NAMD 3 Design Issues

Reflections for a Second Decade of Scalable Molecular Dynamics

Jim Phillips

Theoretical and Computational Biophysics Group

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

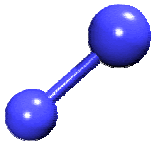
What is NAMD?

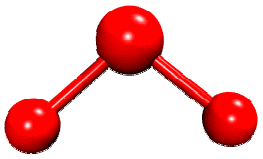
- Molecular dynamics and related algorithms
  - e.g., minimization, steering, locally enhanced sampling, alchemical and conformational free energy perturbation
- Efficient algorithms for full electrostatics
- Effective on affordable commodity hardware
- Read file formats from standard packages:

  X-PLOR (NAMD 1.0), CHARMM (NAMD 2.0),

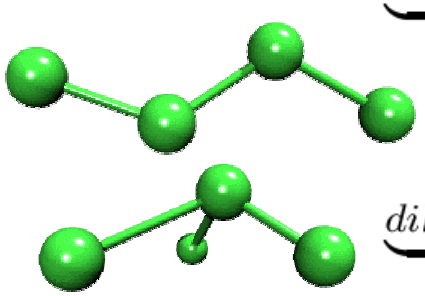
  Amber (NAMD 2.3), GROMACS (NAMD 2.4)
- Building a complete modeling environment

Molecular Mechanics Force Field





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

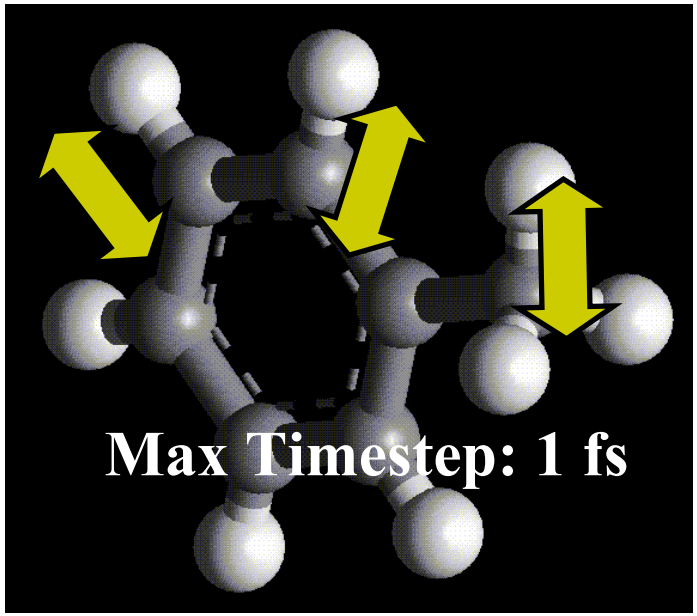


$$\underbrace{\sum_{dihedrals} k_i^{dihe}[1 + \cos(n_i\phi_i + \delta_i)]}_{U_{dihedral}} +$$

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

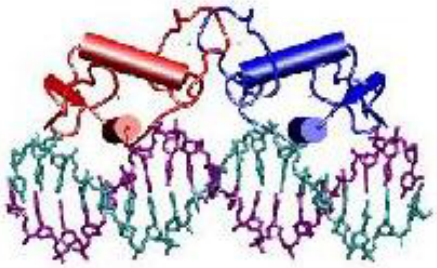
Biomolecular Time Scales

Motion	Time Scale (sec)
Bond stretching	10^{-14} to 10^{-13}
Elastic vibrations	10^{-12} to 10^{-11}
Rotations of surface sidechains	10^{-11} to 10^{-10}
Hinge bending	10^{-11} to 10^{-7}
Rotation of buried side chains	10^{-4} to 1 sec
Allosteric transistions	10^{-5} to 1 sec
Local denaturations	10^{-5} to 10 sec

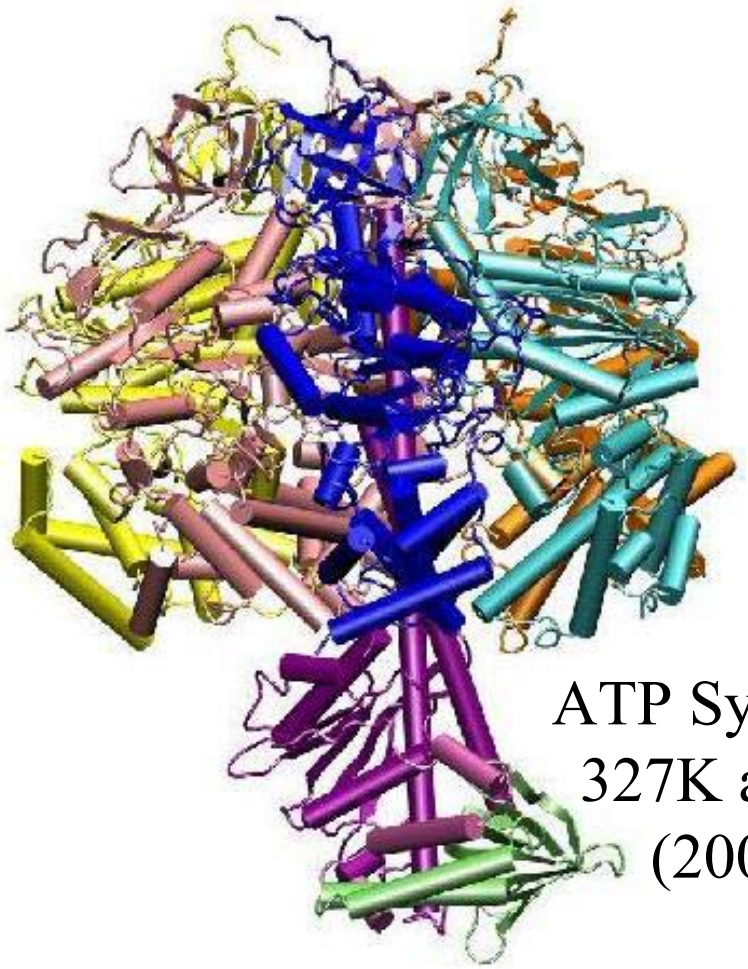


Sizes of Simulations Over Time

BPTI
3K atoms



Estrogen Receptor
36K atoms (1996)



