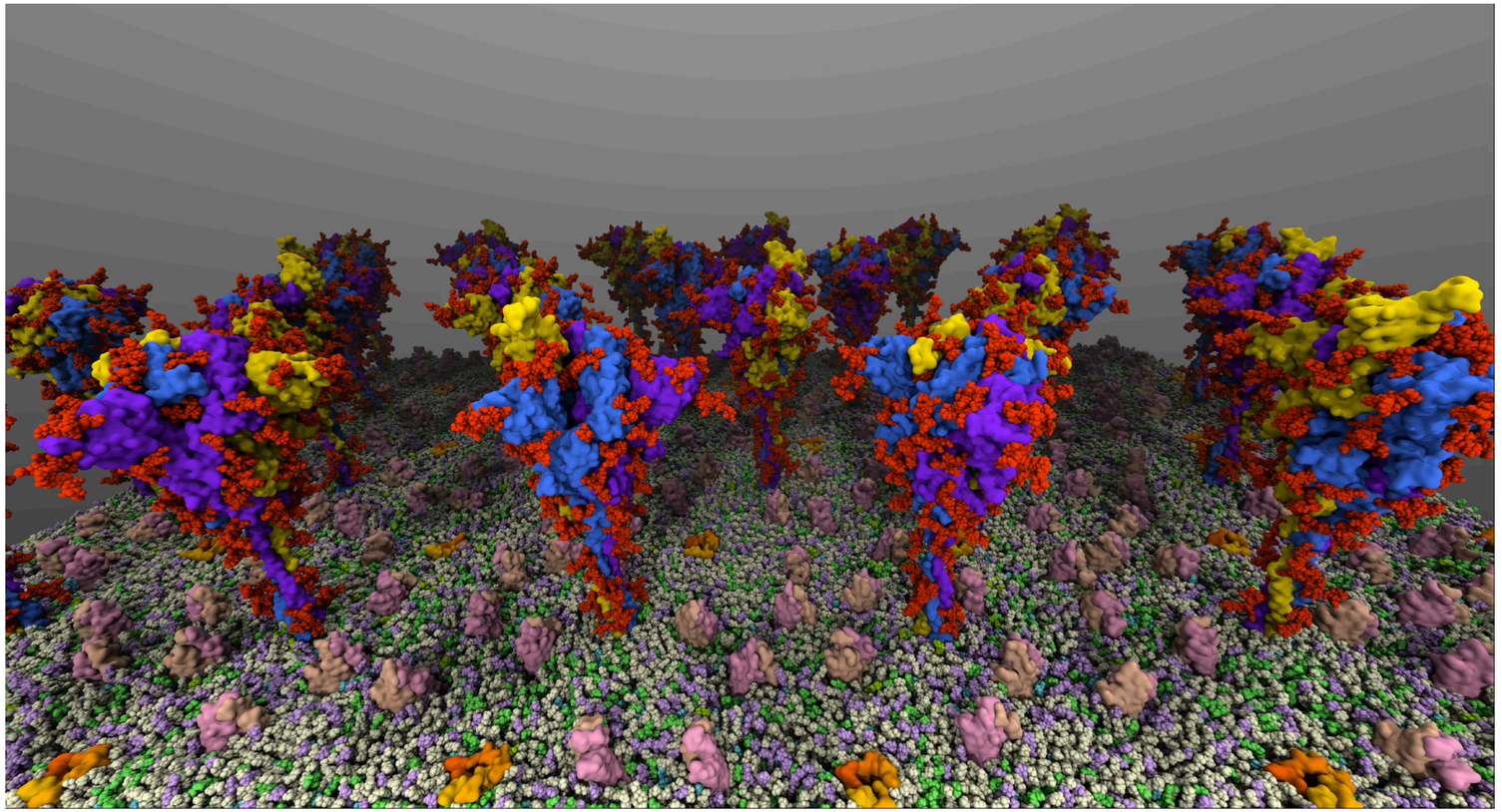
- ~1B atom aerosol



Dommer, et al. *bioRxiv* (2021) https://doi.org/10.1101/2021.11.12.468428

NAMD: Scalable Molecular Dynamics

- Developed and maintained by the Theoretical and Computational Biophysics Group at UIUC
- Code written in C++ with Charm++ parallel objects
  - CUDA for NVIDIA devices
  - HIP (via Hipify) for AMD devices
  - **oneAPI SYCL for Intel devices**
- Simulate movements of biomolecules over time
- Enable parallel scaling
  - **Large systems (single-copy scaling)**
  - **Enhanced sampling (multi-copy scaling)**
- Over 45,000 registered users, over 21,000 citations



Investigations of coronavirus (SARS-CoV-2) spike dynamics.
Credit: Tianle Chen, Karanpal Kapoor, Emad Tajkhorshid (UIUC).
Simulations with NAMD, movie created with VMD.

