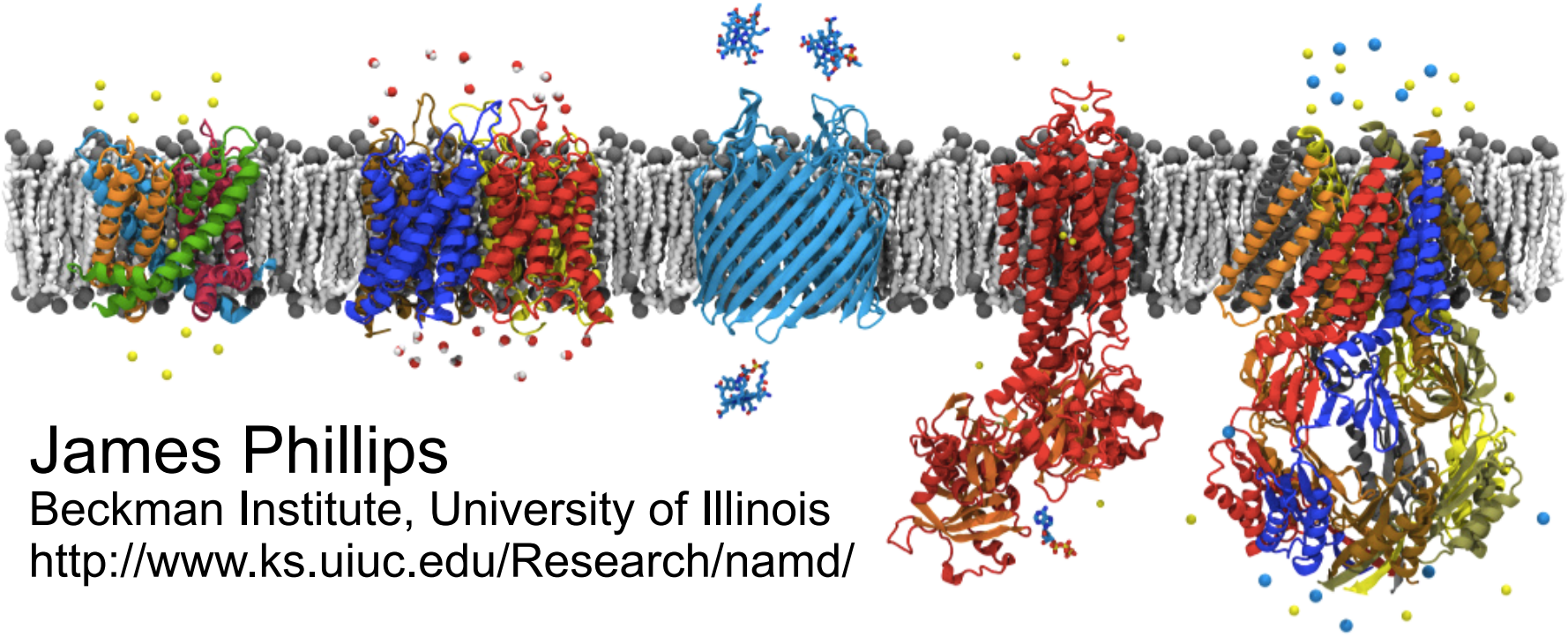
NAMD From the Developer's Perspective



James Phillips
Beckman Institute, University of Illinois
http://www.ks.uiuc.edu/Research/namd/

