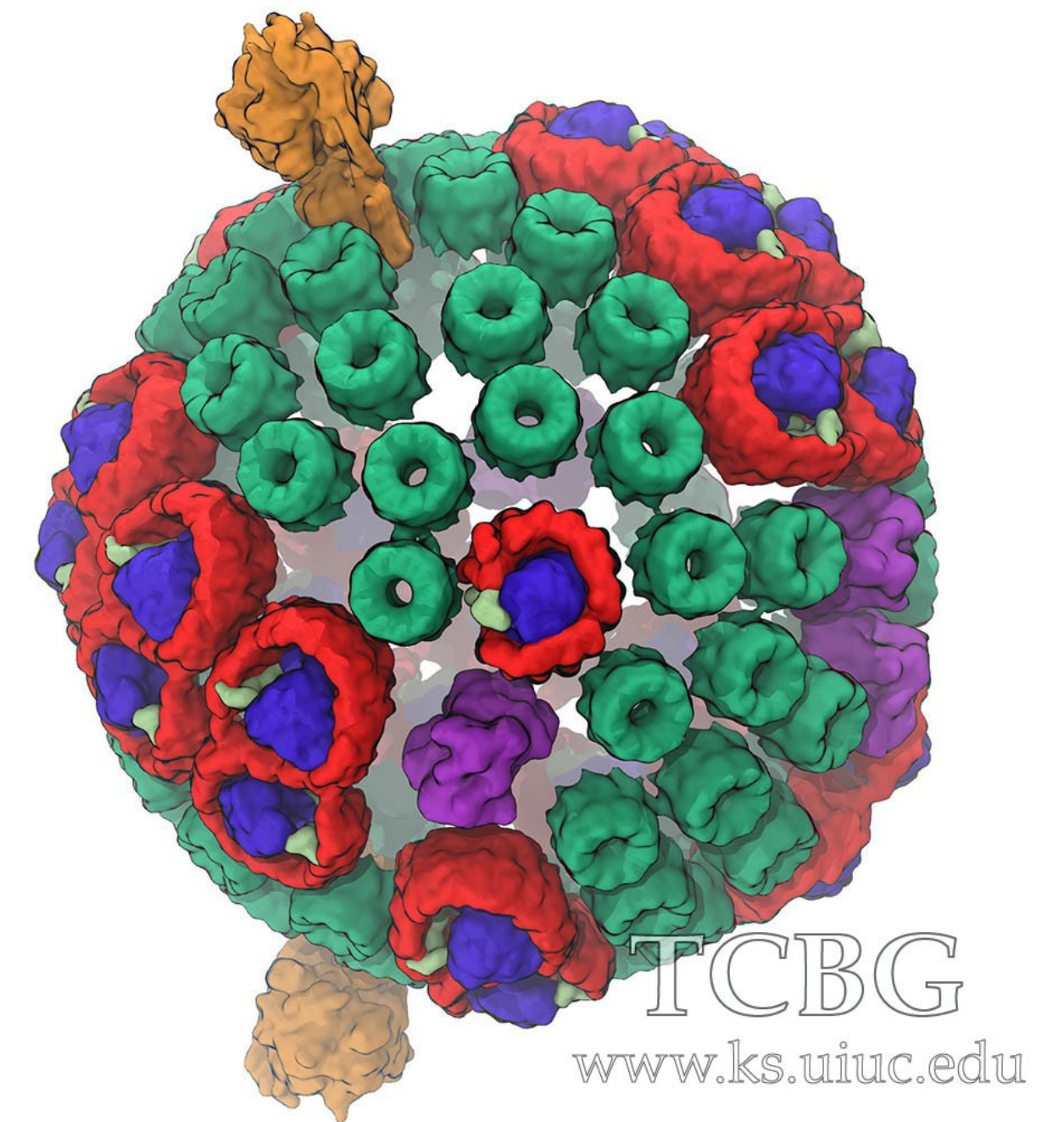




Analyzing NAMD Performance on TACC Vista

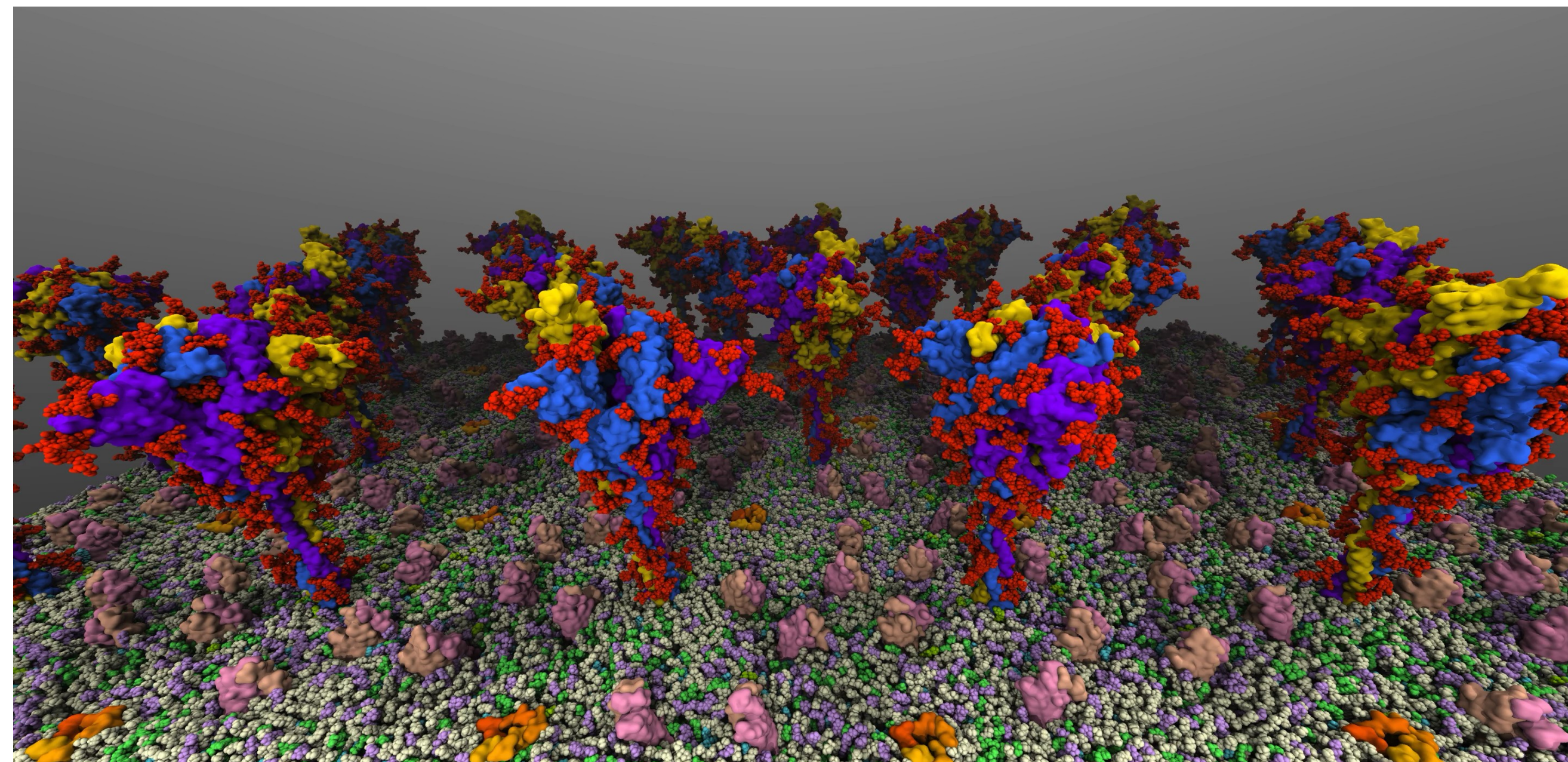
Eric Bohm , David J. Hardy, Hanning Chen, and Amit Ruhela
UIUC TACC