ATP Synthase
327K atoms
(2001)

Our Solution: Parallel Computing

HP 735 cluster
12 processors
(1993)

SGI Origin 2000
128 processors (1997)

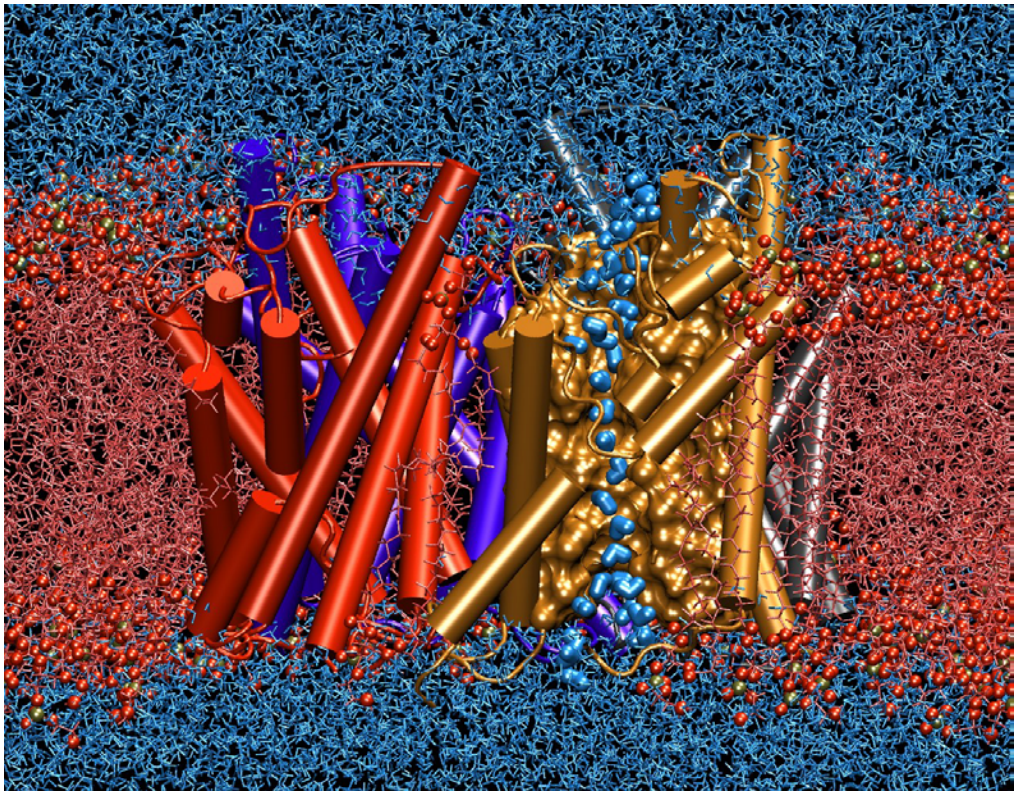
PSC Lemieux AlphaServer SC
3000 processors (2002)

Example Simulation: GlpF

NAMD with PME
Periodic boundary conditions
NPT ensemble at 310 K

Protein:	~	15,000 atoms
Lipids:	~	40,000 atoms
Water:	~	51,000 atoms
Total:	**~**	**106,000 atoms**

1024 PSC TCS CPUs
4 hours per ns



M. Jensen, E. Tajkhorshid, K. Schulten, *Structure* **9,** 1083 (2001)
E. Tajkhorshid et al., *Science* **296**, 525-530 (2002)

Typical Simulation Statistics

- 100,000 atoms (including water, lipid)
- 10-20 MB of data for entire system
- 100 A per side periodic cell
- 12 A cutoff of short-range nonbonded terms
- 10,000,000 timesteps (10 ns)
- 4 s/step on one processor (1.3 years total!)

Parallel MD: Easy or Hard?

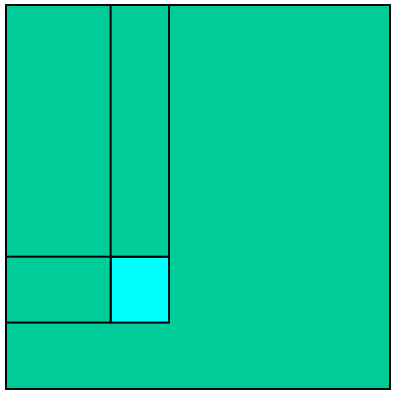
- Easy
  - Tiny working data
  - Spatial locality
  - Uniform atom density
  - Persistent repetition
  - Multiple timestepping
- Hard
  - Sequential timesteps
  - Short iteration time
  - Full electrostatics
  - Fixed problem size