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

Phillips, et al. *J. Comput. Chem.* 26, 1781-1802 (2005)
Phillips, et al. *J. Chem. Phys.* 153, 044130 (2020)

NIH National Institutes of Health

Molecular Dynamics Simulation

Integrate Newton's equations of motion:

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i = -\vec{\nabla} U(\vec{R})$$

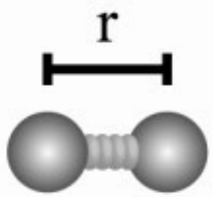
Integrate for millions of time steps

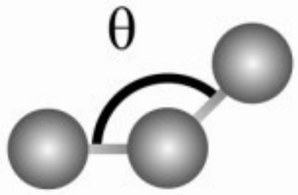
Bonded

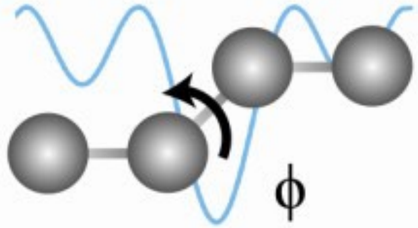
$$E_{total} = \sum_{bonds} K_r(r - r_{eq})^2 + \sum_{angles} K_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)]$$

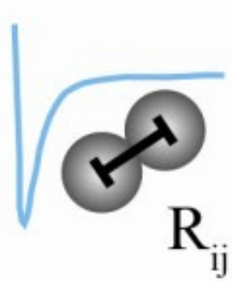
Non-bonded

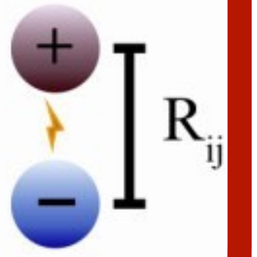
$$+ \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}} \right]$$