NAMD: Scalable Molecular Dynamics

- Developed and maintained by the Theoretical and Computational Biophysics Group at UIUC

Funded by NIH grants P41-GM104601 and R24-GM145965

- Code written in C++ with Charm++ parallel objects
 - CUDA for NVIDIA devices
 - HIP (via Hipify) for AMD devices
 - oneAPI SYCL for Intel devices
- Biomolecule simulations using CHARMM or AMBER force fields
- Enable parallel scaling
 - **Large systems (single-copy scaling)**
 - **Enhanced sampling (multi-copy scaling)**



<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)

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

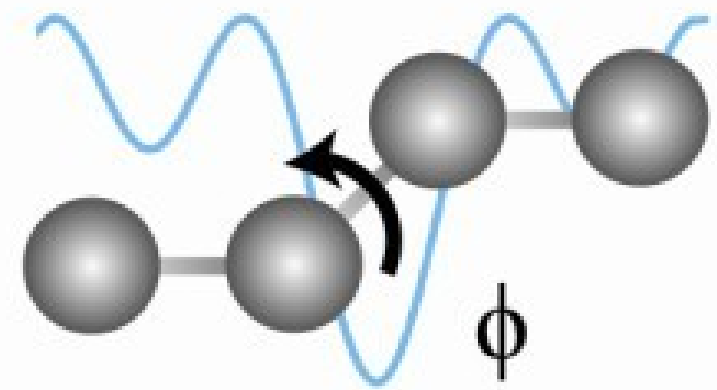
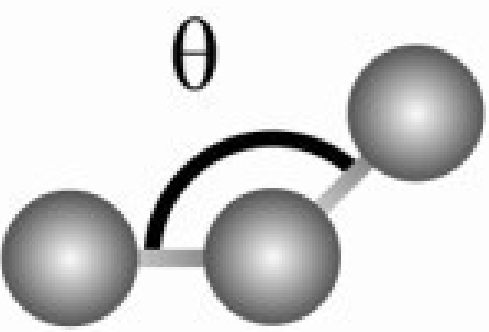
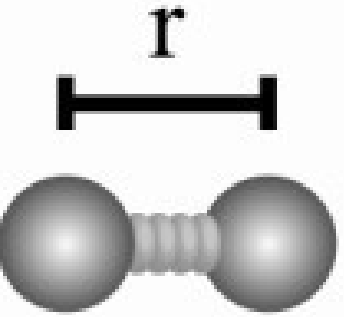
Performing Molecular Dynamics (MD) Simulation

Integrate Newton's equations of motion:

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i = -\vec{\nabla} U(\vec{R}) \quad \leftarrow \text{Integrate for millions of time steps}$$

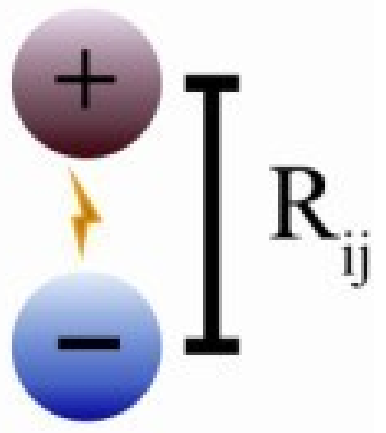
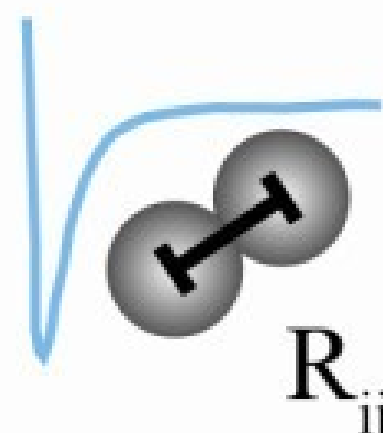
Bonded

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



Non-bonded

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



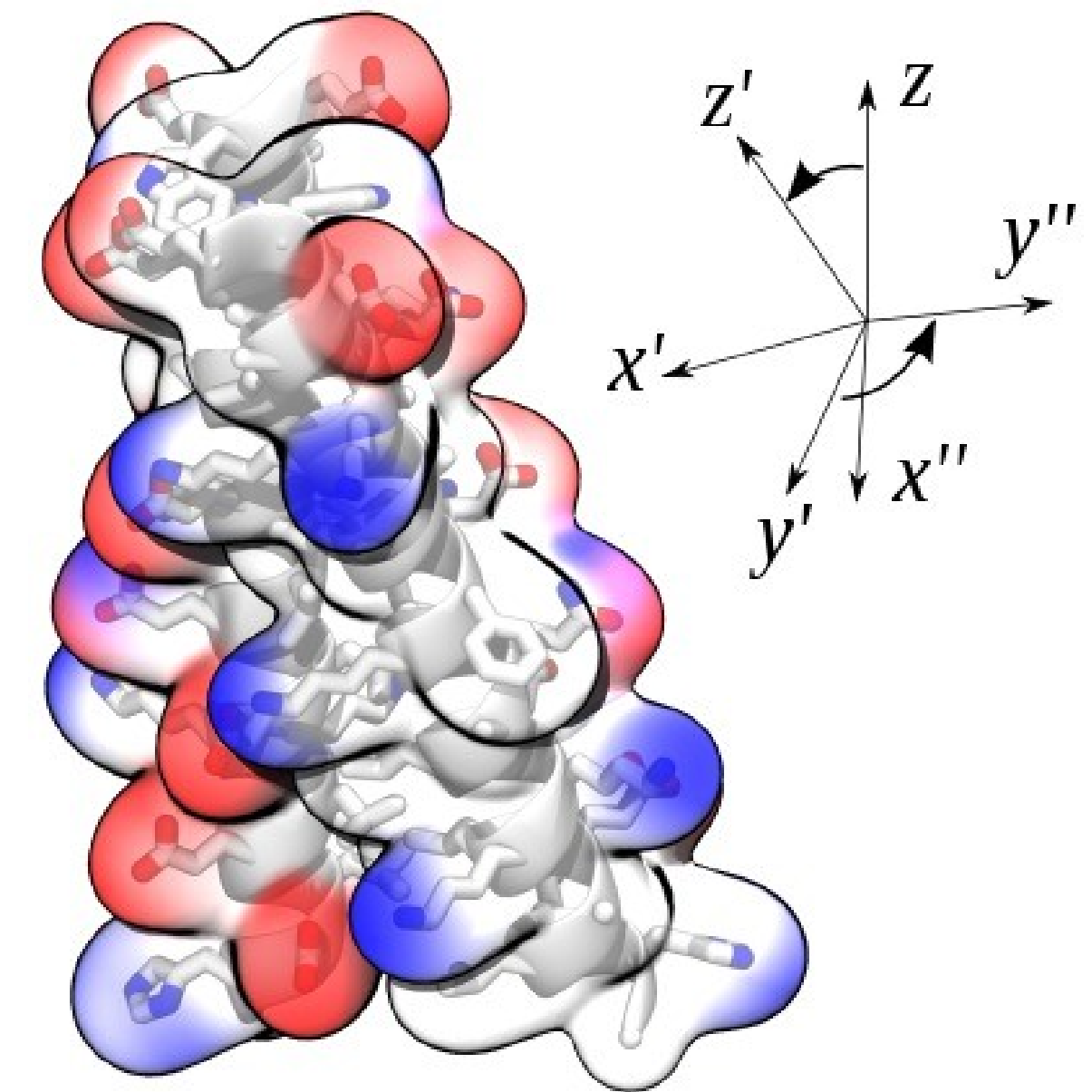
Most computationally intensive part →

- Use cutoff distance for non-bonded interactions O(N)
- Fast O(N log N) approximation for long-range electrostatics (particle mesh Ewald)
- Employ multiple time stepping techniques to avoid calculating PME every step
- Constrain bonds to hydrogen atoms and water angles to permit use of longer time steps

