

CROW FOR LARGE SCALE MACROMOLECULAR SIMULATIONS

MILAN HODOŠČEK, URBAN BORŠTNIK
and DUŠANKA JANEŽIČ*

National Institute of Chemistry, Hajdrihova 19, Ljubljana, Slovenia

Abstract: CROW (Columns and Rows Of Workstations - <http://www.sicmm.org/crow/>) is a parallel computer cluster based on the Beowulf (<http://www.beowulf.org/>) idea, modified to support a larger number of processors. Its architecture is based on point-to-point network architecture, which does not require the use of any network switching equipment in the system. Thus, the cost is lower, and there is no degradation in network performance even for a larger number of processors.

Key Words: GNU/Linux Cluster, Parallel Architectures, QM/MM

INTRODUCTION

In this paper, we present some information about medium- to large-sized parallel cluster architecture. Our objective is to build a medium- to large-sized cluster of PCs without expensive network switching equipment. The cost of such network switching equipment can easily end up higher than that of the processors themselves if the cluster has more than around 30 PCs.

MATERIALS AND METHODS

We built clusters with point-to-point connections between processors. A large variety of network topologies are possible within this framework: hypernet (each CPU connected to every other), hypercube, 3-D torus, 2-D MESH, ring, and others.

Due to price constraints, we chose the Debian (<http://debian.org/>) distribution of the GNU/Linux operating system. Our primary applications are mixed quantum and classical potential (QM/MM) methods to study catalytic reactions in enzymes [1]. To successfully connect all the CPUs, we developed a computer program that writes static routing tables for each CPU in the network. The program also draws connection graphs for easier setup, debugging and maintenance of the cluster. It is available on the Internet (<http://www.sicmm.org/crow/>).

To speed up the parallel calculations, we use our own global communication routines on top of MPICH/LAM MPI libraries (<http://www-unix.mcs.anl.gov/mpi/>). We use only the receive/ transmit primitives from the

*Corresponding author: e-mail: dusa@kihp5.cmm.ki.si

standard communication libraries. Efficient global communication routines were developed for each network topology.

RESULTS AND DISCUSSION

Our strategy was based on price/performance: we tested the currently-available CPUs and chose the one which had the greatest price/performance ratio. The timings of the standard CHARMM benchmark for a 1 ps simulation of the MbCO protein surrounded by 3830 water molecules can be found at <http://www.sicmm.org/crow/>.

The comparison of our (CMPI) library global communication routine and the standard is presented in Table 1. Timings (speedups) in hours for Aldose Reductase QM/MM calculations (53 QM atoms (4-31G), 7302 MM atoms, 4 link atoms) performed on a CROW can be found at <http://www.sicmm.org/crow/>.

Tab. 1. The comparison of CMPI library global communication routine and the standard.

Library/Machine	Communication time for 8 CPUs
CMPI/Hypercube	92s
CMPI/Switch	100s
MPICH/Switch	158s

We decided to use a somewhat different approach based on the hypercube parallel topology for our next cluster. Communication in the hypercube is as follows: all the pairs in one dimension exchange $\frac{1}{2}$ of the data. In the next stage, in the other dimension, the amount of data is only $\frac{1}{4}$, and so on for the rest of the dimensions. For a 32 CPU system one can do the following: for the first stage use an internal bus of the dual processor machines; for the second, use relatively cheap 1000 Mbit/s ethernet controllers; and for the other three dimensions, which in total have to exchange a little over 10% of the data, use a 100 Mbit/s ethernet.

REFERENCE

1. Eurenus, K.P., Chatfield, D.C., Brooks, B.R. and Hodošček, M. Studying enzyme mechanism with hybrid quantum mechanical and molecular mechanical potentials. I. Theoretical considerations. **Int. J. Quant. Chem.** 60 (1996) 1189–1200.