Scaling Behavior of the NAMD Code at AMD's Developer Center

The scaling behavior of the latest released version of the NAMD molecular dynamics simulation code (2.6b1; 30 July 2005) was recently investigated on the AMD Developer Center's Smith cluster. Each of the 32 nodes of this cluster consists of a four-way AMD Opteron 875 (2.2GHz Dual-Core) server, providing eight processor cores per node and a total of 256 64-bit processors for the entire cluster. The nodes are interconnected by three types of network fabric, including Gigabit ethernet, Myrinet, and SilverStorm InfiniBand. Besides these different network architectures, other parameters which could have been tested in the benchmarking process included the compilers used to build the NAMD 2.6b1 binaries; however, due to several technical difficulties in using various other compilers (GNU, Intel, and Pathscale) to build a binary with either Myrinet or InfiniBand support, only the functional Myrinet (MPICH/GM)-enabled NAMD 2.6b1 binary which was built with the Portland Group compilers was used in producing the AMD data presented here.

The remainder of this brief report consists of two figures which provide more detailed information about the benchmarking results (Figs. 1, 2), as well as the hardware features which are most pertinent to the scaling behavior observed for the various clusters (Table 1). Some other important notes and caveats regarding these data are the following:

- All of the non-AMD data – *i.e.* those acquired on the CTBP1, CTBP1m, NBCR, and McCammon clusters – were collected using the previously released version of NAMD (2.5), not the most recent release which was built and tested on the AMD platform (2.6b1).
- All of the simulation data were acquired for exactly the same system under exactly the same conditions (*i.e.*, same point in the MD trajectory), except that the CTBP1m (Myrinet) data points for 10-, 16-, 24-, 40-, and 48-processor runs correspond to different timepoints in the overall trajectory.
- The 61,439-atom system (DNA solvated in explicit water) was simulated under periodic boundary conditions in the N*p*T ensemble, with full electrostatics calculated every second timestep using PME with a 96x96x96-point FFT grid. A 10.0 Å cutoff was used for evaluation of nonbonded interactions.

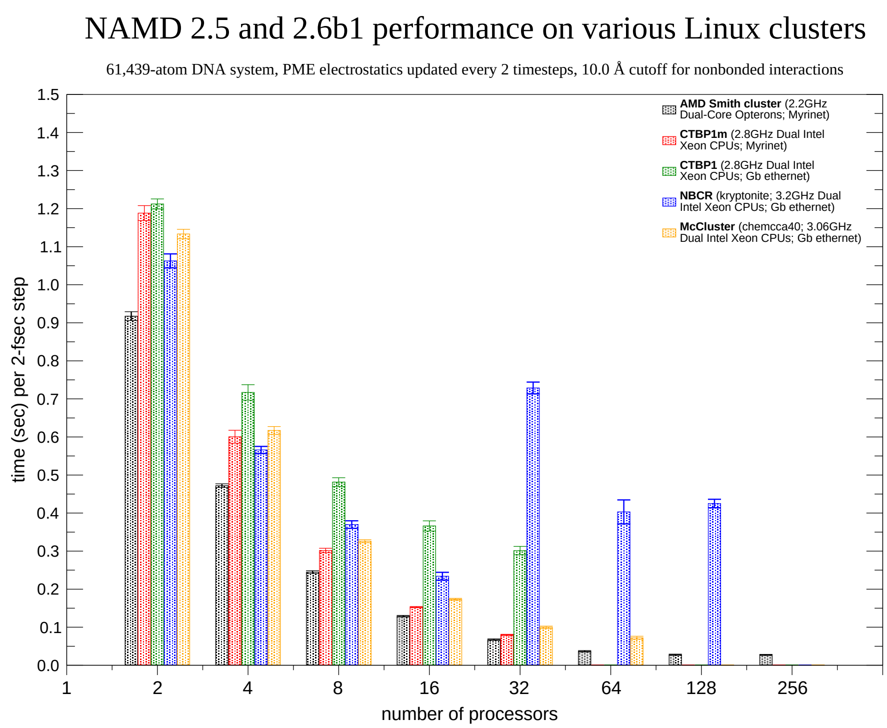

Figure 1: Available data for the performance of NAMD 2.5 and 2.6b1 on various parallel platforms (all Linux clusters) are displayed as the raw wallclock times required to complete a single step of dynamics (2-fsec timesteps). Mean values from (minimally) several hundred timesteps are plotted, as well as the standard deviations of these values (error bars). The reason for the aberrant NBCR datum at 32 processors is unclear, but may arise from the rack-specific scheduling of jobs by the SGE queueing system.