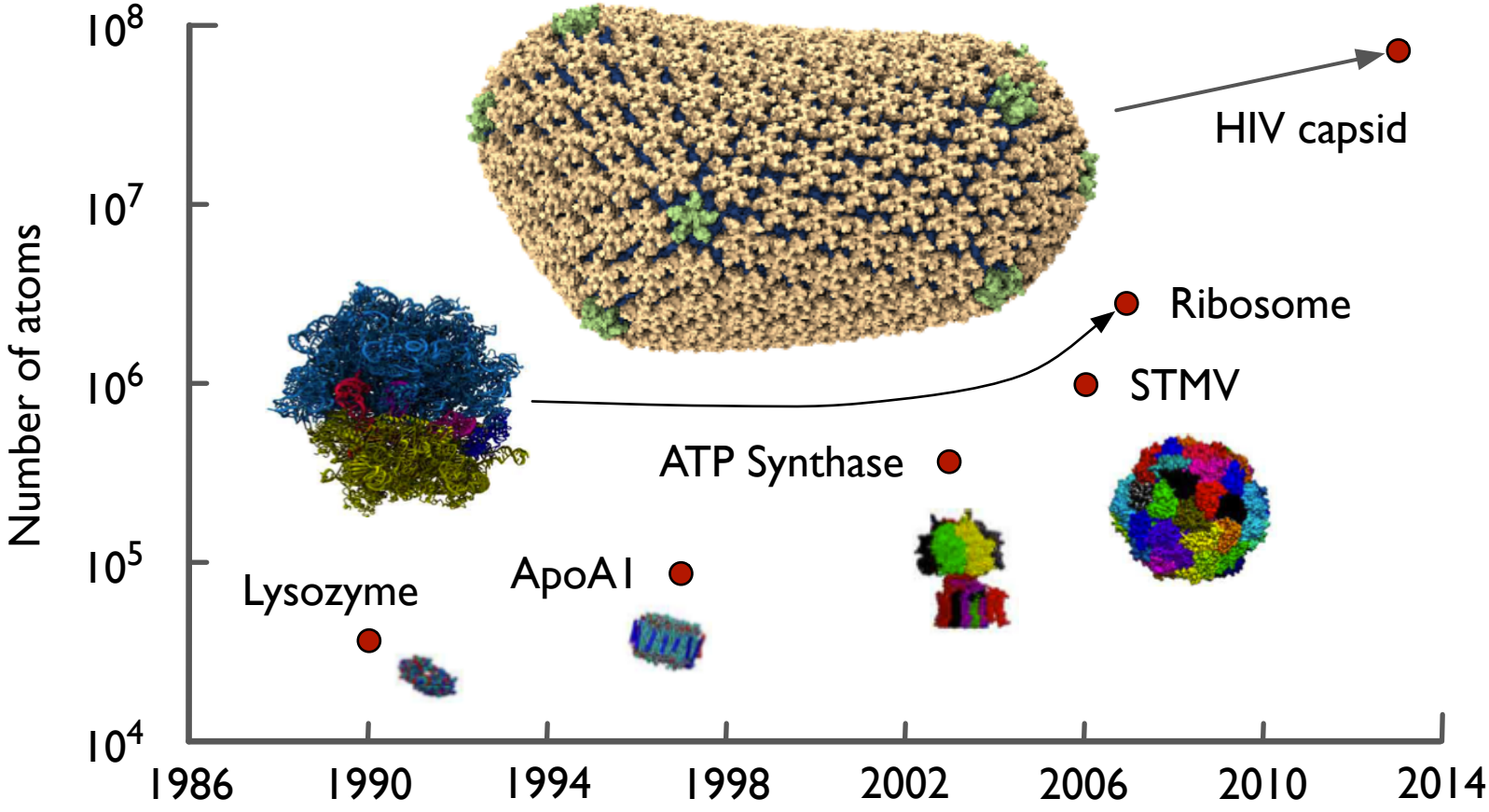
Our Partners: UIUC Parallel Programming Lab

Develops Charm++ Parallel Programming System

Siebel Center for Computer Science

http://charm.cs.illinois.edu/

NAMD Enables Larger Simulations on Larger Machines



NAMD Serves NIH Users and Goals

Practical Supercomputing for Biomedical Research

- 60,000 users can't all be computer experts.
  - 18% are NIH-funded; many in other countries.
  - 17,000 have downloaded more than one version.
  - 4000 citations of NAMD reference papers.
- One program available on all platforms.
  - Desktops and laptops – setup and testing
  - Linux clusters – affordable local workhorses
  - Supercomputers – free allocations on XSEDE
  - Blue Waters – sustained petaflop/s performance
  - GPUs - next-generation supercomputing
- User knowledge is preserved across platforms.
  - No change in input or output files.
  - Run any simulation on **any number of cores.**
- Available free of charge to all.