Basis of NAMD Scalability

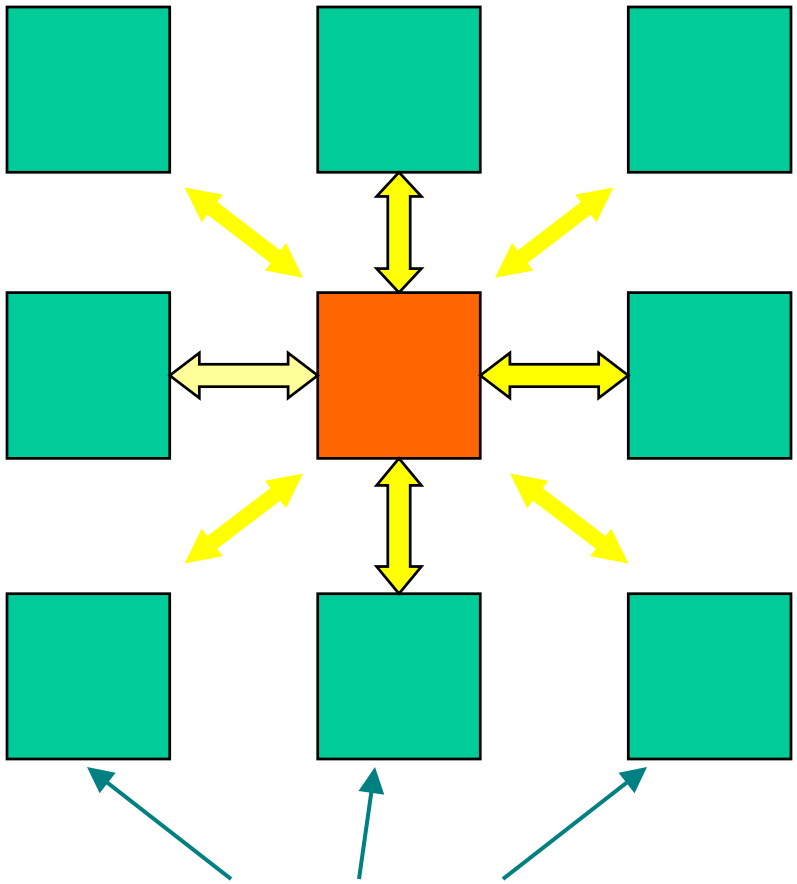
- Very active computer science collaboration
  - UIUC Parallel Programming Lab, since 1992
  - Charm++ system: message driven objects
  - Constant tuning for evolving parallel platforms
- Designed for parallel efficiency
  - NAMD 1: discrete spatial decomposition, fast multipole
  - NAMD 2: hybrid force-spatial decomposition, PME
  - Dependency-driven execution, no barriers
  - Measurement-based load balancing system

Poorly Scaling Approaches

- Replicated data
  - All atom coordinates stored on each processor
  - Communication/Computation ratio: O(P log P)
- Partition the atom array across processors
  - Nearby atoms may not be on the same processor
  - C/C ratio: O(P)
- Distribute force matrix to processors
  - Matrix is sparse, non uniform
  - C/C Ratio: O(sqrt P)

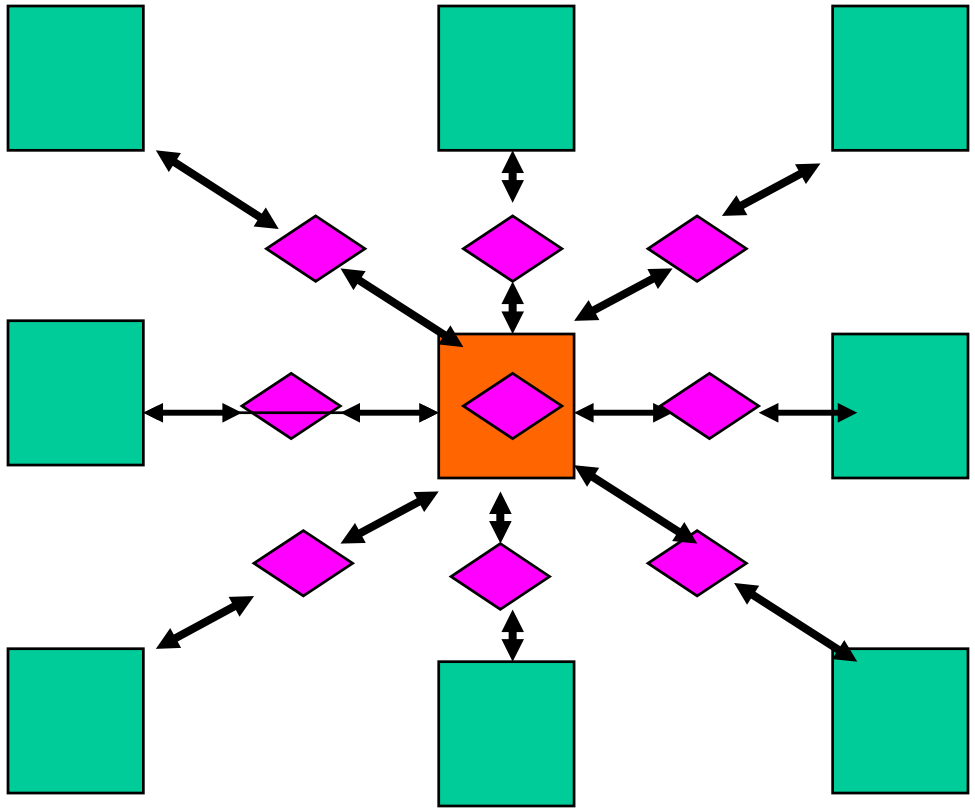


Spatial Decomposition: NAMD 1



- Atoms spatially distributed to cubes
- Size of each cube :
  - Just a larger than cut-off radius
  - Communicate only w/ neighbors
  - Work for each pair of neighbors
- C/C ratio: O(1)
- However:
  - Load Imbalance
  - Limited Parallelism

