NAMD Serial and Parallel Performance

Jim Phillips
Theoretical Biophysics Group
Beckman Institute, UIUC

Serial performance basics

- Main factors affecting serial performance:
  - Molecular system size and composition.
  - Cutoff distance and cycle length.
  - Full electrostatics (PME) parameters.
  - Processor architecture and speed.

System size and density

- Time per step scales
  - linearly with the number of atoms.
  - linearly with density (atoms per volume).

- Example: explicit H vs. united atoms
  - 1/3 the number of atoms.
  - 1/3 the density.
  - Expect 9 X the performance (for protein).

Cutoffs and frequencies

- Time per step scales cubically with cutoff.
- Steps per cycle has a much smaller effect.

- Example: 10 A vs. 14 A
  - 1000 vs. 2744
  - Expect 2.7 X the performance.

PME and grid sizes

- PME with a shorter cutoff breaks even.
- Grid spacing should be about 1 A.
- Grid sizes should be $2^i\ 3^j\ 5^k$. (2^i is best!)
- Only affects FFT (should be minor).

- Example: 47A x 31A x 39A periodic cell
  - Grid sizes: 48 32 40

Processors

- Limited dependence on memory system:
  - Small amount of data.
  - Cache friendly design.
- Performance scales
  - linearly with clock speed (for one architecture).
  - uncertainly with SPEC benchmarks.

- Example:
  - 1333 MHz Althon equals 667 MHz Alpha

Parallel performance basics

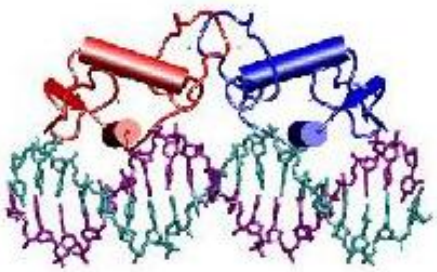
- Speedup = (serial time) / (parallel time).
- Ideal (linear) speedup is number of CPUs.

- Amdahl's law says best speedup possible is (total work) / (non-parallelizable work).

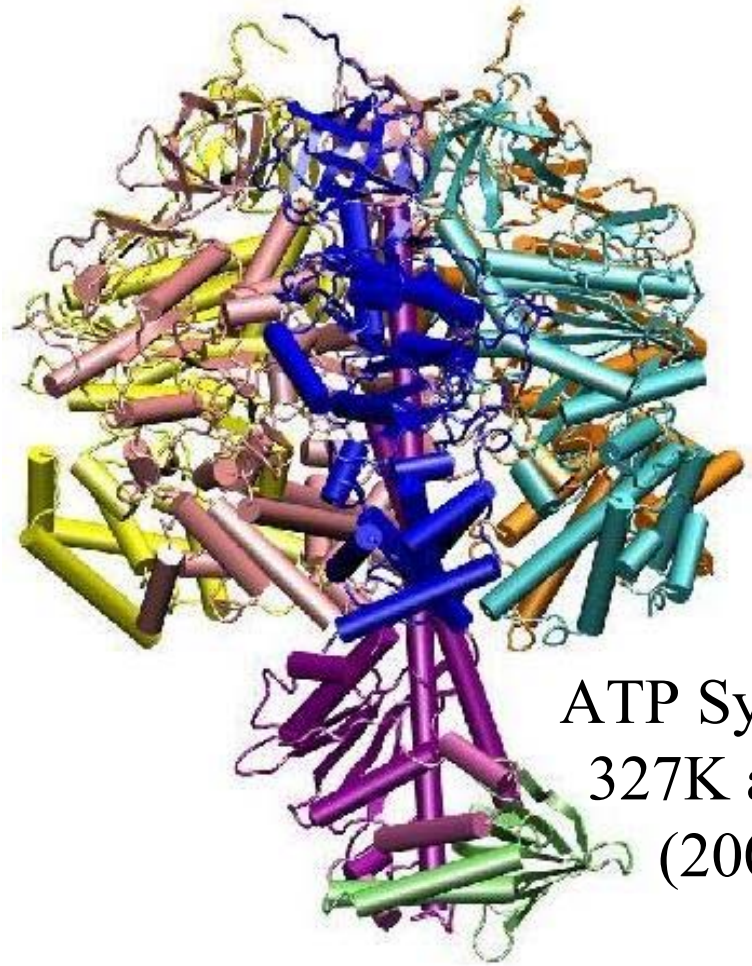
- In MD, each timestep must be parallelized.