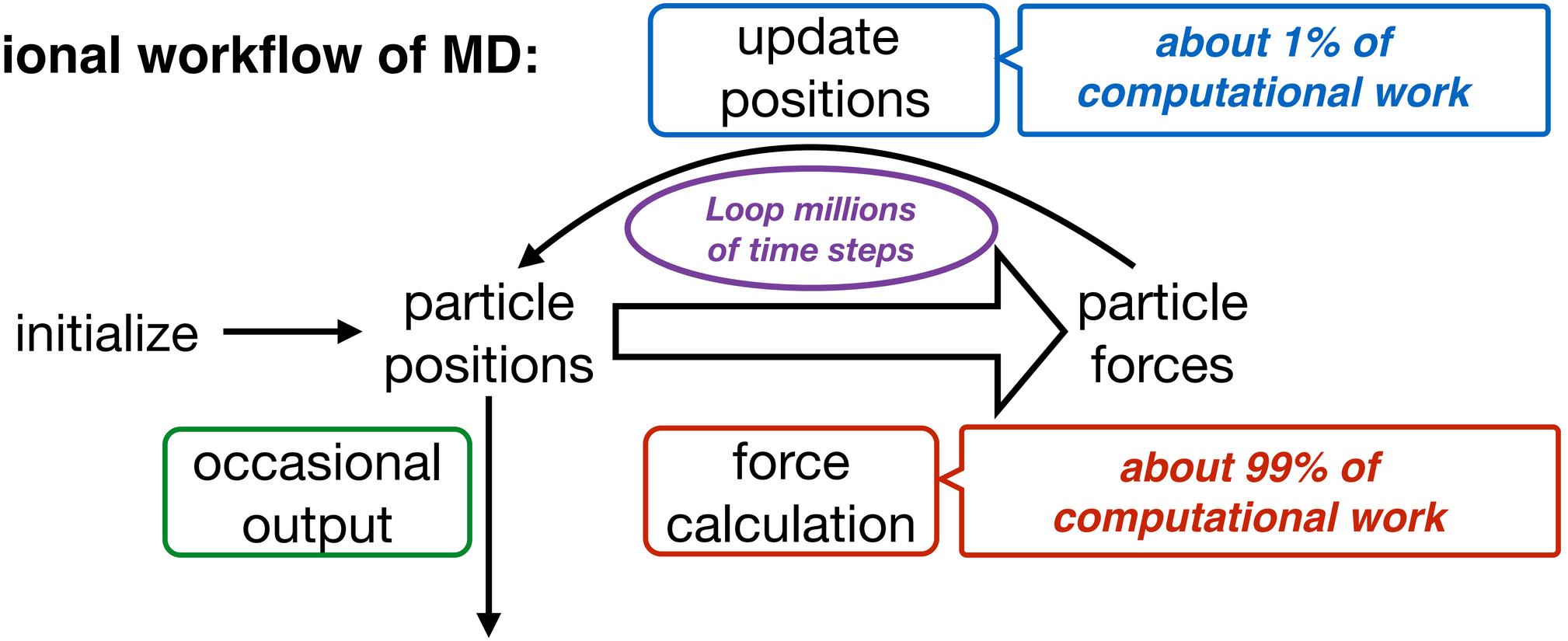




Most computationally intensive part

Parallelism for MD Simulation Limited to Each Time Step

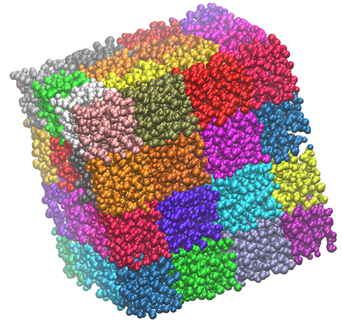
Computational workflow of MD:



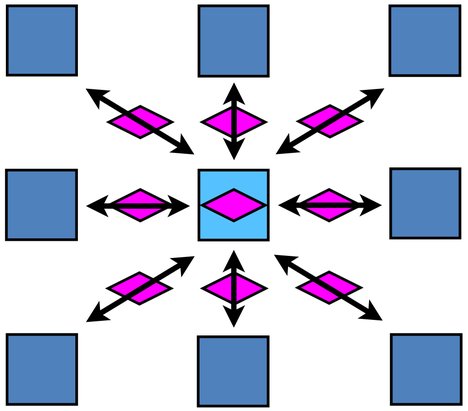
NAMD Parallelizes Domain and Interaction Space

- Decompose atoms into equal volume ***patches***
- Calculate pairwise forces between atoms, treat as interactions between neighboring patches
- Decompose patch-patch interaction ***compute objects***
- Moving atoms: update spatial decomposition by ***migrating atoms*** between adjacent patches
- Load balancing: update work decomposition by ***migrating compute objects*** to keep processors consistently occupied

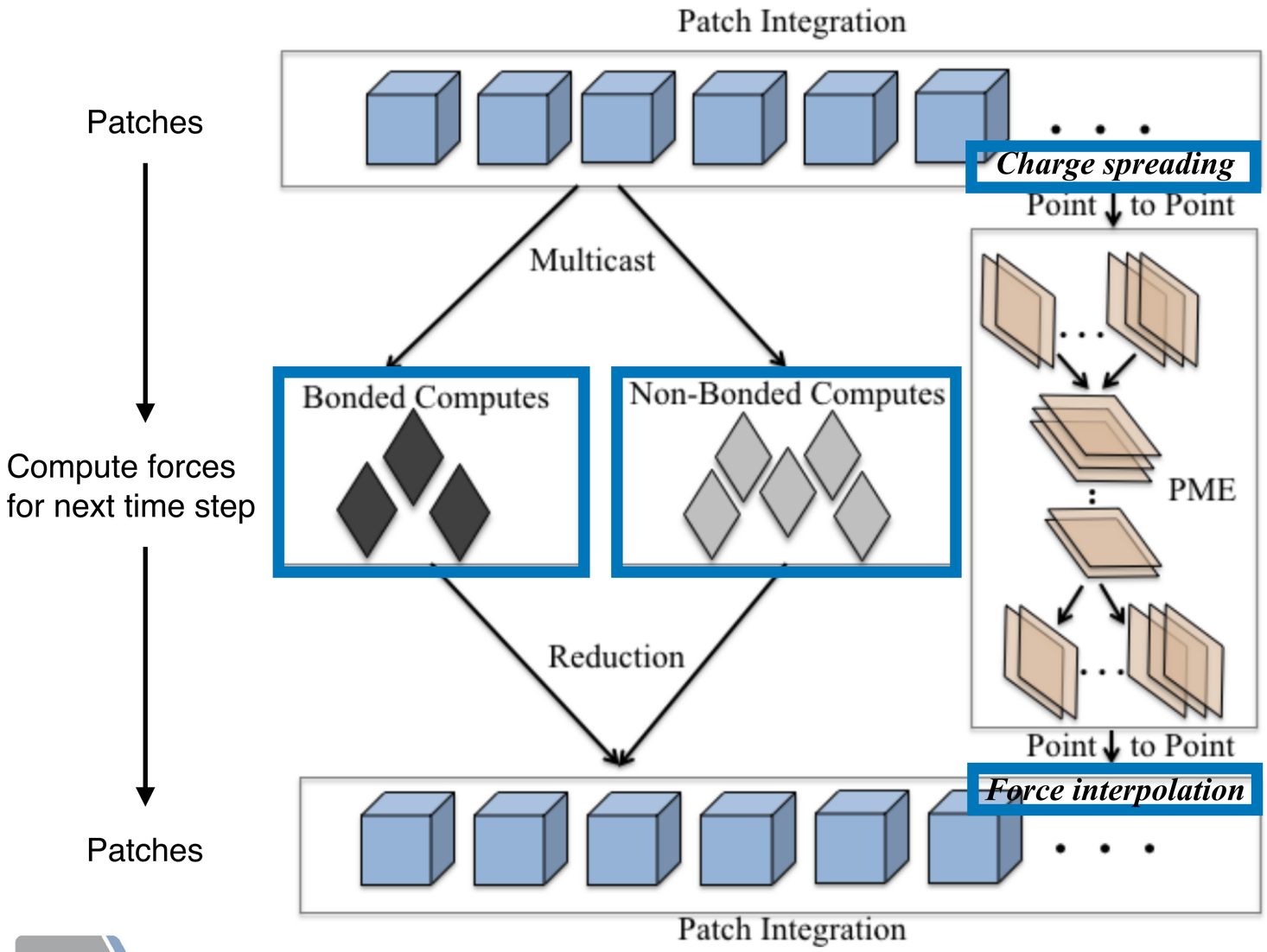
Spatial decomposition of atoms into patches