Hybrid Decomposition: NAMD 2

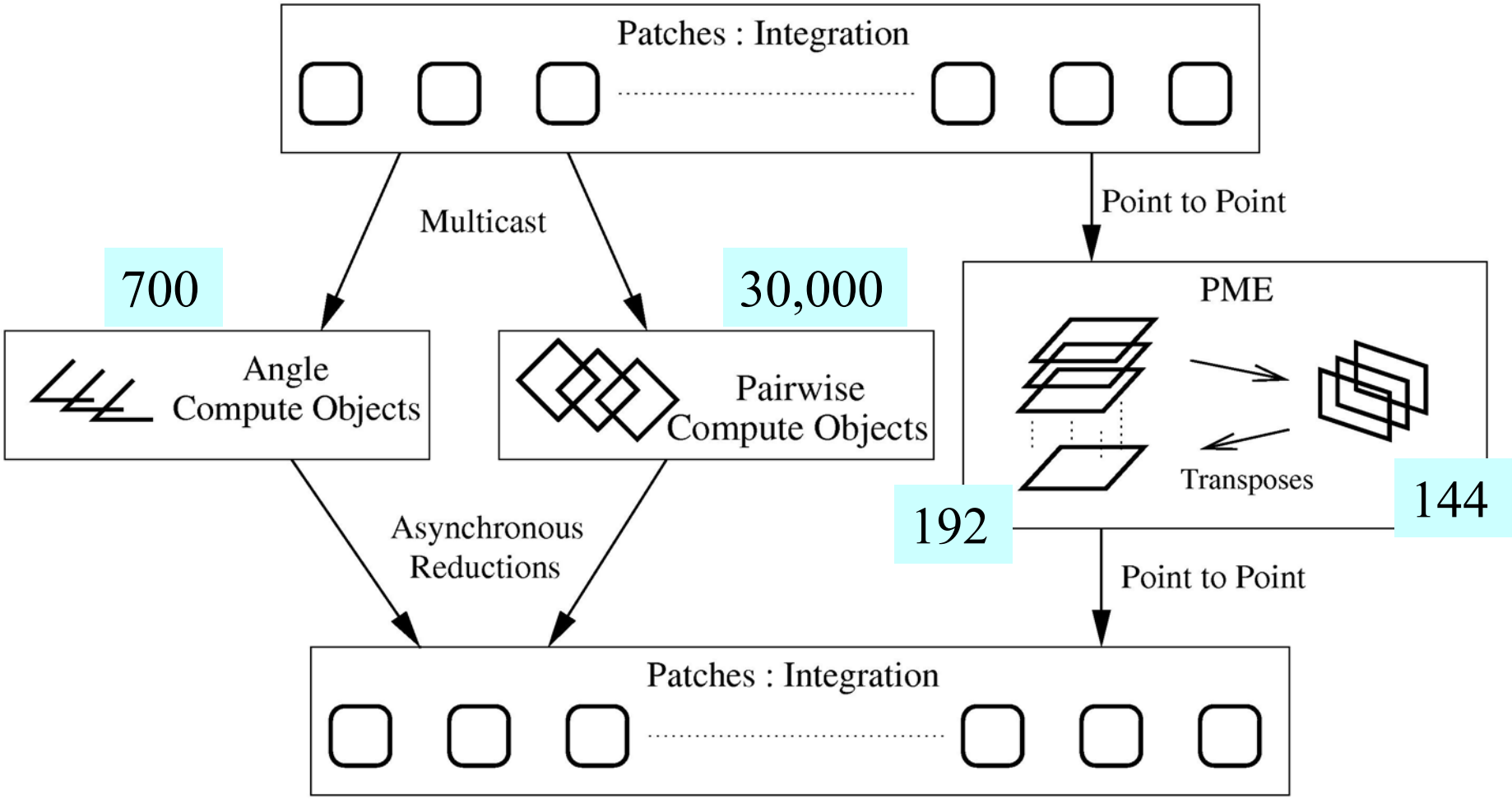


- Spatially decompose data and communication.
- Separate but related work decomposition.
- "Compute objects" facilitate iterative, measurement-based load balancing system.

Particle Mesh Ewald

- Particle Mesh Ewald (PME) calculation adds:
  - A global grid of modest size (e.g. 192x144x144).
  - Distributing charge from each atom to 4x4x4 sub-grid.
  - 3D FFT over the grid, hence O(N log N) performance.
- Strategy:
  - Use a smaller subset of processors for PME.
  - Overlap PME with cutoff computation.
  - Use same processors for both PME and cutoff.
  - Multiple time-step reduces scaling impact.

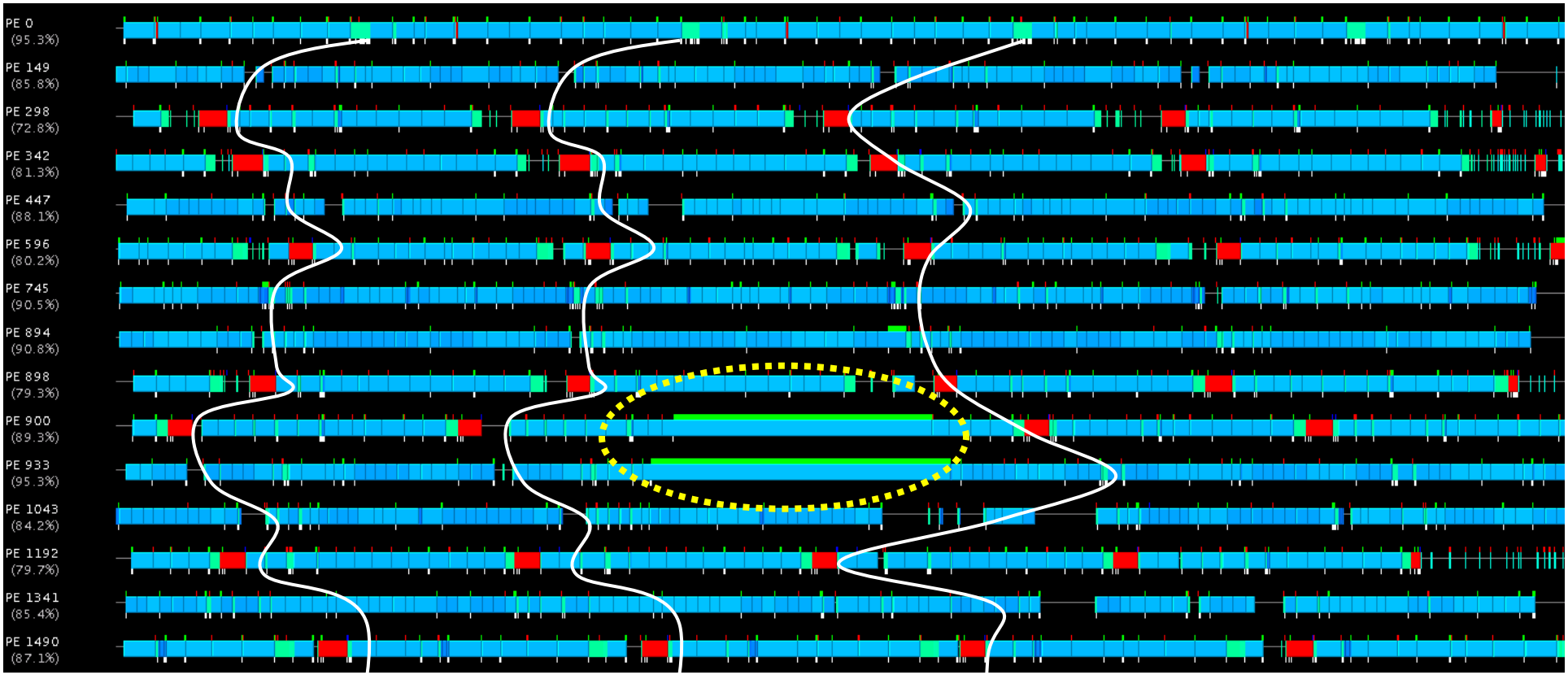
NAMD 2 w/PME Parallelization using Charm++