Driven by Larger Molecules

BPTI
3K atoms



Estrogen Receptor
36K atoms (1996)



ATP Synthase
327K atoms
(2001)

Driven by Larger Machines

HP 735 cluster
12 processors
(1993)

SGI Origin 2000
128 processors (1997)

PSC LeMieux AlphaServer SC
3000 processors (2002)

NAMD parallel design

- Atoms are “spatially decomposed” into cubes which are distributed among CPUs.
  - Generates 25-400 cubes to distribute.
- Interactions between pairs of cubes are also distributed among CPUs.
  - Generates 300-5000 interaction groups.
- Interactions are redistributed among CPUs based on measurement of their run time.

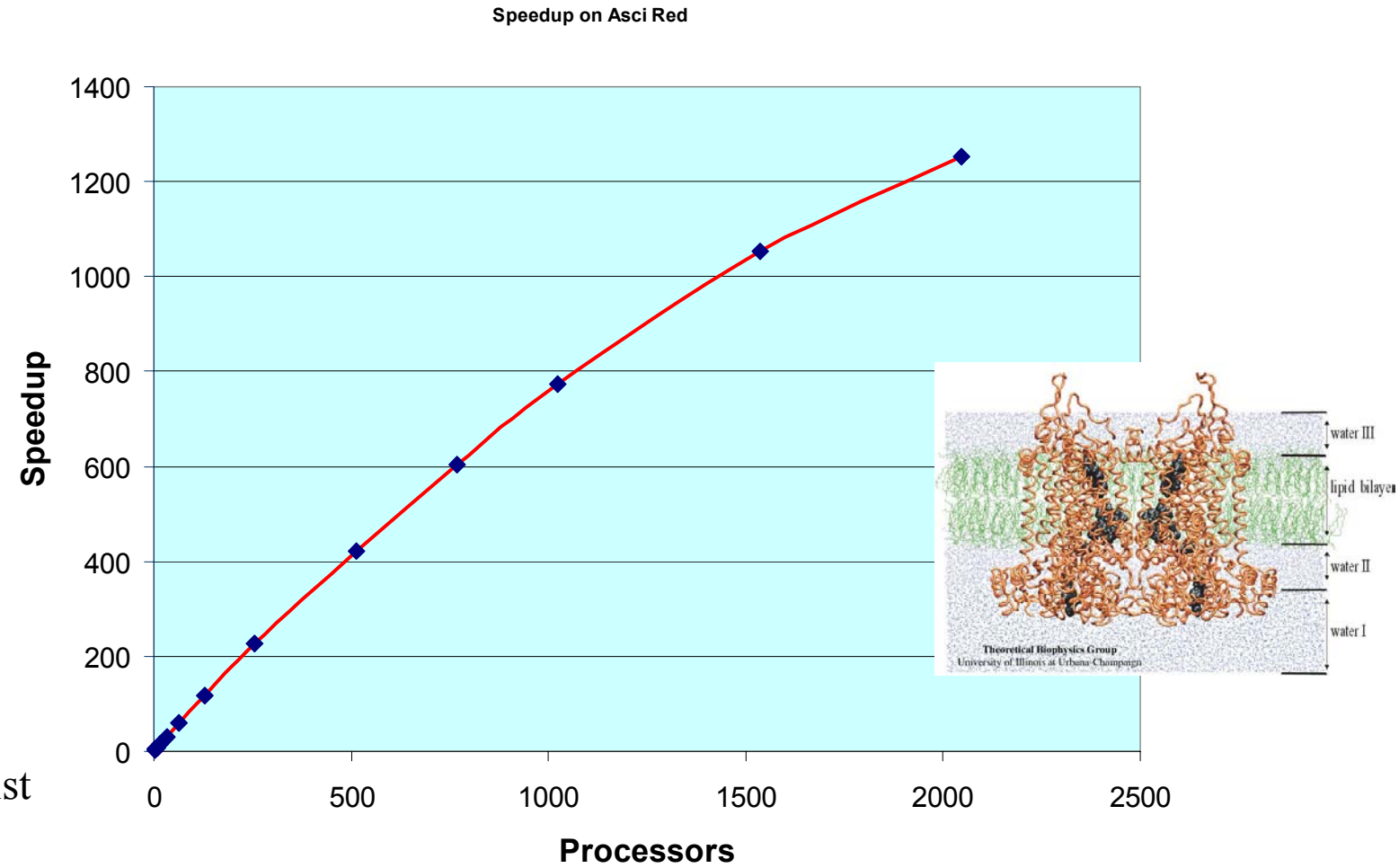
Calculating speedup

- Run 500 steps on 1,2,4,8,… CPUs.