Hands-On Workshops

Oak Ridge TITAN

NAMD 2.10 Ships Soon

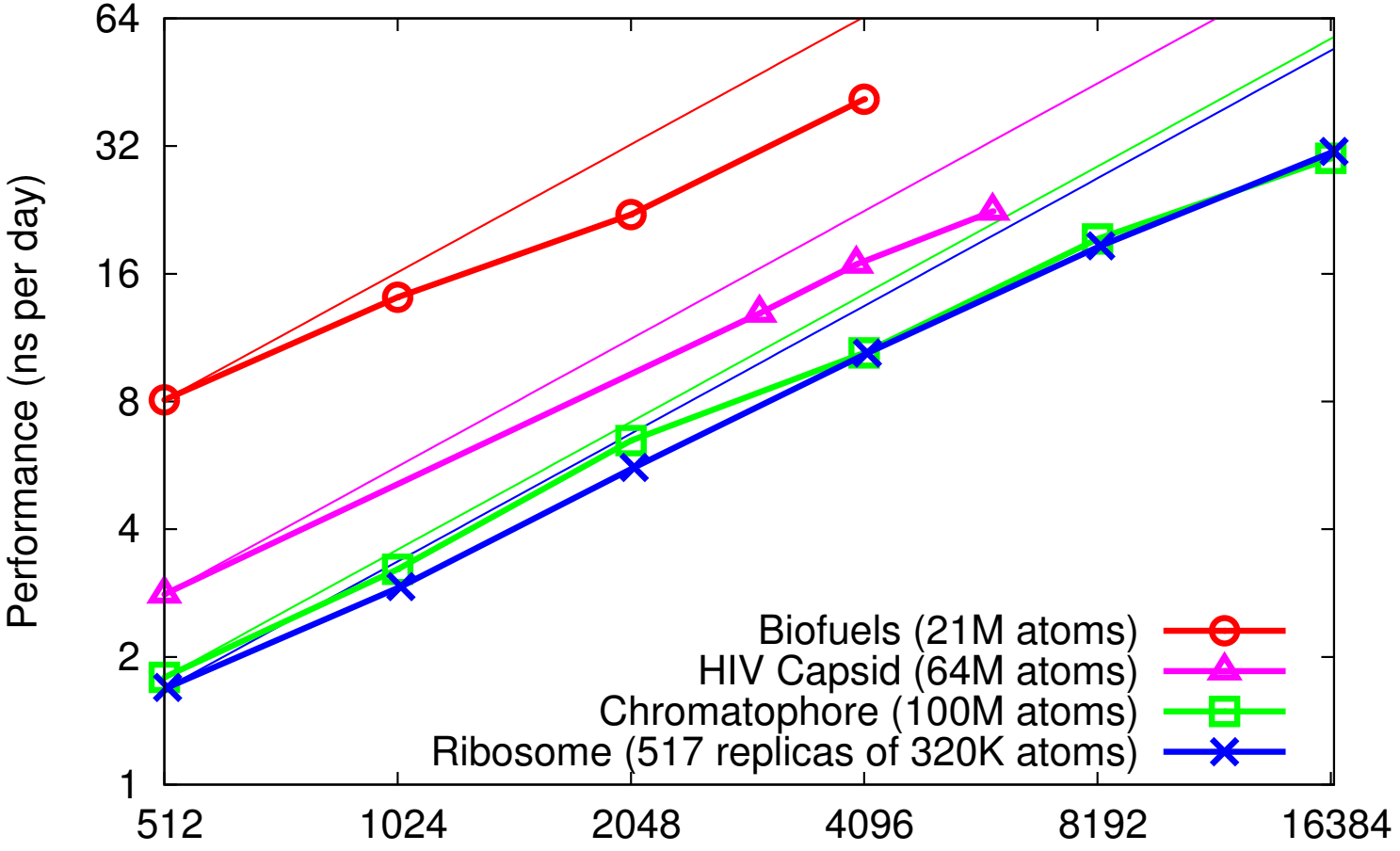
- First beta release hopefully November 2013 (this week)
  - Development code in production on Blue Waters and Titan
  - Nightly builds available publicly on NAMD download site
- Focus on enabling petascale simulations
- Type 1: Large systems of ~100 million atoms
  - Scalable to all of Blue Waters or Titan
  - Most steering methods now working
  - Minimizer more stable for large systems
- Type 2: Replica exchange simulations of smaller systems
  - Charm++ replica support improves scaling on Cray
- Various other improvements
  - Xeon Phi port, GPU improvements, Multilevel summation