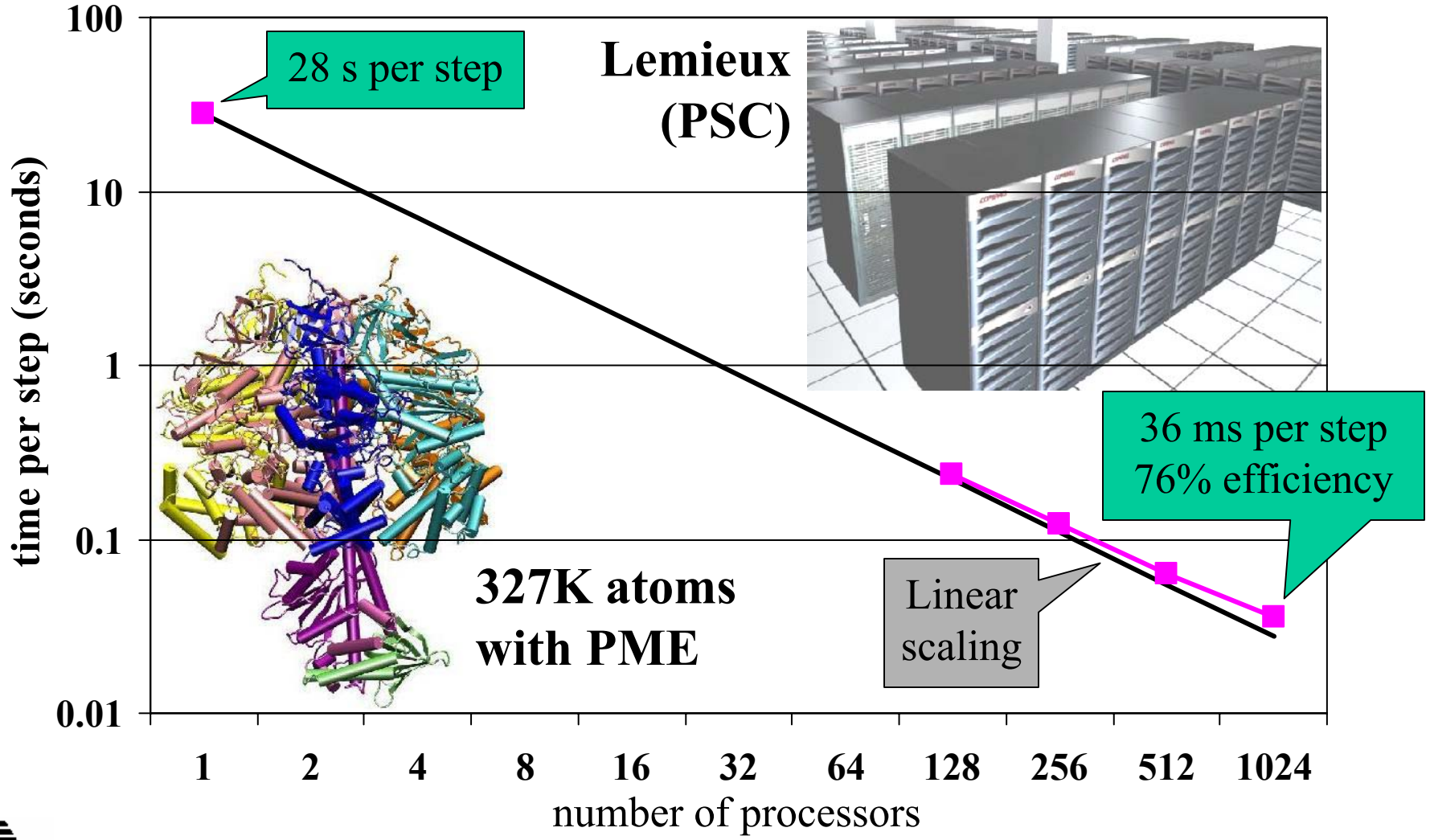
Avoiding Barriers

- In NAMD:
  - The energy reductions were made asynchronous.
  - No other global barriers are used in cut-off simulations.
- This came handy when:
  - Running on Pittsburgh Lemieux (3000 processors).
  - The machine (and how Converse uses the network) produced unpredictable, random communication delay.
    - A send call would remain stuck for 20 ms, for example.
  - Each timestep, ideally, was 12-14 ms.