cmura@mccammon.ucsd.edu

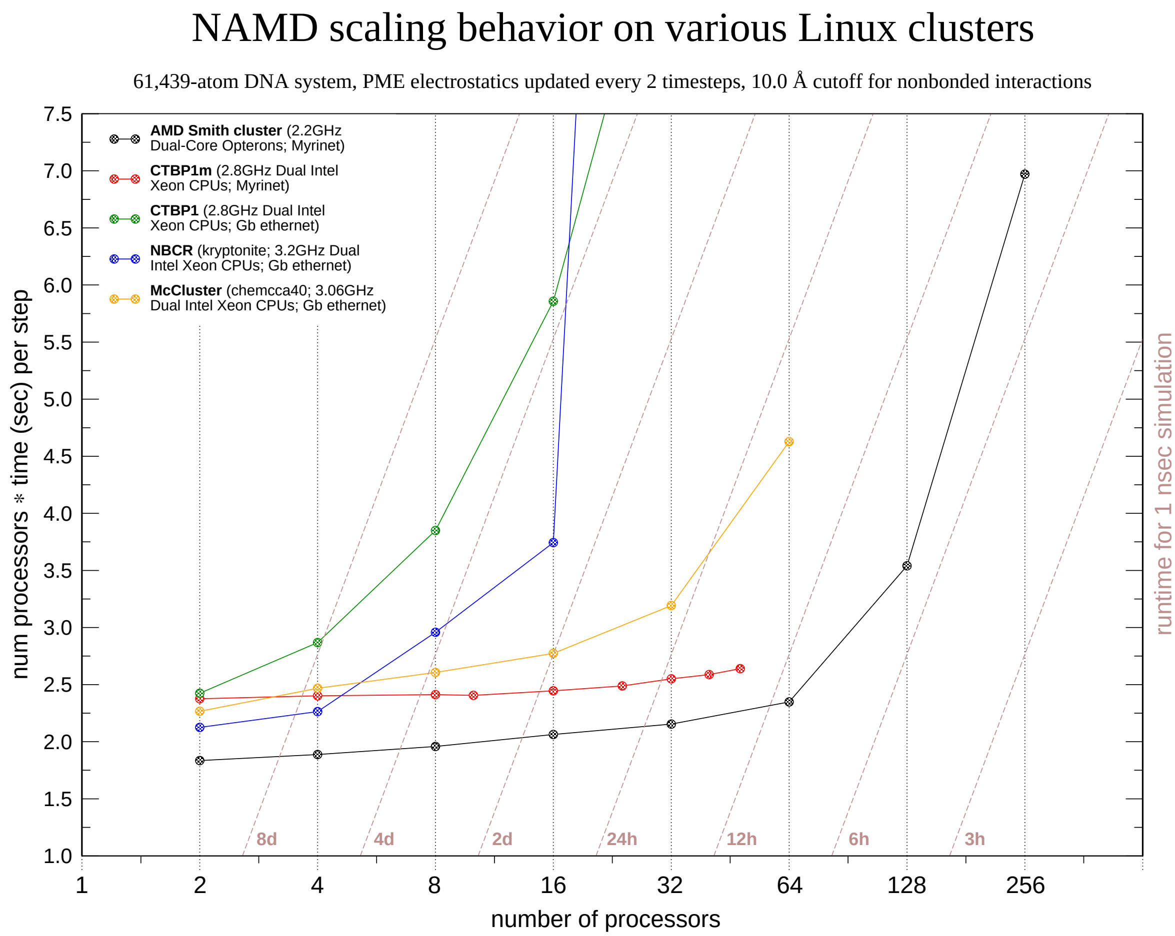

Figure 2: A plot of the scaling behavior of NAMD on various Linux clusters clearly shows that this code scales much better when utilizing Myrinet interconnects (AMD and CTBP1m clusters) rather than the traditional Gigabit ethernet. This is essentially the same data as in Fig. 1, but with the runtime per MD step multiplied by the number of processors used (a horizontal line would represent perfectly efficient scaling). Note that the AMD Smith cluster out-performs all of the others, with the efficiency beginning to degrade between 64 and 128 processors. Also note that the reasonable scaling on the non-Myrinet McCammon cluster is probably due to the network topology of this smaller (32-node) cluster leading to the scheduler deploying jobs on the same rack (fewer number of total racks).

Table 1: The following table lists some hardware characteristics of the Linux clusters used in this report. In particular, those features which are most likely to influence the parallel performance of the NAMD code are listed – *i.e.*, speed and architecture of the CPU as well as the type of network interconnect.

Cluster	Pertinent hardware information (per node)	Network architecture
AMD Smith cluster	4-way AMD Opteron 875 (2.2GHz Dual-Core) servers (8 proc cores per node)	32-node Myrinet network (M3-E64 swtiches)
CTBP1m (Myrinet)	Dual Intel Xeon 2.8 GHz CPUs (Dell PowerEdge 2650)	Myrinet interconnects
CTBP1	Dual Intel Xeon 2.8 GHz CPUs (Dell PowerEdge 2650)	Gigabit ethernet
NBCR (kryptonite)	Dual Intel Xeon 3.2 GHz CPUs (Dell PowerEdge 1425)	Gigabit ethernet
McCluster (chemcca40)	Dual Intel Xeon 3.06 GHz CPUs	Gigabit ethernet

Acknowledgements — This work was performed in the McCammon group at UCSD and was partly funded by a Sloan/DOE postdoctoral fellowship in computational molecular biology.

cmura@mccammon.ucsd.edu