- Final performance, ignore startup time.
- Be aware of dynamic load balancing:
  - Don’t use timing at start of simulation.
  - Don’t use timing including load balancing step.
  - Don’t use “Initial timing: …” value.
  - Best estimate is “Benchmark timing: …”.

Cutoff and system size

- Smaller cutoff may improve efficiency.
  - Data is distibuted to more CPUs.
  - But there is less work to distribute.
- A larger simulation can employ more CPUs.
  - More work to divide among CPUs.
  - More data to distribute as well.
- Length of the simulation doesn't matter.

Best recorded scaling (cutoff)



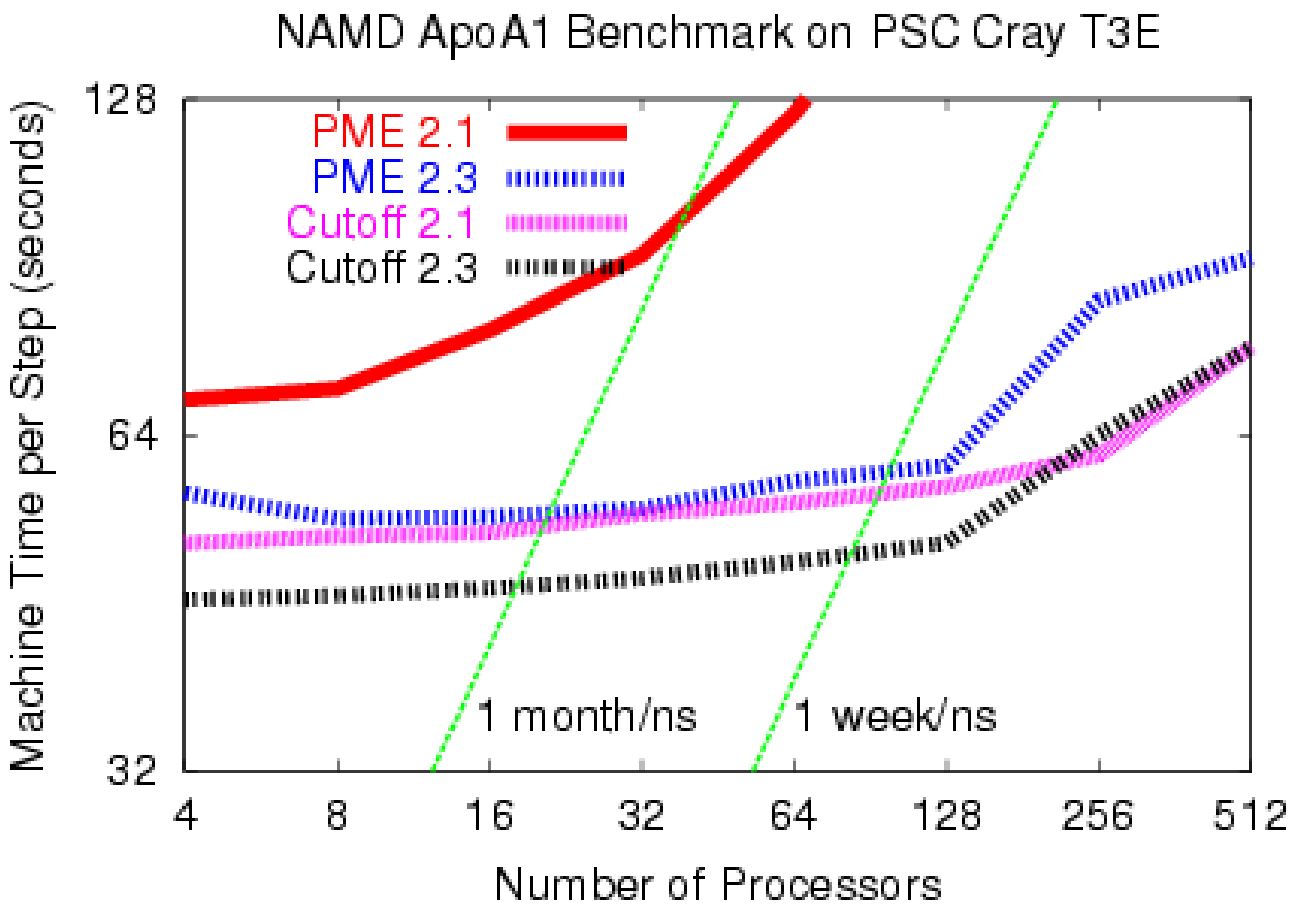
Published in SC2000: Gordon Bell Award Finalist