Handling Network Delays

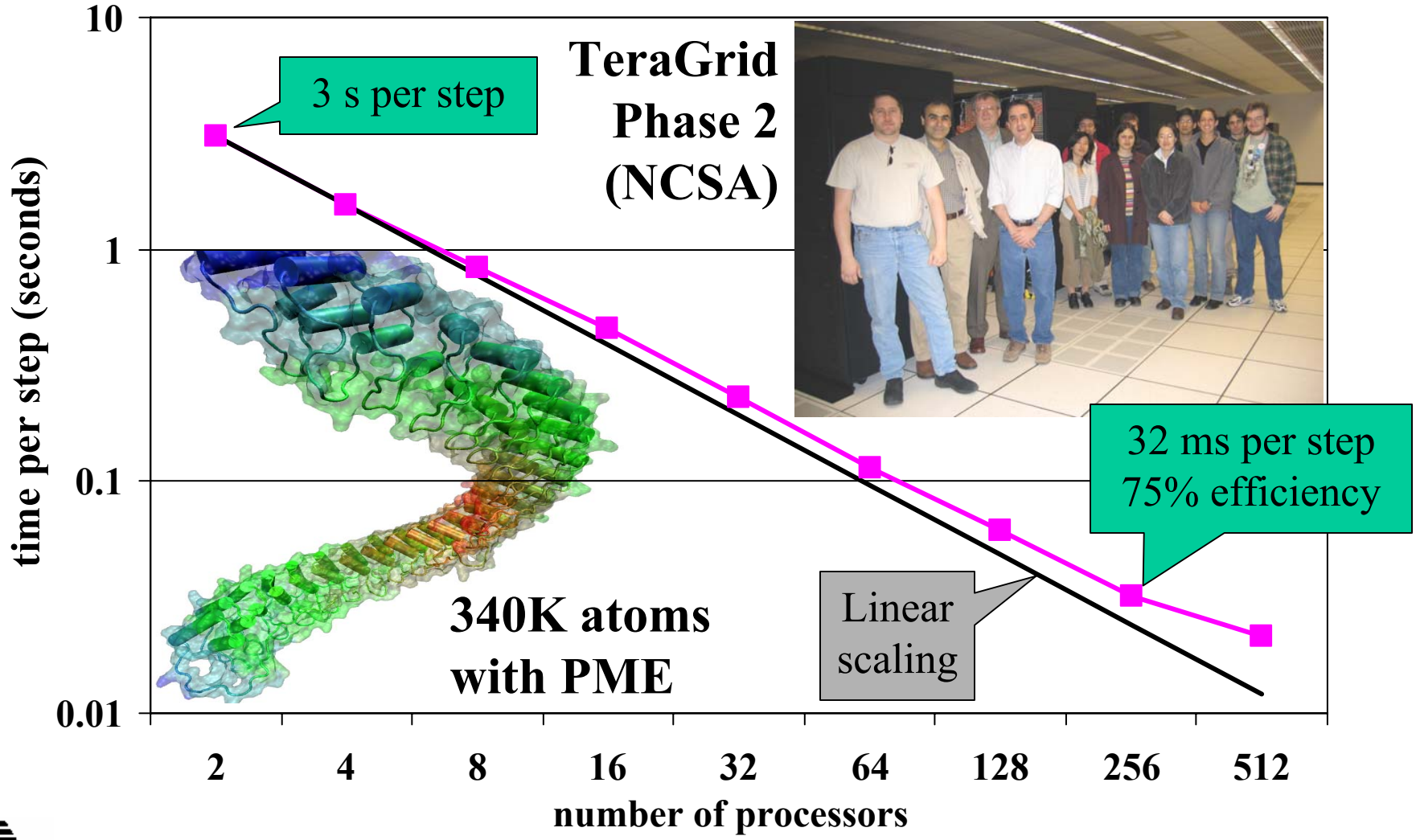


SC2002 Gordon Bell Award

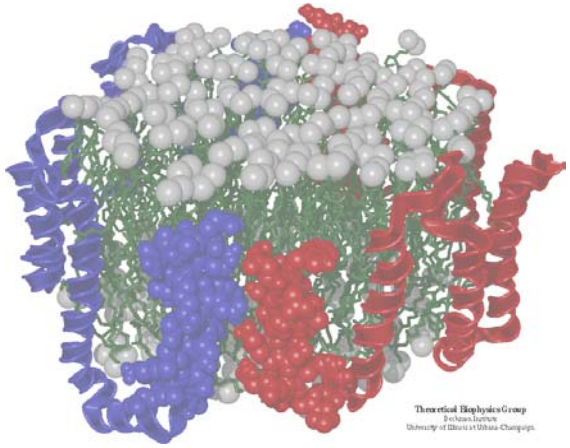


National Center for Research Resources

Ankyrin on TeraGrid Phase 2



Performance on Local Linux Clusters



92K atoms w/ PME

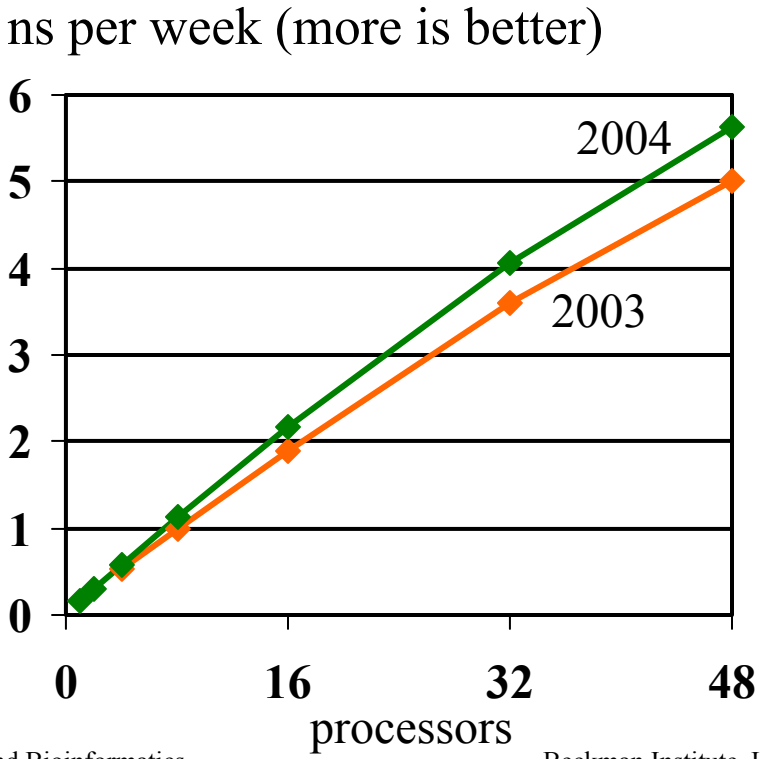
New Xeon clusters are only 12% faster for typical production runs, but much more compact, versatile, and manageable.

Older (2003):

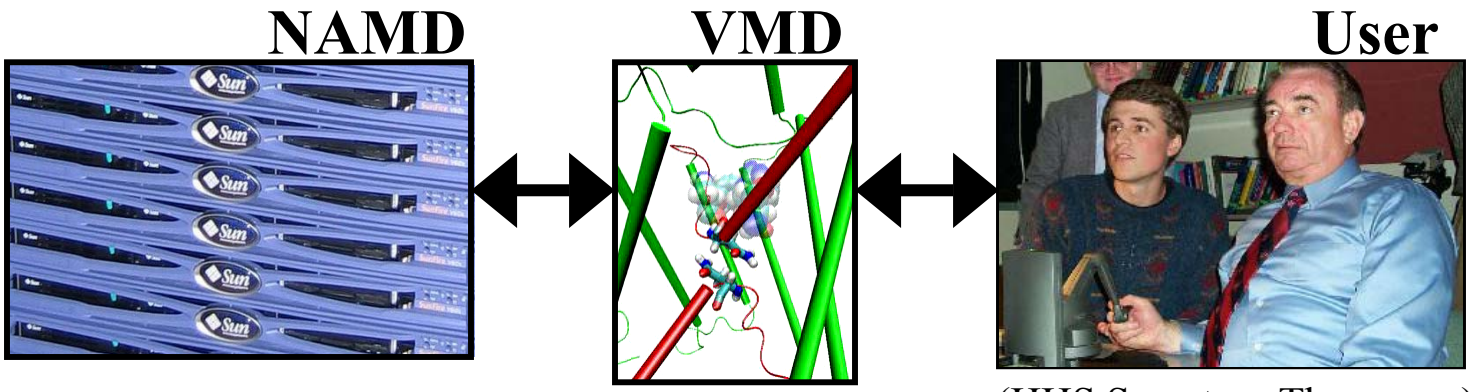
- 24 Dual Athlon 2133 MHz
- Desktop tower
- No hard drives
- Clustermatic 3 on RedHat 8.0
- Floppy boot
- $1200 per node

Newer (2004):

- 24 Dual Xeon 3.06 GHz
- Rackmount
- Fully loaded
- Clustermatic 4 on RedHat 9.0
- Network boot
- $2000 per node

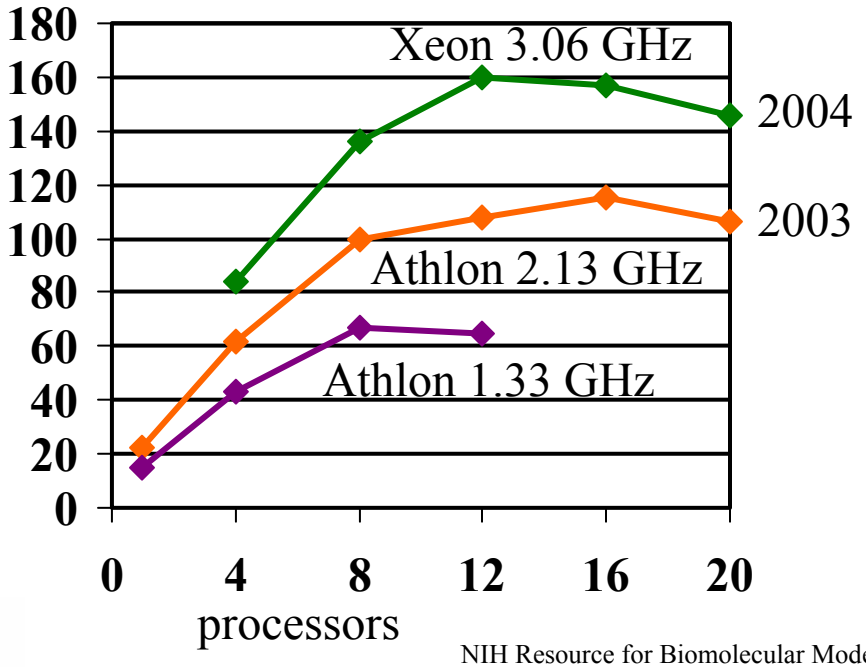


Interactive Molecular Dynamics



(HHS Secretary Thompson)

steps per second (more is better)



New Xeon clusters give a 39% boost to IMD performance.

GlpF IMD Benchmark:

- 4210 atoms
- 3295 fixed atoms
- 10A cutoff, no PME
- Limited by network latency

Platforms on the Horizon

- IBM Blue Gene/L (2005)
  - Currently scaling to 1000 (slow) CPUs, collaboration in place.
- UIUC Macintosh G5 cluster (2005)
  - Hardware sitting in boxes, waiting on machine room.
- PSC Cray Red Storm Opteron cluster (2005)
  - Should be fast, but new OS may have Charm thread problems.
- Graphics Cards with Programmable Shaders (2005)
  - Low cost, ubiquitous acceleration, but painful programming.
- ClearSpeed Accelerator Boards (2005)
  - Small memory SIMD, Gromacs port already underway.
- MDGRAPE-3 Accelerator Boards (2006?)
  - Even faster, if we can use the algorithms as implemented.