Blue Waters Posed Many Challenges

- Scale NAMD to 100M atoms
  - Read new .js file format
  - Distribute or compress static molecular structure data
  - Parallel atomic data input
  - Use shared memory in a node
  - Parallel load balancing
  - Parallel, asynchronous trajectory and restart file output
  - 2D decomposition of 3D FFT
  - Limit steering force messages
  - Fix minimizer stability issues
- Also build benchmarks…

- Scale NAMD to 300K cores
  - Charm++ shared memory tuning
  - IBM Power7 network layer
  - IBM BlueGene/Q network layer
  - Cray Gemini network layer
  - Cray torus topology information
  - Charm++ replica layers
  - Optimize for physical nodes
  - Adapt trees to avoid throttling
  - Optimize for torus topology
  - Optimize for parallel filesystem
- Also optimize for GPUs…

NAMD Enables Varied Projects on Titan Cray XK7



(2fs timestep)

100 Million Atom Simulations Are Not Routine

- Simulation setup is a black art
  - Tools for adding solvent and ions don't scale
  - Need to move tools and users towards new "js" file format
- Still some rough edges
  - Not all NAMD features usable at scale
- Trajectory and restart output performance
  - New Charm++ I/O library will help address this
- Simulations require leadership machines
  - Available resources are limited, allocation process is slow
- Lack of setup/visualization/analysis facilities

Center Facilities Enable Petascale Biology

External Resources,
90% of our
Computer Power

Over the past five years the Center has assembled all necessary hardware and infrastructure to prepare and analyze petascale molecular dynamics simulations, and ***makes these facilities available to visiting researchers***.

Simulation Output

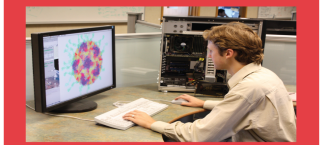
10 Gigabit Network

Petascale Gateway Facility

Storage

Compute

Visualization



High-End Workstations
Accessible to Visitors

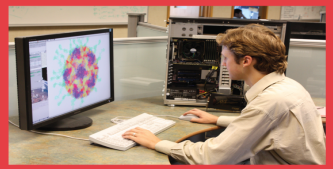
Virtual Facilities Enable Petascale Anywhere

High-end visualization and analysis workstations currently available only in person at the Beckman Institute must be ***virtualized and embedded at supercomputer centers***.