PME and parallel scaling

- PME is harder to parallelize:
  - More communication.
  - More stages of communication.
  - Only uses number of CPUs equal to grid size.

- But…
  - Superlinear speedup observed on T3E.

PME performance increase



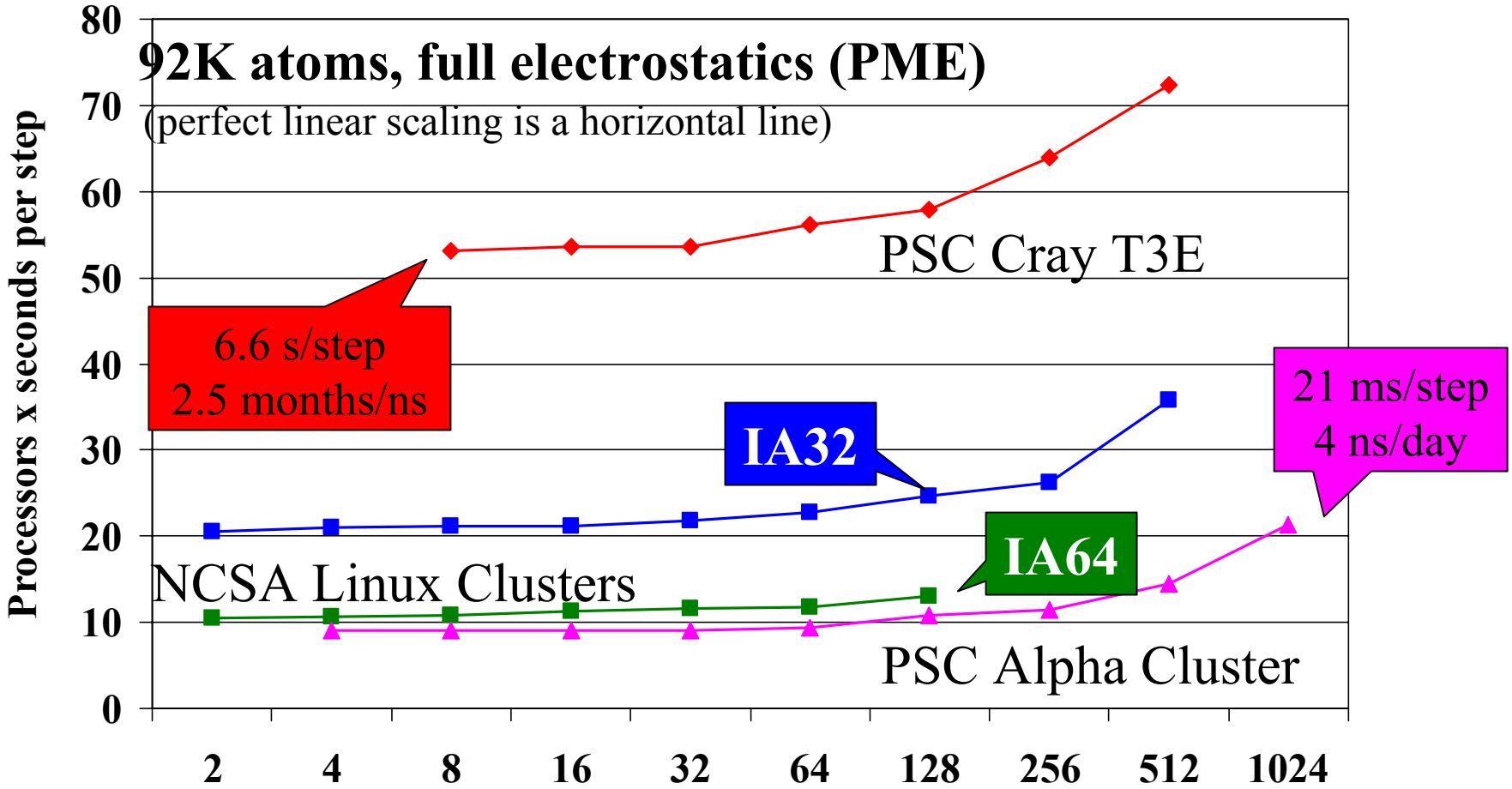
Parallel machines

- Three dimensions to machine performance:
  - Performance of the individual processors.
  - Bandwidth of the communication network.
  - Latency of the communication network.
- CPUs are improving faster than networks.
- Slow CPUs and fast networks give best speedups, but not best price/performance.