Overcoming Timescale Limitations with Non-equilibrium MD

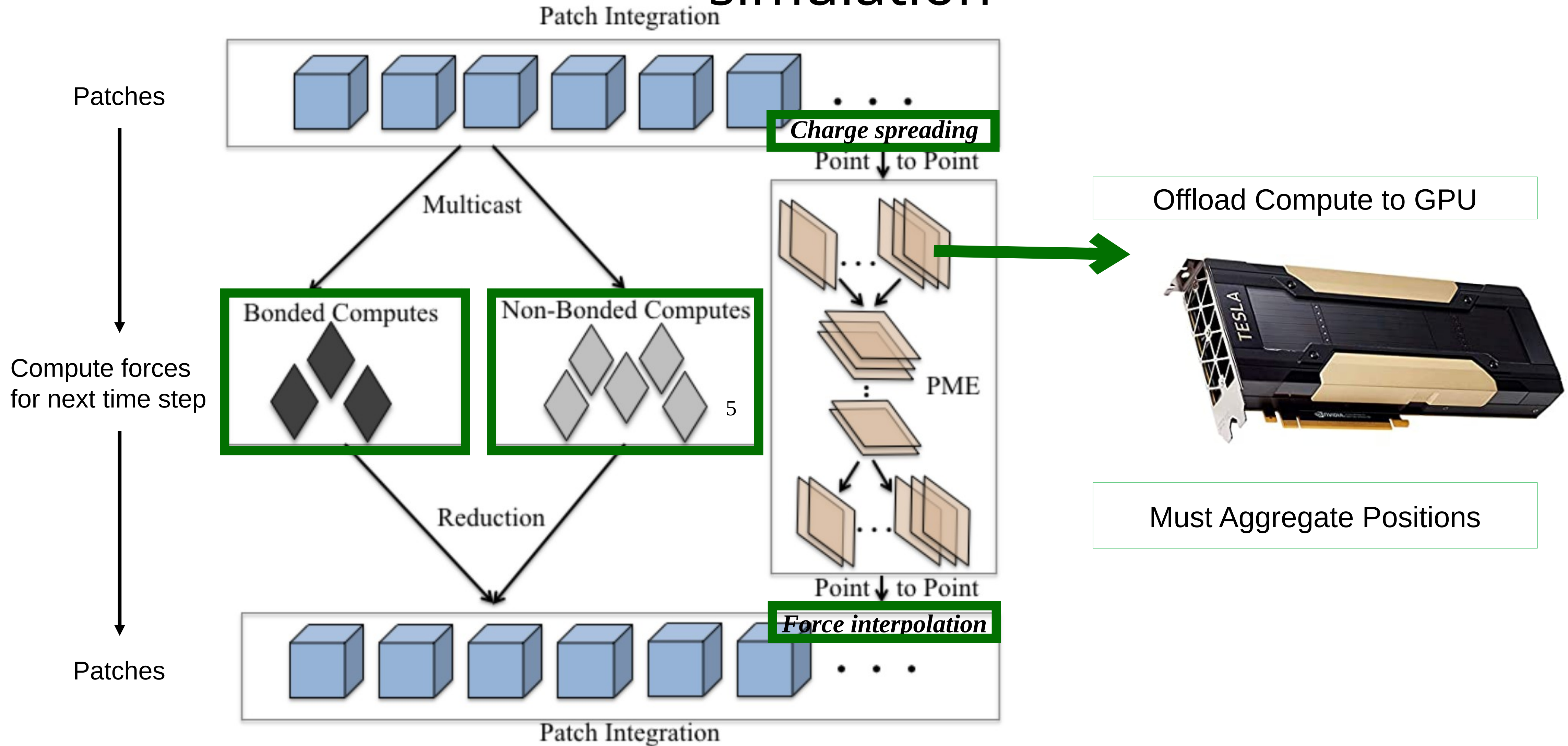
Methods to enhance sampling or extend modeling:

- User-defined forces
 - E.g. steered MD, grid forces, Tcl forces
- Collective variables (Colvars) module
 - E.g. apply controlled biasing potential
- Accelerated sampling methods
 - E.g. accelerated MD, solute scaling (REST2), replica-exchange MD
- Alchemical free energy methods
 - Free energy perturbation (FEP) and thermodynamic integration (TI)
- Constant pH simulations
 - Treating variable protonation states
- Hybrid QM/MM simulations
 - Treating chemical reactions or charge redistribution
 - Call external QM code or (very recent) library call to MLFF (machine-learned force field)



Colvars orientation collective variable (CV)