Storage

Compute

Visualization

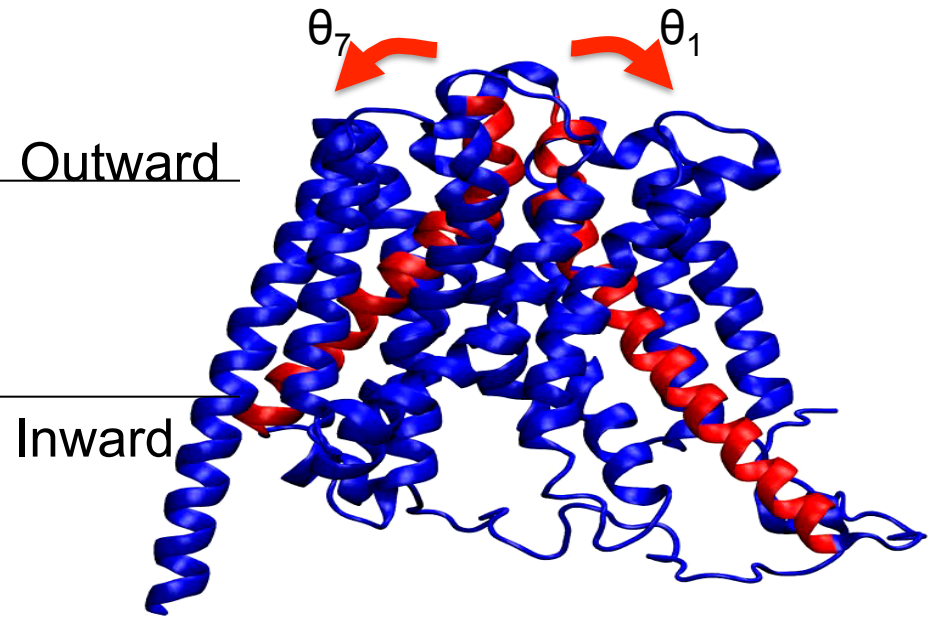


Compressed Video

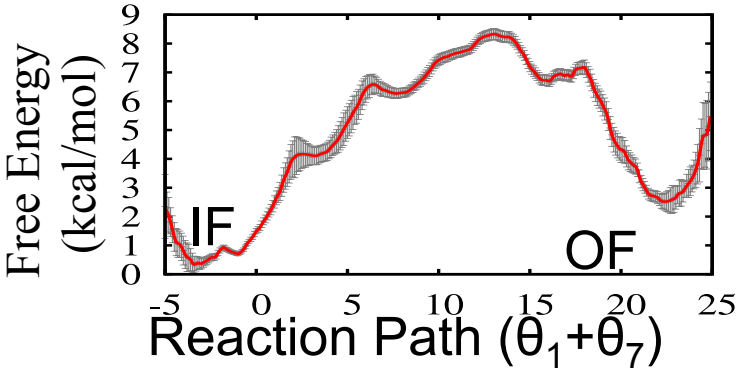
1 Gigabit Network

Replica Exchange Enables Advanced Sampling

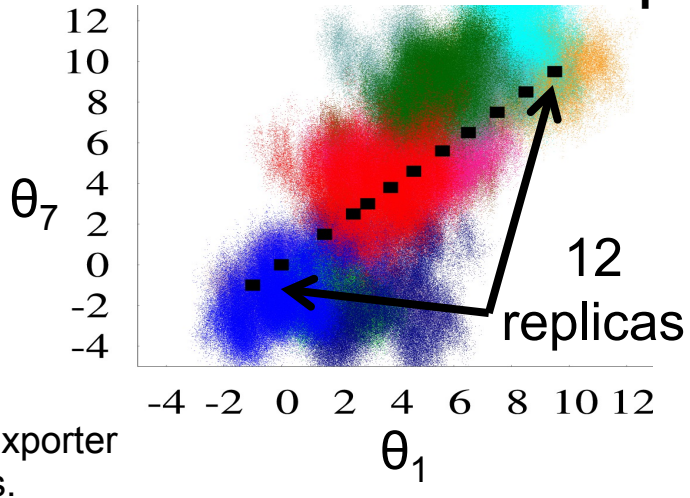
Bias-Exchange Umbrella Sampling
on quaternion-based order parameters



GlpT transporter in membrane



Efficient Reaction Path Sampling



M. Moradi and E. Tajkhorshid, Conformational Transition of ABC Exporter MsbA: A Mechanistic Picture at Atomic Resolution. PNAS, in press.