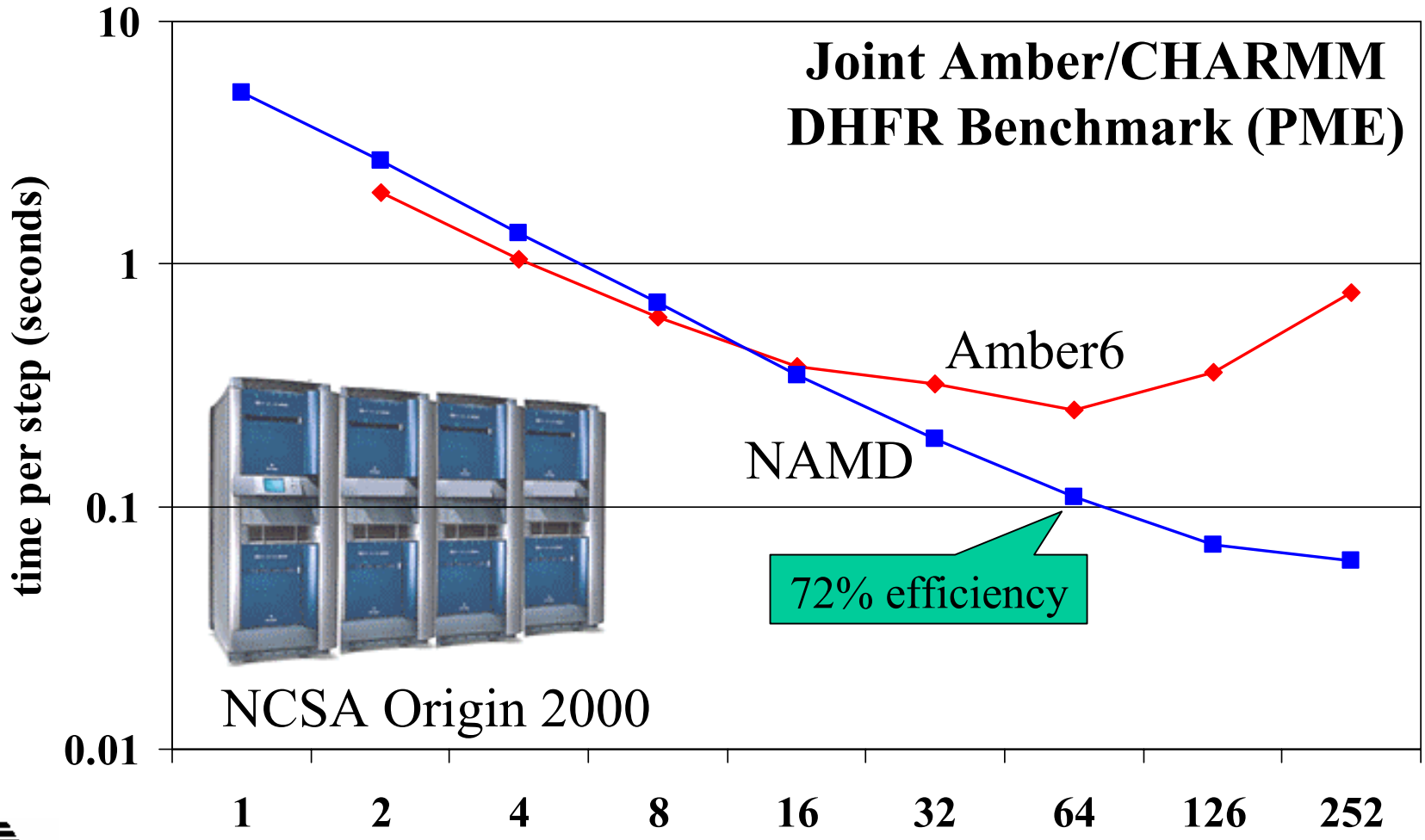
Latency tolerant design

- Some communication is necessary, but we don't want to sit idle waiting for it.
- Work is broken down into interaction sets which require different parts of the incoming data.
- Interactions are calculated as soon as required data is received, not in any particular order.

NAMD 2.4 Performance



NAMD Relative Performance



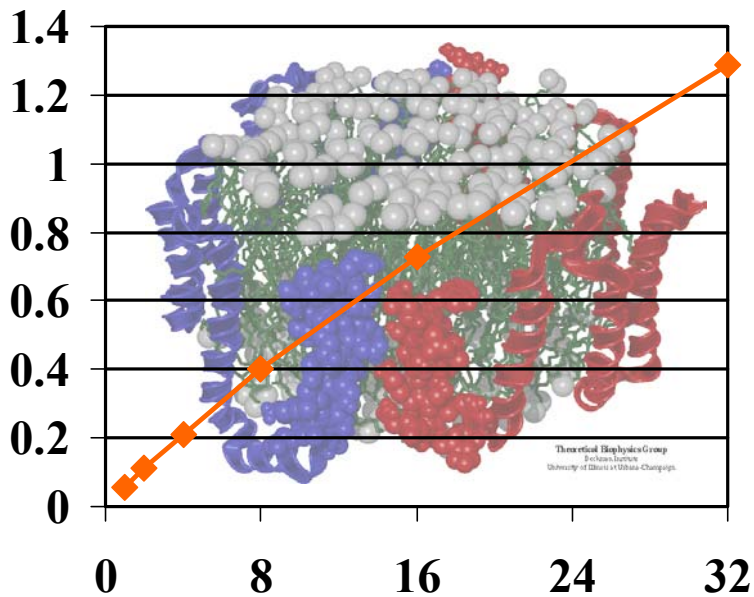
Cost of large simulations

- 100,000 atom simulation with PME.
- 20s per timestep on one 1GHz CPU.
- One nanosecond of simulated time requires:
  - 5000 CPU-hours
  - 8 CPU-months
  - 1 week on 32 CPUs
- You can buy 32 processor machines but they cost $300K (or more, we don't ask).

NAMD 2.4 on Athlon cluster

- 67% efficiency on 32, *commodity hardware*.
- NAMD design is
  - latency tolerant,
  - cache friendly.
- Can simulate 100K atoms at 50 ns/year on 32 CPUs.
- Equivalent to owning a 100 CPU Cray T3E for only $32K.

92K atoms with PME
(ns simulated per week)



NAMD: Scalable Molecular Dynamics