Work decomposition of patch-patch interactions into migratable compute objects



NAMD Parallel Workflow Incorporates GPUs



Offload force compute to GPU

Must aggregate positions

Years of Intel Support for NAMD Development

- Established TCBG at UIUC as **oneAPI Academic Center of Excellence**
- Developed AVX-512 acceleration (3.5x improvement on TACC Frontera) (Mike Brown)
- Initial porting of CUDA force kernels to oneAPI/SYCL (Tareq Malas)
- Fixing PME SYCL kernel implementation with FFTs using oneMKL (Jaemin Choi)
- **Optimization of SYCL GPU-offload code path (4.5x improvement) (Ke Yue)**

Current NAMD SYCL Developers

Aditya Bhosale (UIUC)

Eric Bohm (UIUC)

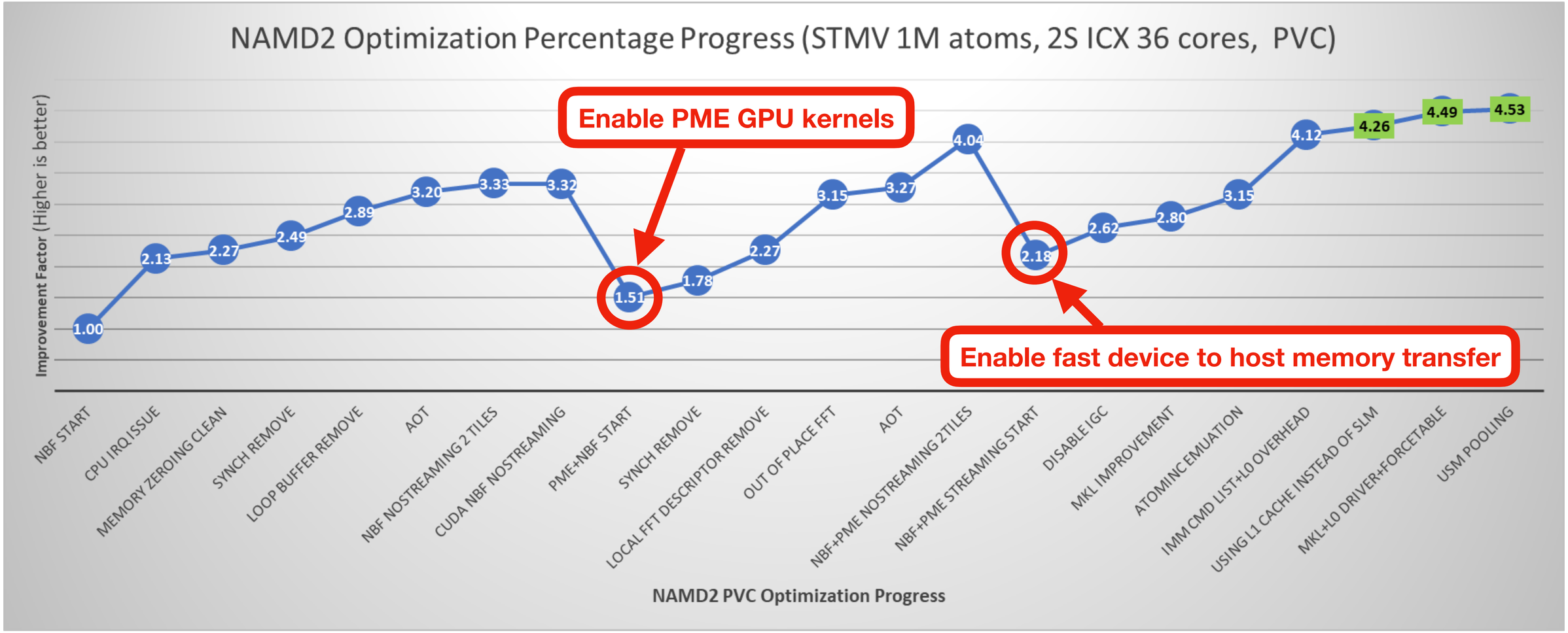
David Hardy (UIUC)