Using GPU-offload approach for multi-node simulation



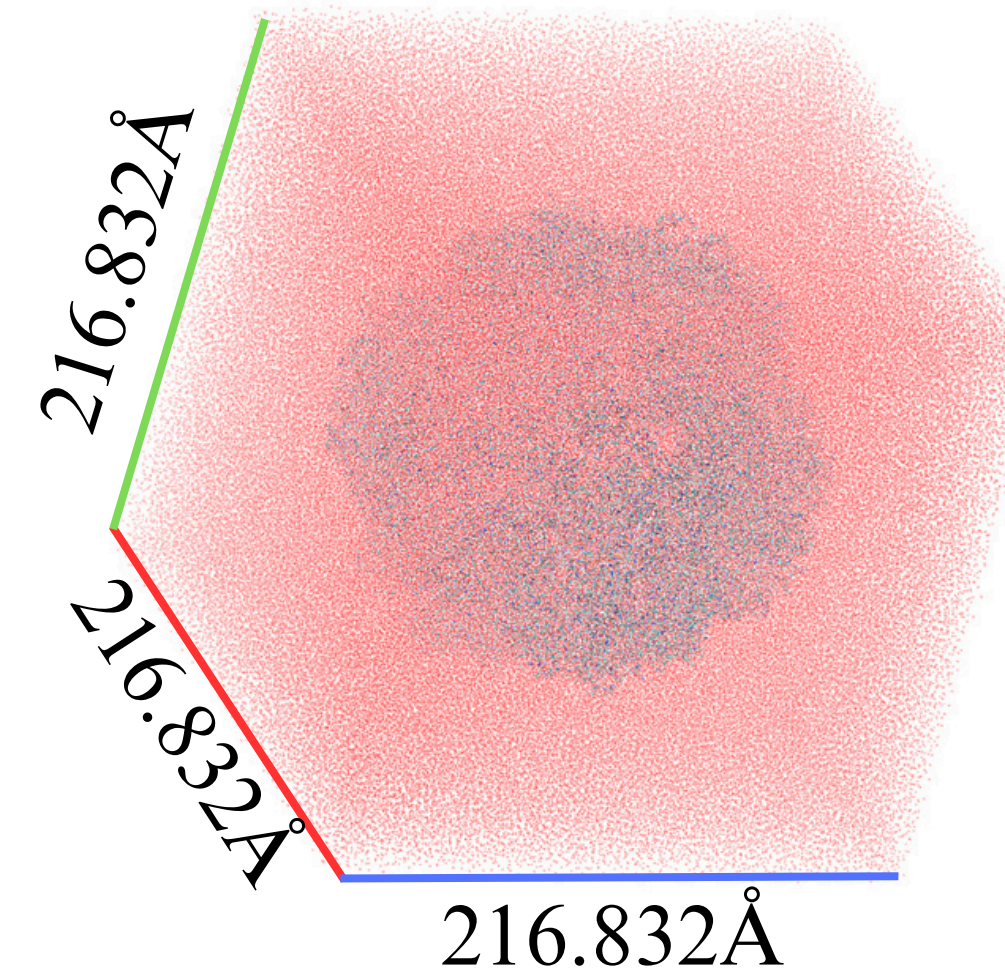
STMV Benchmark

Satellite Tobacco Mosaic Virus

STMV : 1,066,628 atoms

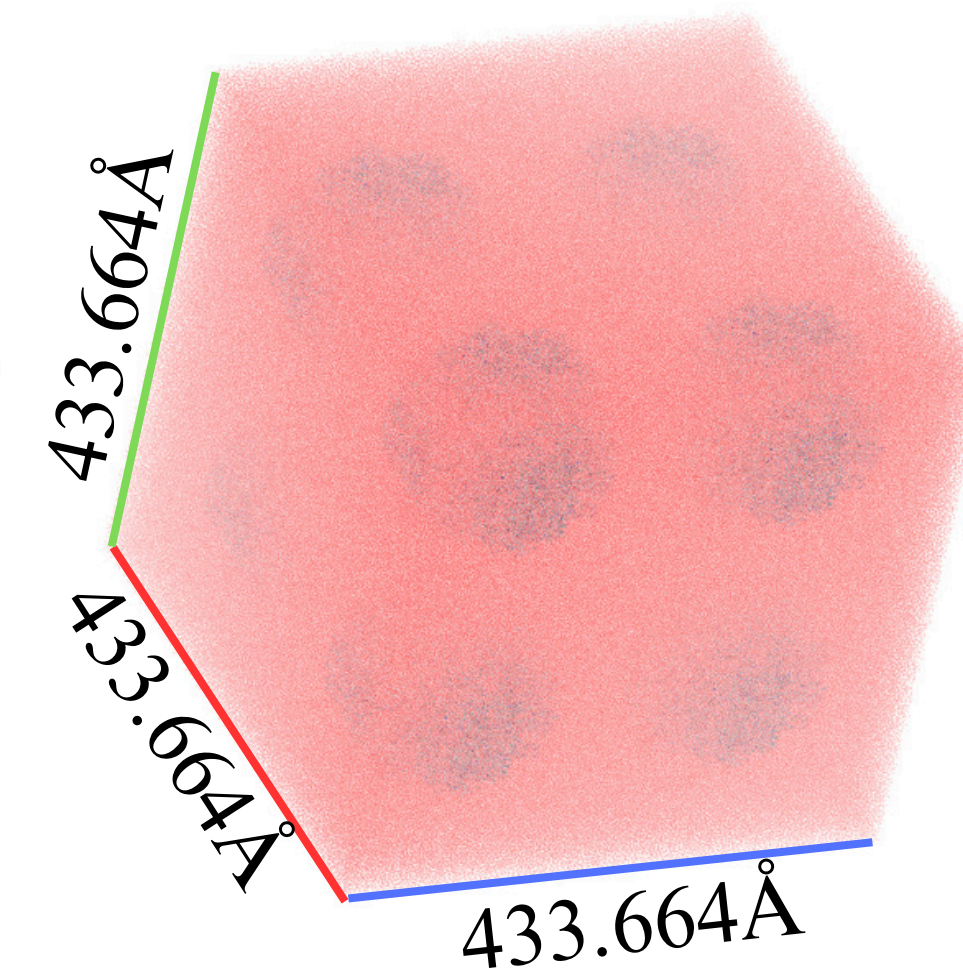
Tiled to 8M, 64M, 1000M
Provides targets for scaling
experiments.