NAMD 2.10 Replica Exchange

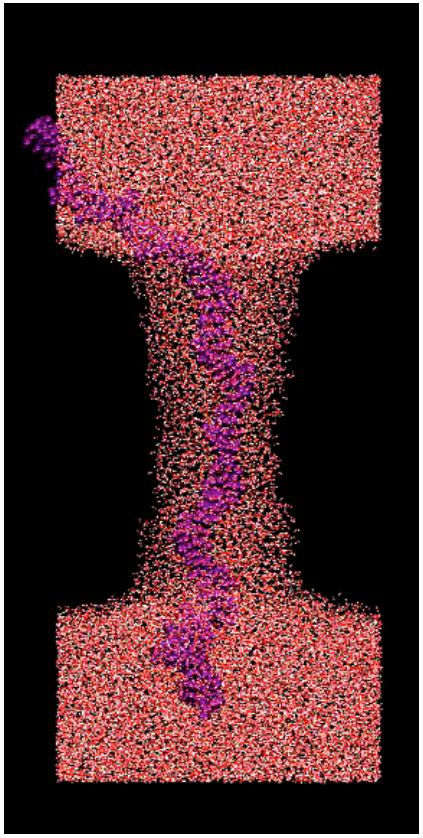
- More general Charm++ integration:
  - NAMD 2.9 used MPI communicator splitting, required MPI version
  - NAMD 2.10 splits replicas in Charm++ low-level runtime (LRTS)
  - LRTS underlies MPI, Cray (uGNI), and BlueGene/Q (PAMI) implementations
  - New LRTS machine layers added for InfiniBand (“verbs”) and ethernet (“netlrts”)
- Basis for many enhanced sampling methods:
  - Parallel tempering (temperature exchange)
  - Umbrella sampling for free-energy calculations
  - Hamiltonian exchange (alchemical or conformational)
  - Finite Temperature String method
  - Nudged elastic band
- Better scaling for individual replicas:
  - **Native Charm++ layers essential for multi-node GPU replicas**
  - Topology-aware placement on Cray and BlueGene torus networks

Same Tcl scripts as NAMD 2.9

Future work enabled by Charm++ integration

Adaptability Through Scripting

- Tcl customizations are **portable**
- Top-level protocols:
  - Minimize, heat, equilibrate
  - Simulated annealing
  - Replica exchange (originally via sockets)
- Long-range forces on selected atoms
  - Torques and other steering forces
  - Adaptive bias free energy perturbation
  - Coupling to external coarse-grain model
- Special boundary forces
  - Applies potentially to every atom
  - Several optimizations for efficiency
  - Shrinking phantom pore for DNA



NAMD GPU: What is accelerated?

Accelerated

- Short-range non-bonded
  - Cutoff or with PME
  - w/ or w/o energy calculation
- Implicit solvent
- NVIDIA GPUs only

Not Accelerated

- Bonded terms
- PME reciprocal sum
- Integration
- Rigid bonds
- Grid forces
- Collective variables
- Etc.

NAMD GPU: What is disabled?

Disabled

- Alchemical (FEP and TI)
- Locally enhanced sampling
- Tabulated energies
- Drude (nonbonded Thole)
- Go forces
- Pairwise interaction
- Pressure profile

Not Disabled

- Memory optimized builds
- Conformational free energy
- Collective variables
- Grid forces
- Steering forces
- Almost everything else

NAMD GPU: What is different?