Wei Jiang (ANL)

Ke Yue (Intel)

Optimizing NAMD SYCL Performance on PVC

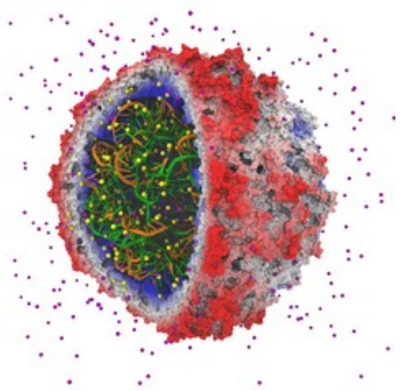
Ke Yue (Intel)



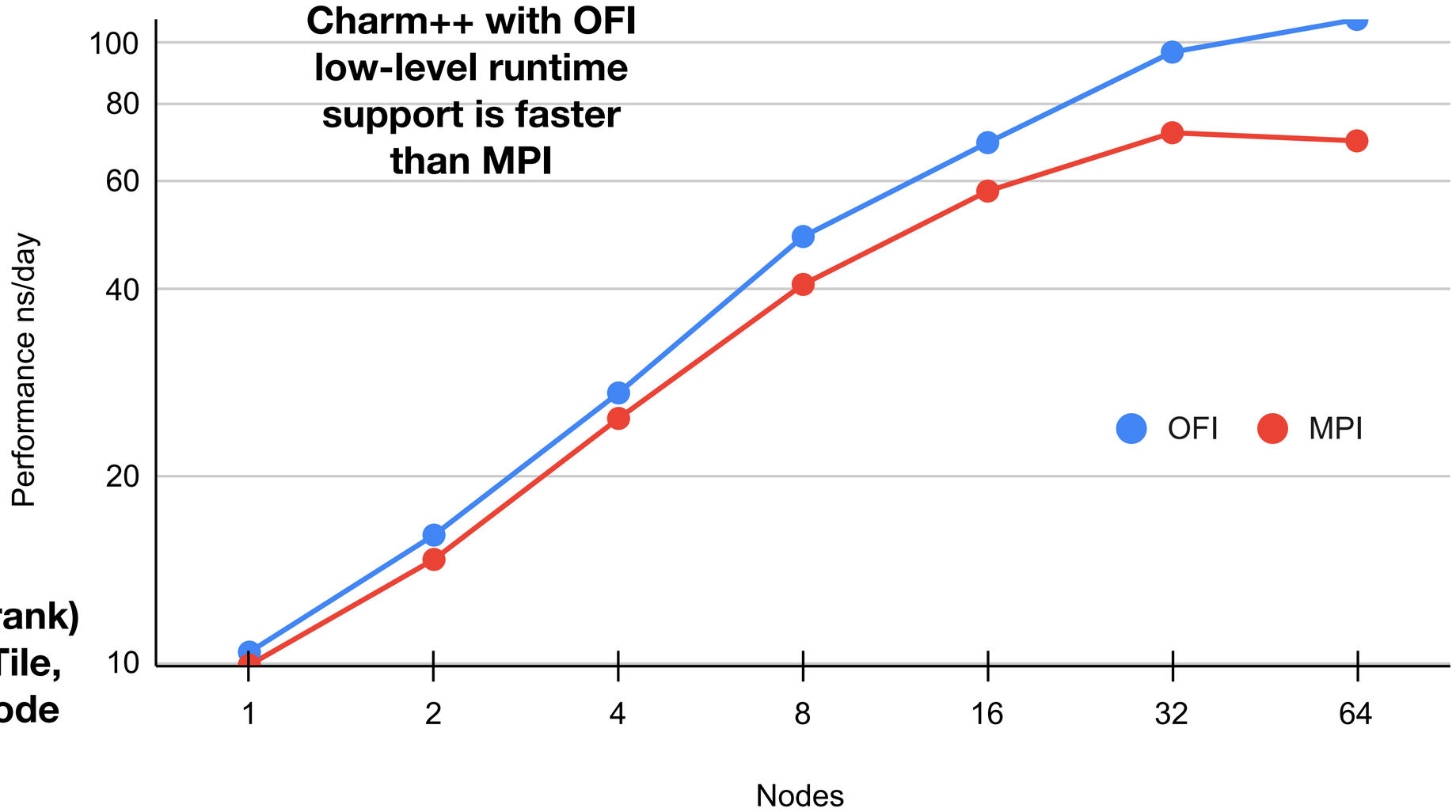
NAMD SYCL Scaling on Sunspot

Eric Bohm (UIUC)