$1 \times 1 \times 1$



(a) STMV_1M

$2 \times 2 \times 2$

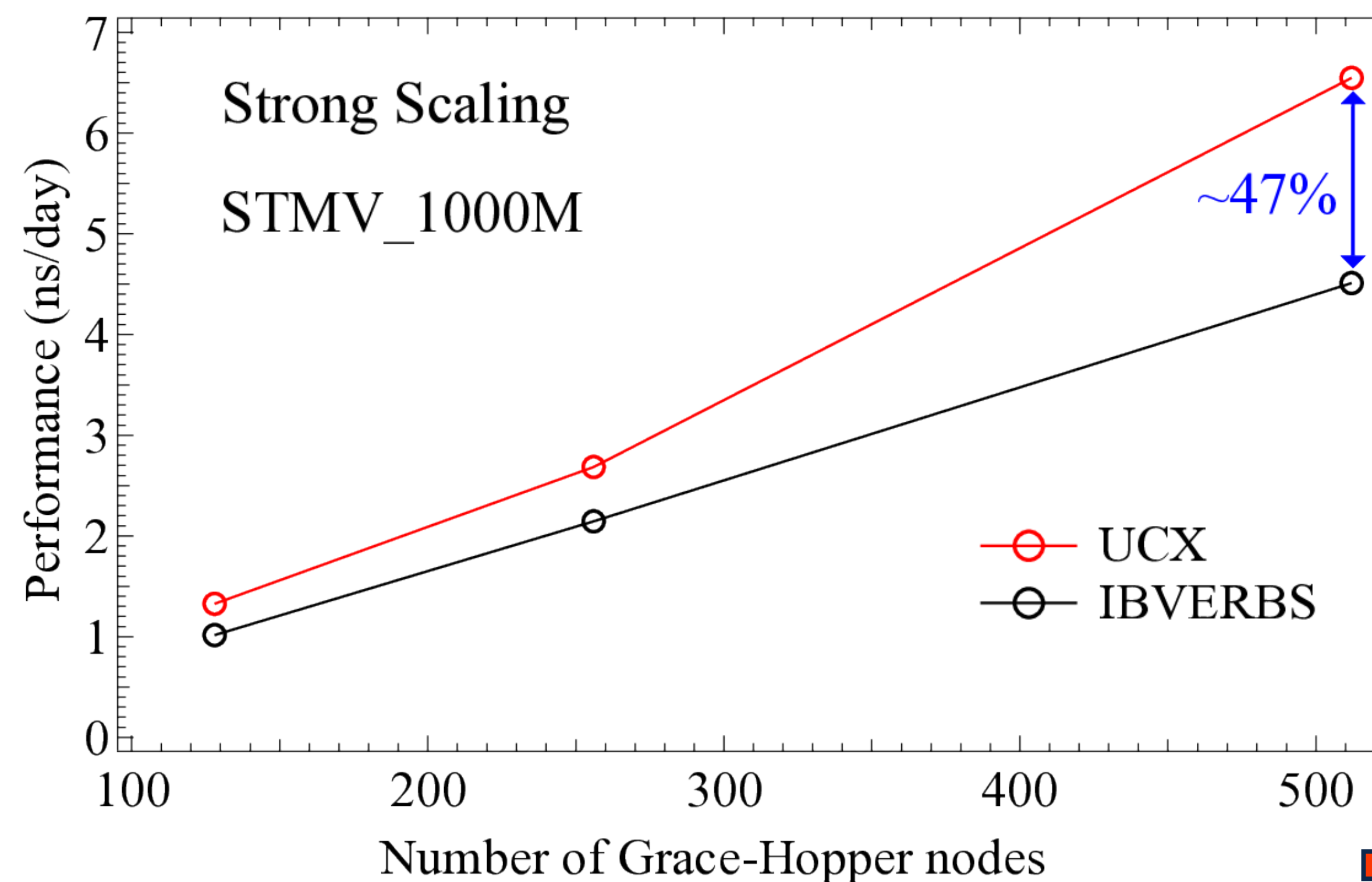
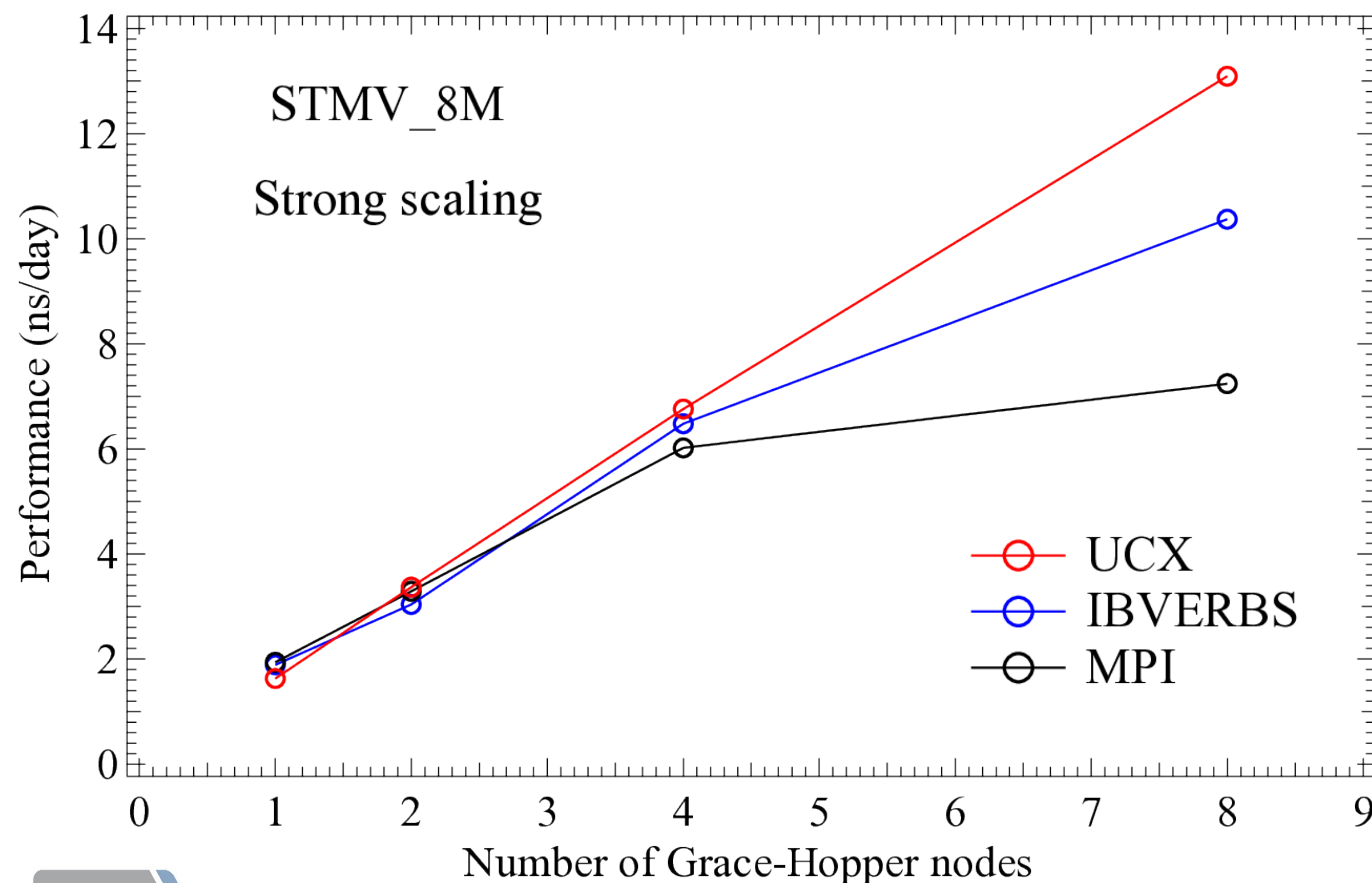


(b) STMV_8M

NAMD Strong Scaling Performance on TACC Vista

TACC Vista: 1 Grace Hopper (GH200) superchip per node
72 ARM64 cores per Grace
NVLink 900 GB/s connection between Grace and Hopper
512 nodes connected by 400 GB NDR InfiniBand

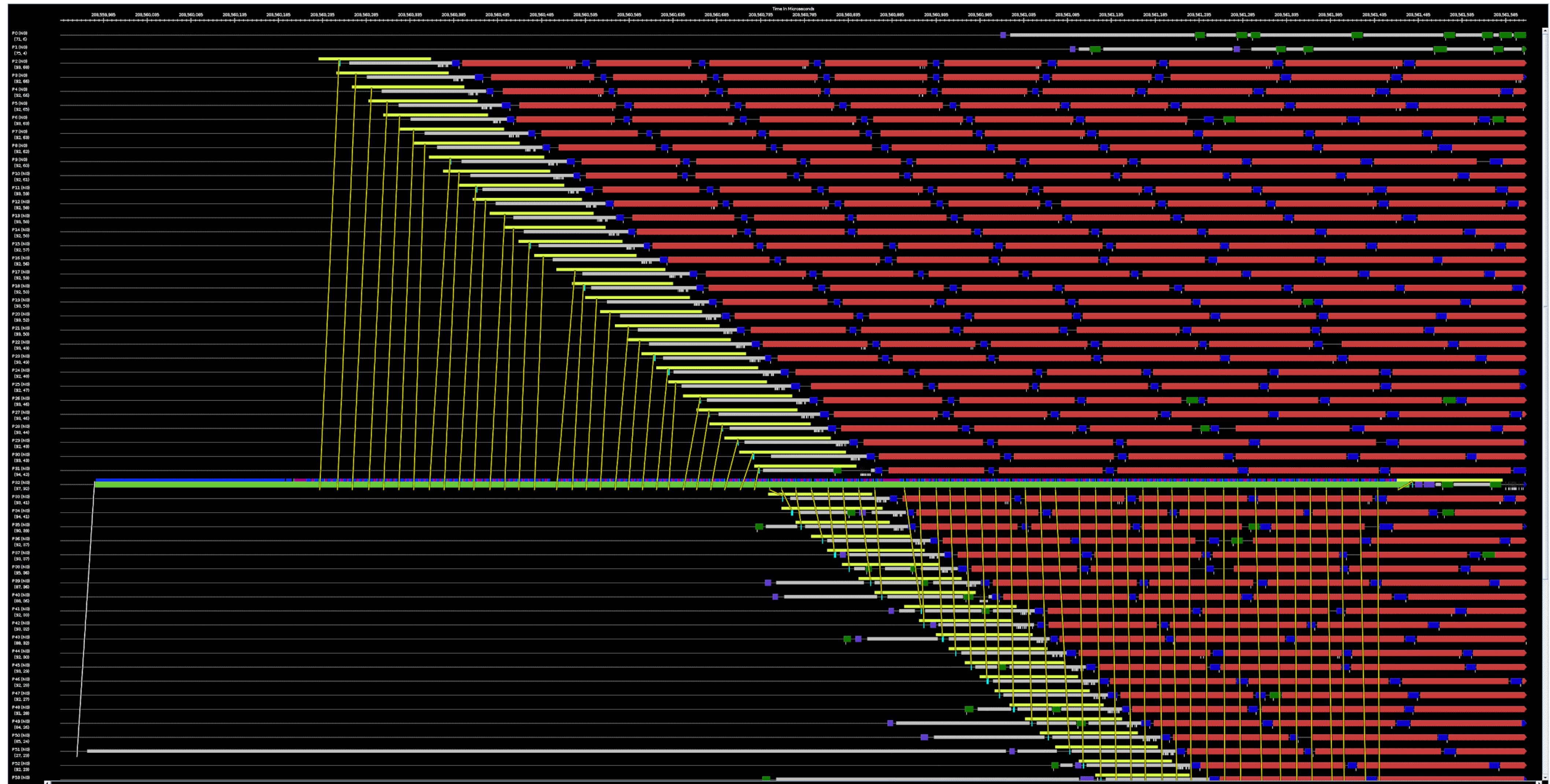
**Large system
benchmarking
performed by
Hanning Chen
(TACC)**