NAMD 3 Vision

- Make NAMD a widely used MD program
  - For large molecular systems,
  - Scaling from PCs, clusters, to large parallel machines
  - For interactive molecular dynamics
- Specific Goals for NAMD 3:
  - High performance: sub-ms timesteps
  - Ease of use: simple to configure, set-up, and run
  - Ease of modification (for developers and **users**)
  - Incorporation of features needed by scientists

Step One: Modern Charm++

- NAMD 2:
  - Groups explicitly control data and work objects.
  - Proxies make data available on nodes as required.
  - Explicit messages and entry points drive progress.
- NAMD 3 (and LeanMD already):
  - System-managed arrays distribute data and work.
  - Communication optimized by multicast and reduction.
  - Explicit control flow expressed in structured dagger.
- User-visible benefits:
  - Advanced checkpoint/restart and node-death tolerance.
  - Dynamically resize running jobs via faucets scheduler.

Step Two: Modular Design

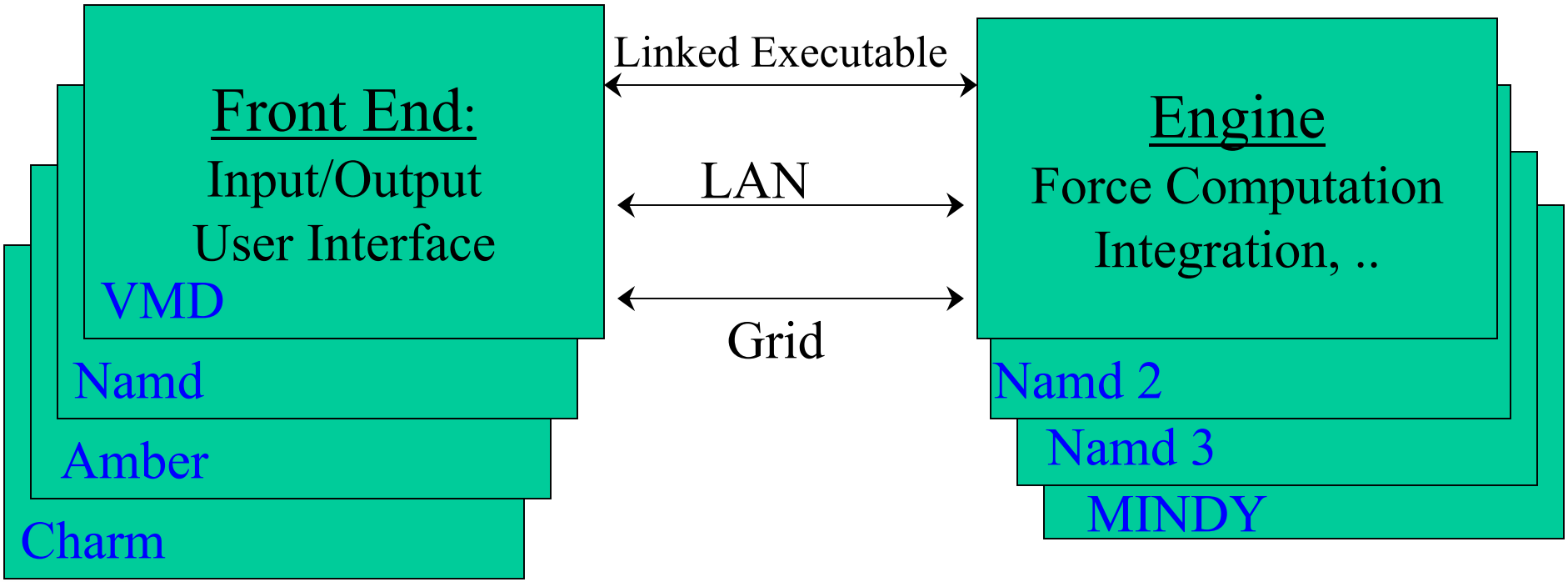
- Modifications for a new force in NAMD 2:
  - SimParameters, Parameters, ComputeMgr, WorkDistrib
  - Possibly also ReductionMgr, Sequencer, LdbCoordinator, …
- NAMD 3:
  - Write a single module to define all aspects of the force.
  - Multiple instances of the same force in one simulation.
- User-visible benefits:
  - Redesigned and legacy modules coexist in one code.
  - Customization and extension are more convenient.
  - Experiment with new methods and analysis in scripts.

Step Three: Orthogonal Layers

- NAMD 2:
  - All forces deal with periodic boundaries, pressure calculation, chemical free energy perturbation, interaction analysis, locally enhanced sampling, etc.
  - Specific forces specialize from parallel base classes.
- NAMD 3:
  - Isolate as much as possible into "physics layer."
    - Some forces, such as PME, require special methods.
  - Similarly isolate parallelization and integration.
- User-visible benefits:
  - Minimal knowledge required to implement new forces.

Step Four: Front End Separation

MDAPI with dynamic discovery of engine capabilities:



Step Five: New Science

- Implicit solvent models (e.g, generalized Born)
- Replica exchange (e.g., 10 replicas on 16 processors)
- Hybrid quantum/classical mechanics
- Self-consistent polarizability
  - with a (sequential) CPU penalty of less than 100%.
- Fast nonperiodic (and periodic) electrostatics
  - using multiple grid methods.
- A Langevin integrator that permits larger time steps
  - by being exact for constant forces
- An integrator module that computes shadow energy.

NAMD: Scalable Molecular Dynamics

Interdisciplinary Software for the Simulation of Large Biomolecular Systems