STMV (Satellite Tobacco Mosaic Virus) 1.06M atoms



Running 1 process (rank) per GPU Compute Tile, 12 processes per node

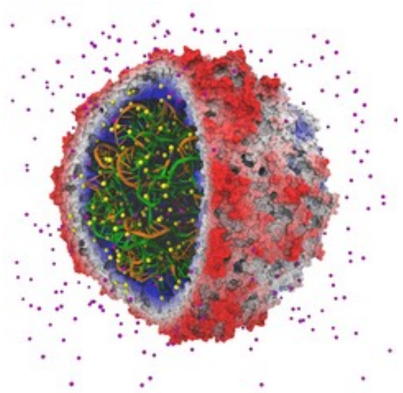


NIH National Institutes of Health

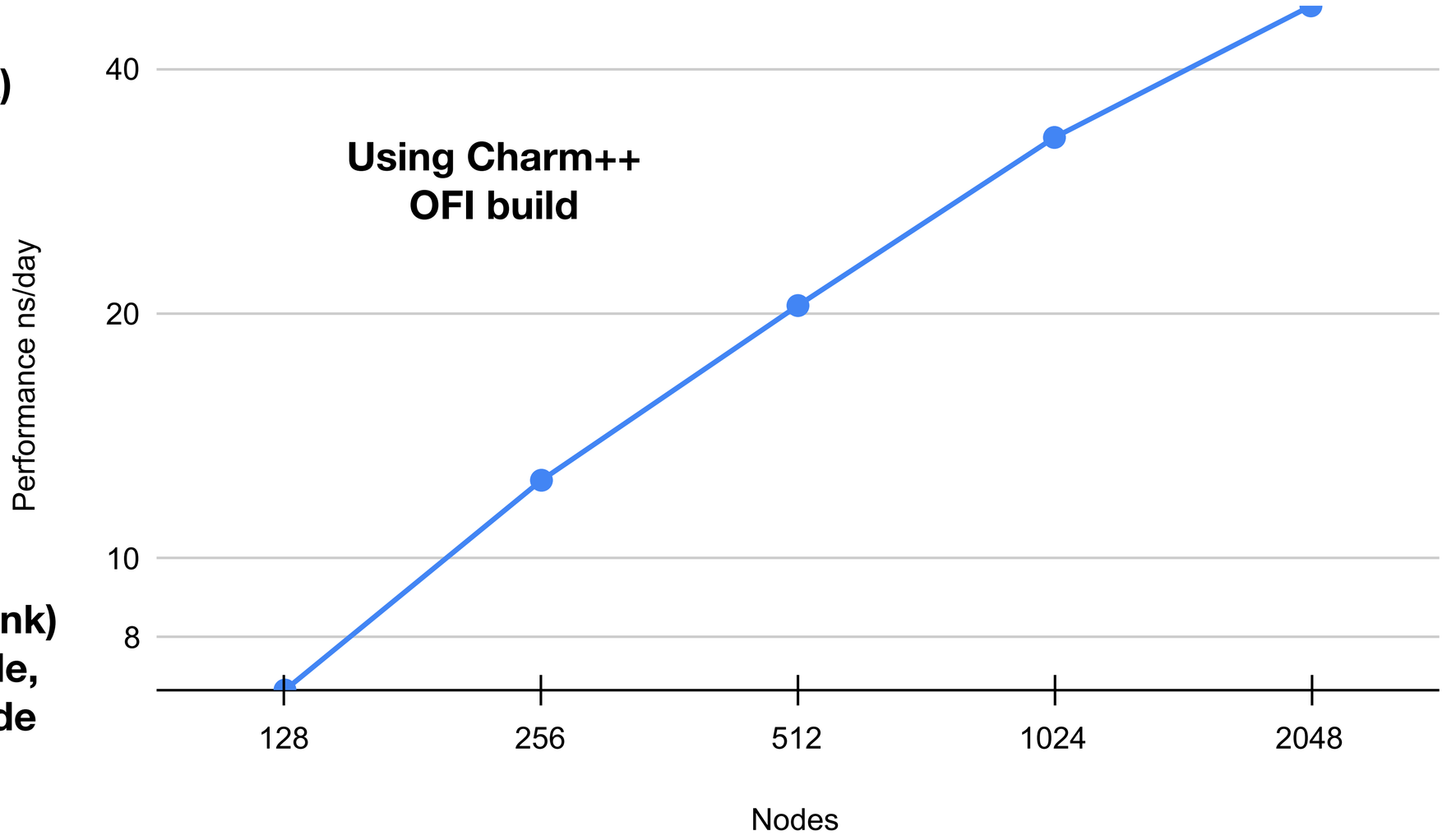
NAMD SYCL Strong Scaling on Aurora

Wei Jiang (ANL)