Performance Profile UCX vs IB VERBS vs MPI



PME Ungrid Calc Bottleneck



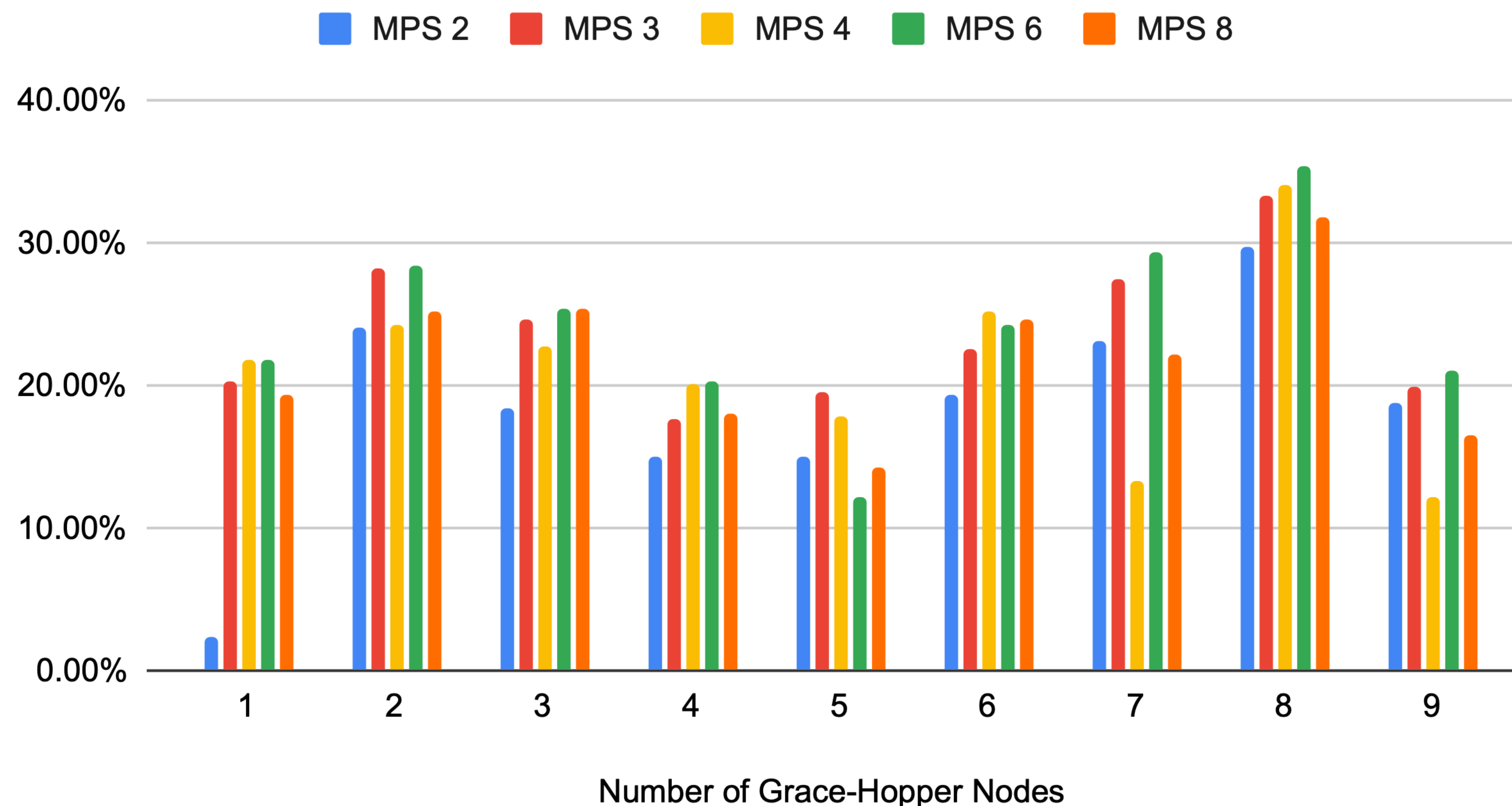
Bottleneck discussion

- The bottleneck in PME UngridCalc is due to serializing to a single GPU
- The design made sense on BlueWaters XK7 nodes
 - There were only effectively 8 cores per node
 - The Kepler in an XK7 had only 6 Gb of memory
- It has little impact on architectures with a similar ratio of cores to GPUs
- It has a large impact on Grace-Hopper with its 72 cores to 1 GPU ratio
- PME UngridCalc should be rewritten to be handled entirely within the GPU
 - That will require substantial amounts of developer time.
 - It is also the last link in a chain of bottlenecks stemming from PME