- Forces
  - Slightly less accurate than CPU
    - Different interpolation scheme, single precision
    - Also affects pressure calculation
- Energies
  - Don’t match forces as closely as on CPU
  - Constant for interactions less than 1 Angstrom
    - Minimizer issues in NAMD 2.9 fixed in 2.10
    - Velocity-quenching minimizer should work in 2.9

NAMD GPU: Performance

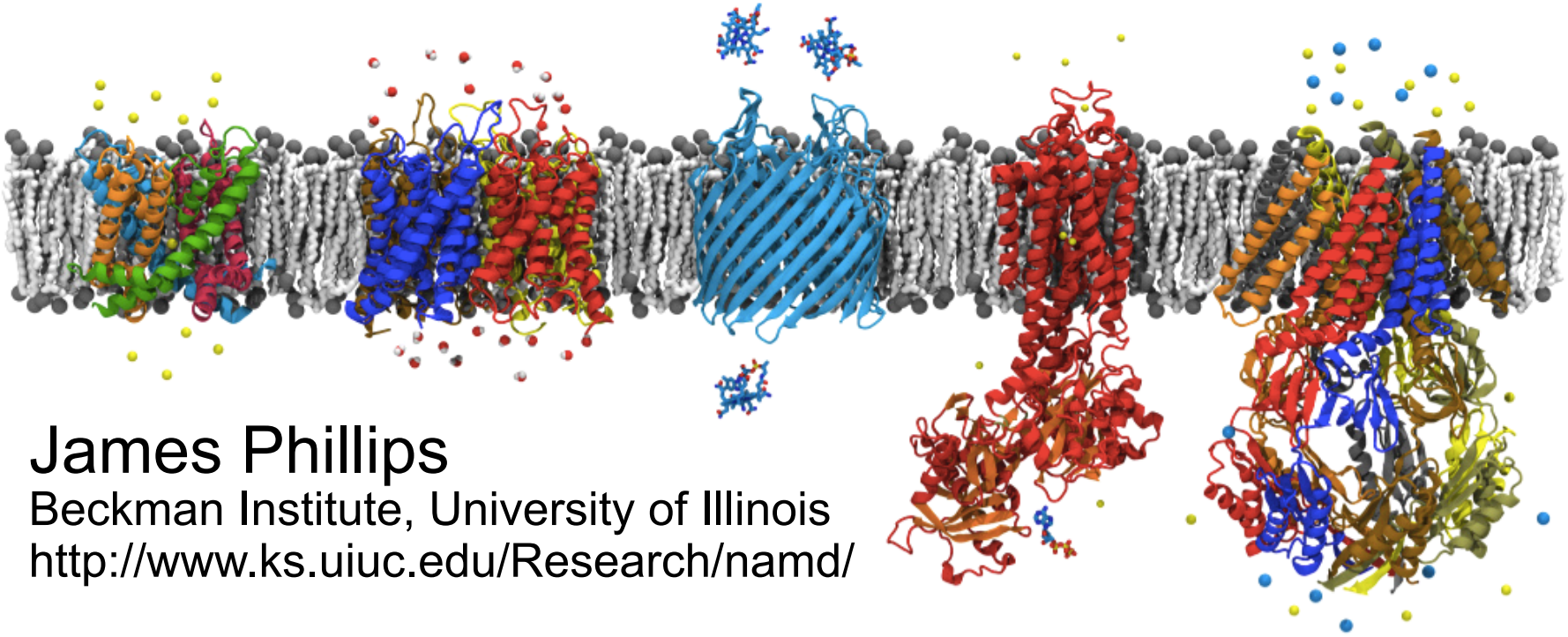
What to expect

- 1 GPU = ~24 CPU cores
  - Depending on CPU and GPU
- Scaling to 10K atoms/GPU
  - Assuming fast network
- Must use smp/multicore
  - Many cores share each GPU
  - Use multicore for single node
  - At most one process per GPU

Why it may be worse

- Weak GPU (e.g., laptop)
- Too few CPU cores used
- Coarse-grained simulation
- Too few atoms per GPU
- Limited by network
- Limited by MPI (use ibverbs)
- Limited by special features

Thank You!



James Phillips
Beckman Institute, University of Illinois
http://www.ks.uiuc.edu/Research/namd/