5x5x4 STMV matrix (synthetic benchmark) ~107M atoms



Running 1 process (rank) per GPU Compute Tile, 12 processes per node

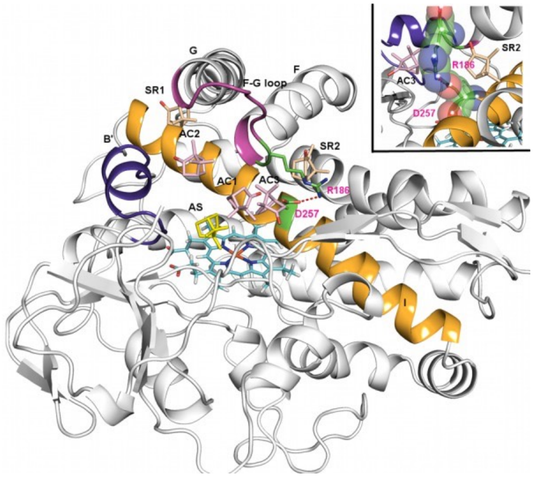


NIH National Institutes of Health

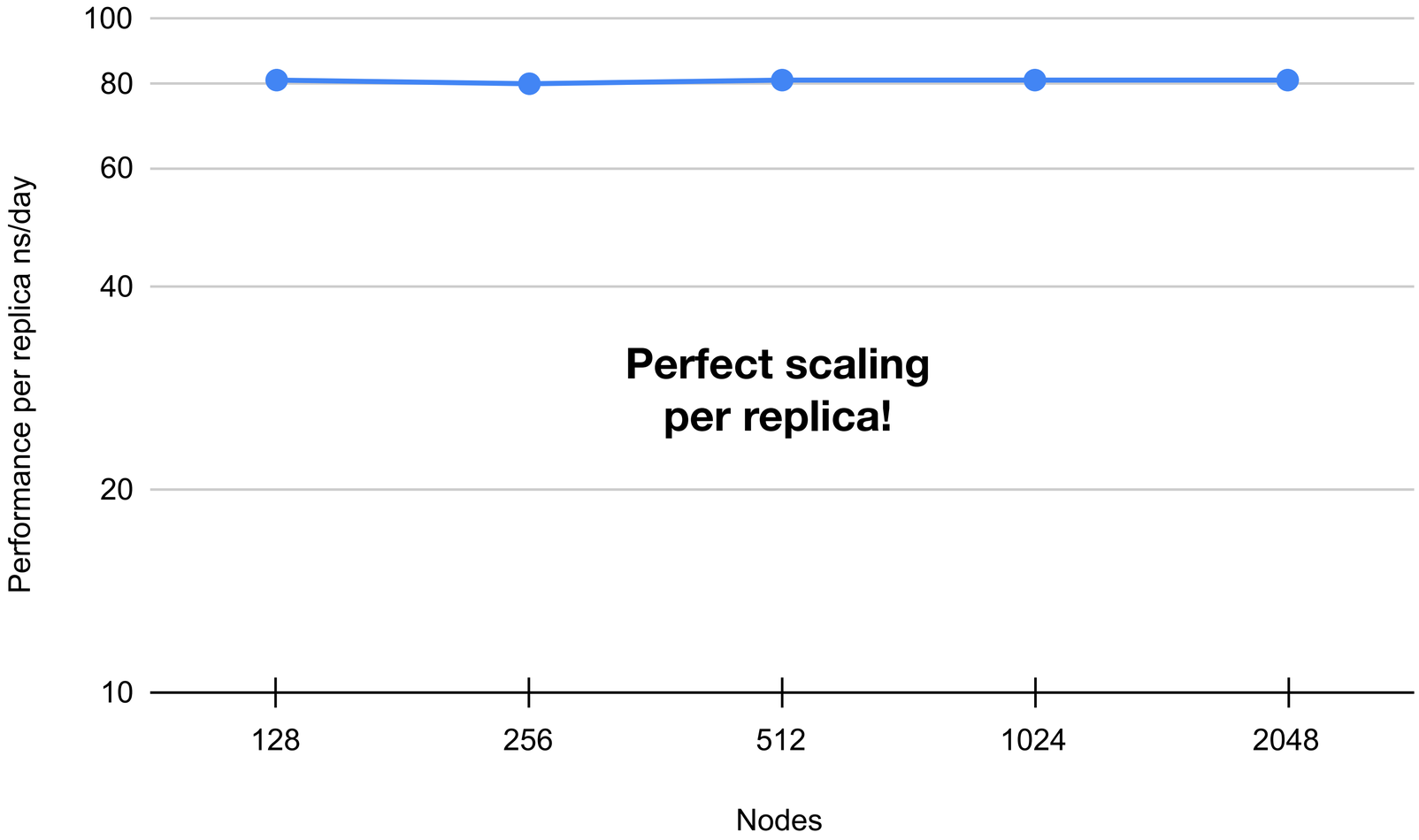
NAMD SYCL Weak Scaling on Aurora

Wei Jiang (ANL)