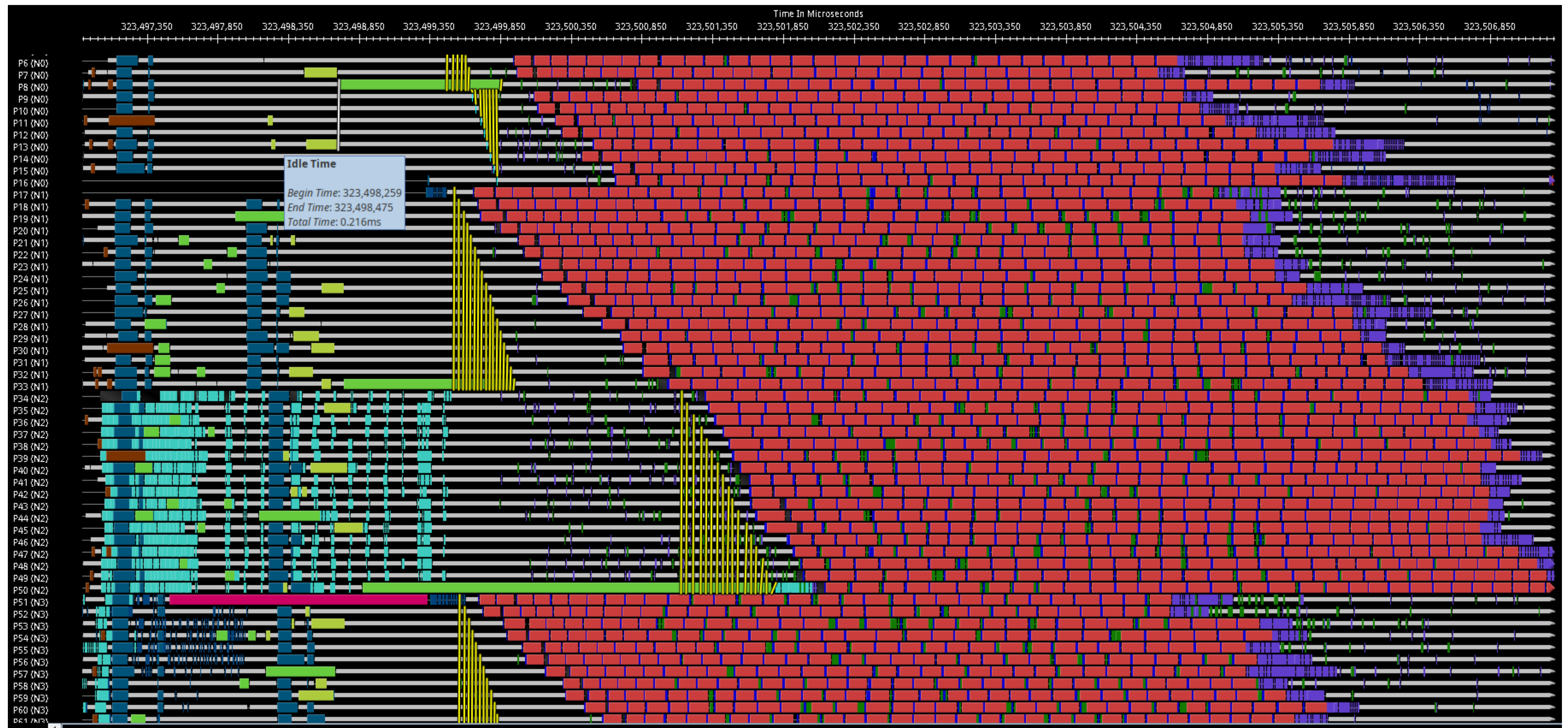
Mitigation: What if we pretend there are more GPUs?

- NVIDIA Multi-Process Service (MPS)
 - Virtualize a GPU to support multiple simultaneous accessors

UCX STMV8 Strong Scaling MPS Improvement

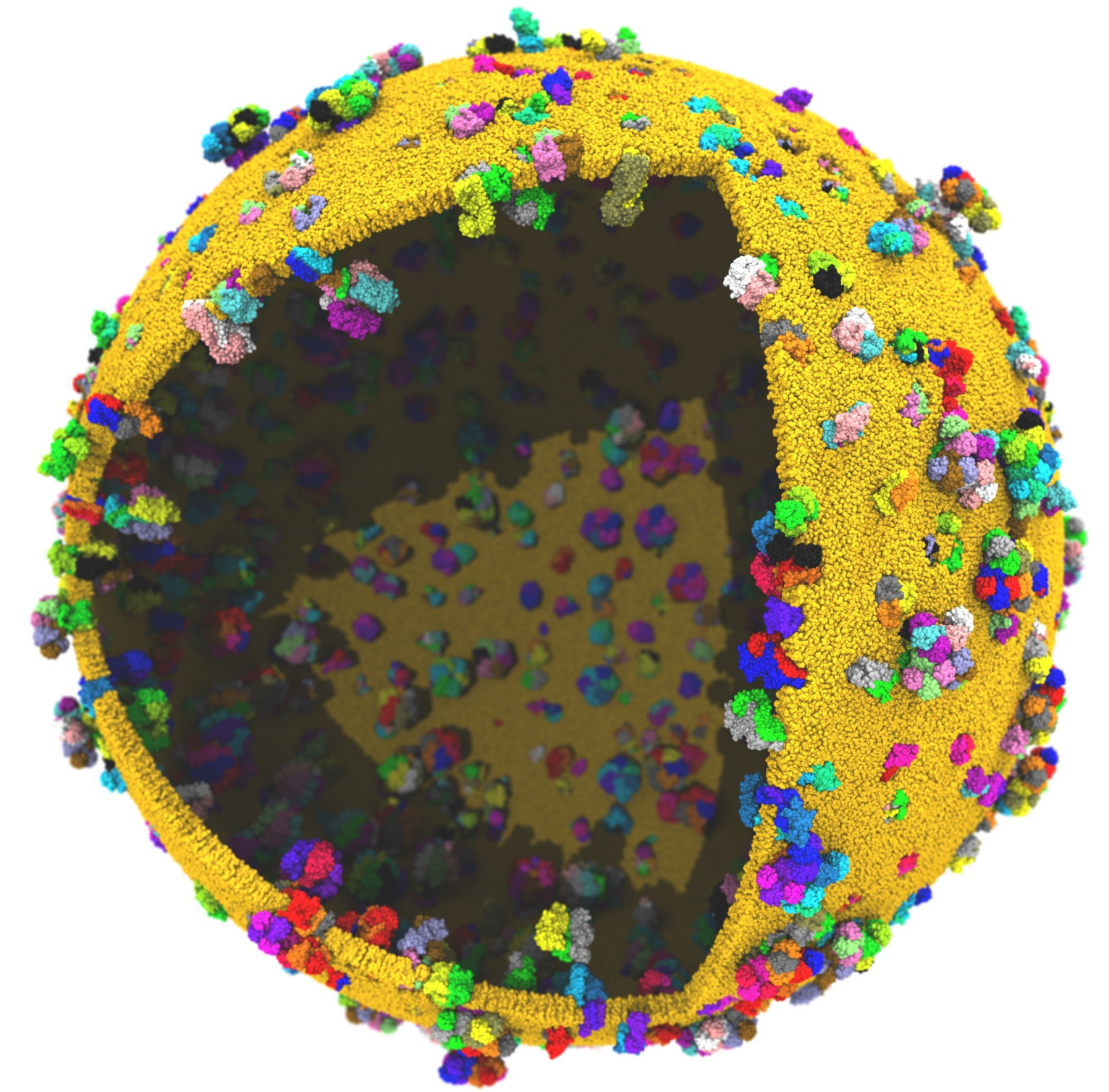
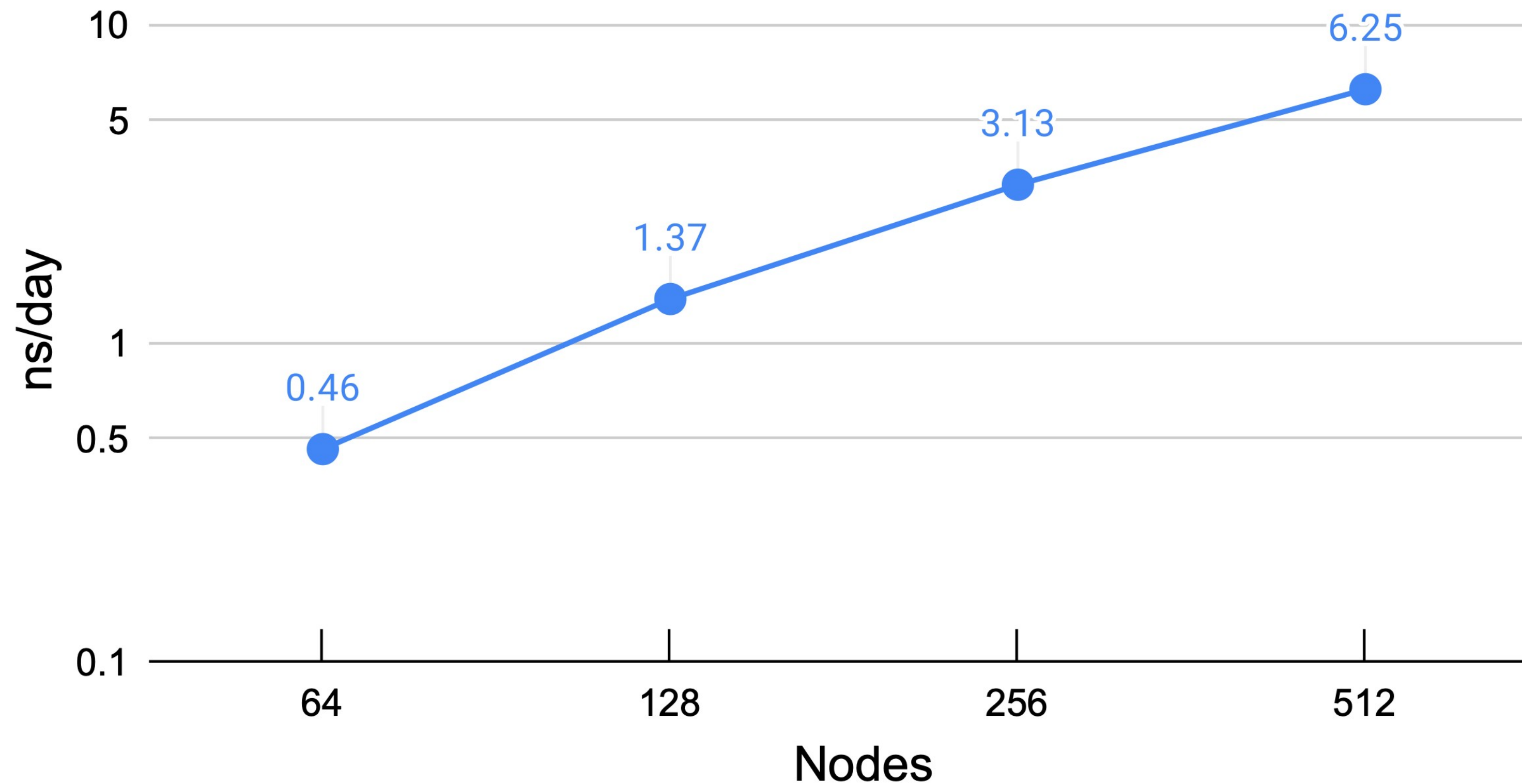


