2D−5

Parallelization of molecular dynamics program and its performance on VPP500

Kazuhiro Matsumoto Shigeo Orii
Supercomputer Applications Section
Fujitsu Limited

1. Introduction

Molecular dynamics (MD) is known as a method simulating the behavior of molecules which are moved according to Newton equations with preferably decided potential energy function. Following is the general expression of Newton equations with a potential energy function.

$d^2 \vec{r}_i / dt^2 = -\nabla P$ where $\vec{r}_i$ are the coordinates of atoms. Suffix i denotes the atom.

$P = P_{bonded} + P_{non\text{-}bonded(neutral)} + P_{non\text{-}bonded(charged)}$

In the programs of molecular dynamics, simulations are done going through several calculation stages. In the case of famous MD program AMBER, they are as follows [1].

Stage.1) Pair list generation for non-bonded pairs
Stage.2) Force calculation for non-bonded and bonded pairs
Stage.3) Scaling of coordinates
Stage.4) Time integration and SHAKE processing

2. Program description

Stage.2) explained in the previous chapter is one of the most time-consuming parts of MD programs. We made a sample program testing a parallel implementation on the force calculation for non-bonded pairs in this stage. We take the following potential energy function.

$P(r_{ij}) = 4*((1/r_{ij})**12 - (1/r_{ij})**6) - Pcut \quad (r_{ij} \leqq Rcut)$
$P(r_{ij}) = 0 \quad (r_{ij} > Rcut)$

where r_{ij} is a distance between the atoms i and j. The unit cell length L is given as $(1/\rho)**1/3$. We set Rcut L/2. Finally Pcut is calculated as $4*((1/Rcut)**12-(1/Rcut)**6)$. We simulate the model with 8000 same atoms, the number density ρ is 0.5, and the maximum time iteration number is 100.

3. Specific features of VPP500

We summarize the specific features of VPP500 [2].

- The VPP500 is a highly parallel distributed memory system.
- A crossbar network interconnects up to 222 PEs.
- Each PE has the vector performance of 1.6 GFlops.
- FORTRAN77 EX/VPP compiler presents programmers a data-parallel shared memory paradigm on the distributed local memory architecture.

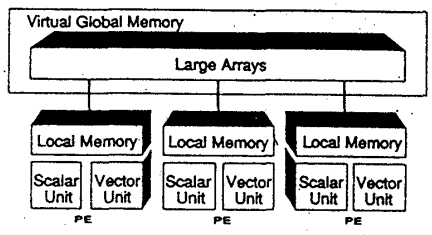


Fig.1. Configuration of VPP500

4. Parallelization and its performance

Parallelization is done with a particle decomposition method and its application to the program is explained as follows.

((Data decomposition))	Data of the coordinates of atoms are used as replicated local arrays.
((Procedure decomposition))	Outer loop of 2-nested DO loops for force calculations is parallelized in CYCLIC.
((Data transfer))	Data segments of forces and potential energy calculated on each PE are transferred to all other PEs.

We will show how to program for parallelization with FORTRAN77 EX/VPP.

Fujitsu Limited, 9-3, Nakase 1-Chome, Mihama-Ku, Chiba-Shi, Chiba 261, Japan

Comparison of the parallelized code with the original one

```
<< Original code >>
      ::
    parameter (n=8000)
      ::
      ::
      ::
    do 10 i=1,n
      ::
    do 20 j=i+1,n
```

Calculation of forces and potential energy

```
  20 continue
      ::
  10 continue
```

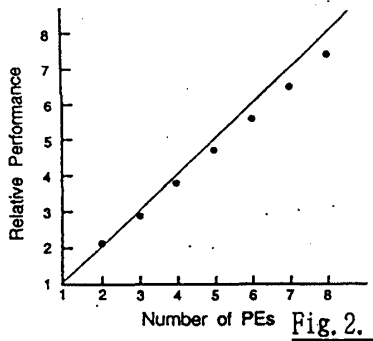
```
<< Parallelized code >>
!XOCL PROCESSOR P(4)
      parameter (n=8000)
!XOCL INDEX PARTITION Q=(P, INDEX=n, PART=CYCLIC)
        ::
!XOCL SPREAD DO /Q
      do 10 i=1,n
        ::
      do 20 j=i+1,n
```

Calculation of forces and potential energy

```
   20 continue
       ::
   10 continue
!XOCL END SPREAD
```

Data transfer of forces and potential energy and summation of them

We will show the parallel performance below.



Relative performance is given as the parallel execution time on 1 PE devided by that on N PEs where N is the number of PEs. Solid line shows the ideal performance.

Fig. 2. Parallelization performance

4. Discussion

- We discuss the 2-nested DO loops which are parallelized. The calculation of inner DO loop is done in the range for j to run from i+1 to n. The computational amount of the inner DO loop decreases as i increases. Parallelization where computational segment with respect to i is allocated cyclicly to PEs is suitable in this case. We obtain homogeneous granularity by this parallelization.
- Data transfer from one PE to all other PEs is needed just after the calculation of the parallelized DO loops at each time step. This may not cause significant difficulties, because VPP500 has crossbar network which realizes homogeneous and high-speed data transfer. In the case using 8 PEs, time for data transfer occupies as small as 4 % of the total elapsed time.
- The computational amount of the 2-nested DO loops is characterized at the order of (the number of atoms)**2. That of the rests is characterized at the order of (the number of atoms). In the case where the number of atoms is 8000, we can focus our parallelization efforts on the 2-nested DO loops above.

With these reasons we have achieved the scalable performance on VPP500 about the parallelization of a MD program.

5. Acknowledgement

Authors thank Dr. R. Nobes of the Autralian National University for fruitful discussion especially on the matter of computational chemistry.

References:

[1] H. Sato, Y. Tanaka, H. Iwama, S. Kawakita, M. Saito, K, Saito, K. Morikami, T. Yao and S. Tatsumi, Proc. of IEEE Scalable High Performance Computing '92 conference, (April 1992) 113-120

[2] K. Miura, M. Takamura, Y. Sakamoto, and S. Okada, Comp. Con. Spring 93 (February 1993) 128-130