Camphor/P450 complex ~70k atoms with Replica-Exchange Solute Scaling (REST2)



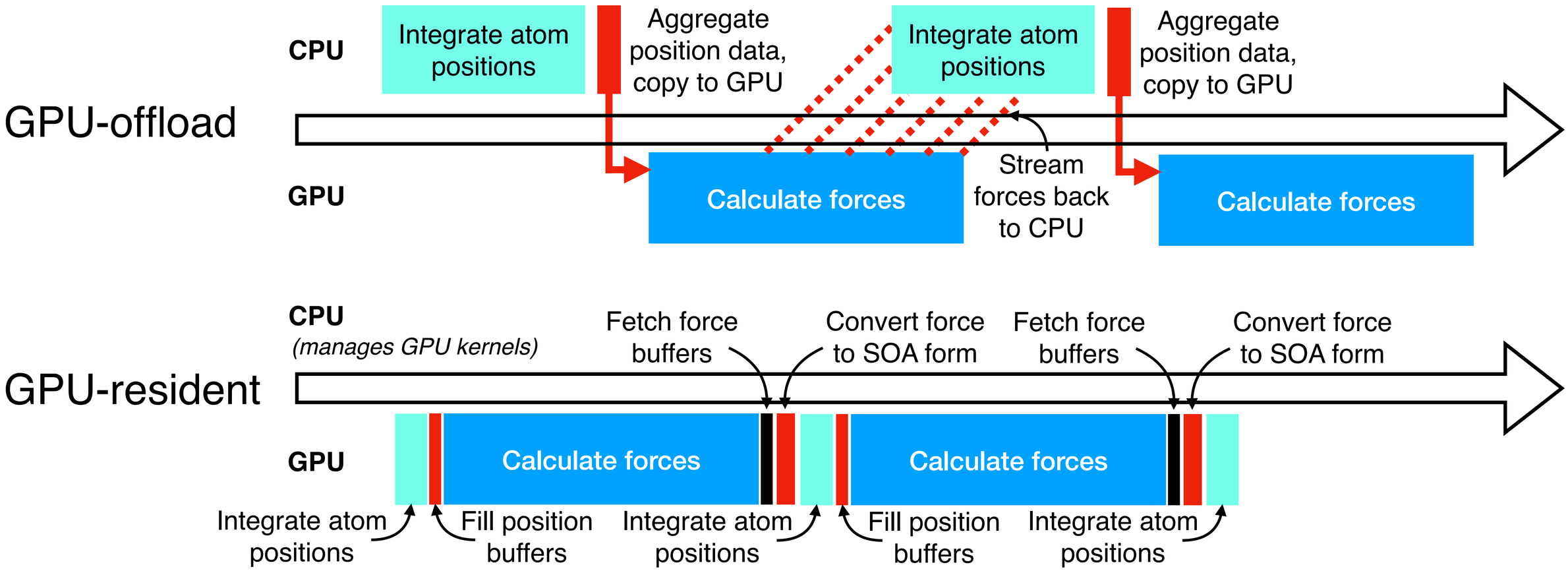
Running 2 GPUs per replica (3 replicas per node) single neighbor exchange every 500 steps



National Institutes of Health

Work in Progress: GPU-resident approach for SYCL

Move integrator to GPU and maintain data between time steps



Expect at least 2.5x performance improvement for SYCL!

Porting to SYCL: Sometimes it's easy to get tripped up by the little things!

Question: Why are we getting the wrong answer when we enable PME?

Answer:

src/SyclRecord.h

```
...  ...   @@ -51,7 +51,8 @@ struct SyclAtom {
 51   51   };
 52   52
 53   53   struct SyclForce {
 54        -   float x, y, z;
      54   +   // pad needed to match sizeof(sycl::float3) which is 16 bytes
      55   +   float x, y, z, pad;
 55   56   };
```

Acknowledgments

- Theoretical and Computational Biophysics Group, The Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign
- National Institutes of Health (NIH) grants P41-GM104601 and R24-GM145965
- Argonne Leadership Computing Facility (ALCF), Aurora Early Science Project Award
- Texas Advanced Computing Center (TACC)
- Intel Corporation

NIH National Institutes of Health