PME Ungrid Calc Bottleneck MPS 4



Simulating Protocell on TACC Vista

NAMD performance on Vista (GH200) simulating 1B atom protocell



**Protocell model
~1B atoms**

Summary and Future Work

NAMD3 scales large molecular systems reasonably well on TACC Vista

There are several performance limitations with the GPU Offload approach

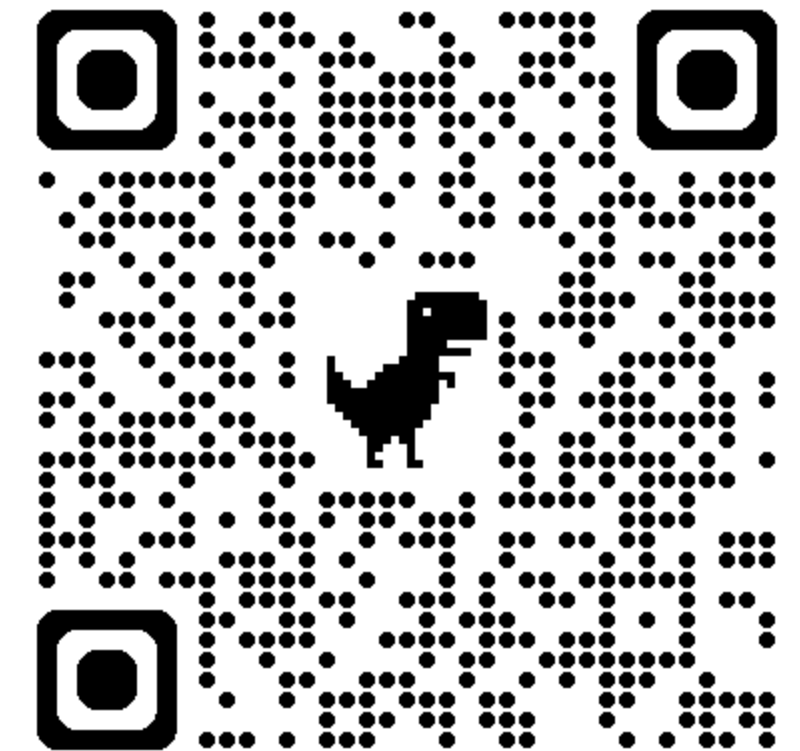
- MPS can mitigate some of them

Future Work (long term)

- Extending GPU Resident node to multiple nodes
- Use Charm++ reConverse's LCI to replace machine layers

Future Work (short term)

- Refactoring bottlenecks will improve strong scaling
- Improve GPU kernel performance
- Performance Optimization for TACC Horizon



Acknowledgments

- PIs: **Emad Tajkhorshid** (UIUC), **Laxmikant Kalé** (UIUC)
- Leadership: **Chris Chipot** (Université de Lorraine, UIUC), **Rafael Bernardi** (Auburn, UIUC), **Benoit Roux** (U Chicago)
- CUDA & HIP: **Eric Bohm** (UIUC), **Haochuan Chen** (UIUC), **David Clark** (NVIDIA), **David Hardy** (UIUC), Antti-Pekka Hynninen (NVIDIA, ORNL), **Julio Maia** (AMD, UIUC), Jim Phillips (UIUC), Mohammad Soroush Barhaghi (Schrödinger, UIUC), John Stone (NVIDIA, UIUC), David Tanner (Google, UIUC), Josh Vermaas (MSU, UIUC), Peng Wang (NVIDIA)
- SYCL/oneAPI: **Aditya Bhosale** (UIUC), Jaemin Choi (NVIDIA, UIUC), **David Hardy** (UIUC), **Wei Jiang** (ANL), Tareq Malas (Meta, Intel), **Ke Yue** (Intel)
- AVX-512: Mike Brown (Intel)
- TACC collaboration: John Cazes (TACC), **Hanning Chen** (TACC), **Amit Ruhela** (TACC)
- NIH grants P41-GM104601 and R24-GM145965
- Support from NVIDIA, Intel, AMD
- Computing resources from TACC, NCSA, OLCF, ALCF



NIH Center for Macromolecular Modeling and Bioinformatics (2019)
Beckman Institute, University of Illinois at Urbana-Champaign