

PC クラスタを用いた *ab initio* 結晶軌道法の並列処理

寺前 裕之*, 大田原 一成

ATR 適応コミュニケーション研究所, 〒 619-0288 京都府相楽郡精華町光台 2 丁目 2-2

*e-mail: teramae@atr.co.jp

(Received: November 26, 2001; Accepted for publication: December 19, 2001; Published on Web: February 23, 2002)

メッセージパッセンジャーとして MPI(Message Passing Interface) の一種である LAM(Local Area Multicomputer) を用いた、並列計算機環境の構築について述べる。実際の計算としては、一次元ポリマーのエネルギーバンド計算の並列処理について示した。PC としては、Pentium III デュアル CPU 仕様のマザーボードを使用したパーソナルコンピュータ 8 台 (CPU 数は 16) からなる PC を 100BaseT の 8 ポートスイッチを使用してネットワーク接続を行い、クラスタ構成とした。Poly-tetrafluoro-ethylene (C_2F_4)_x の計算では 8CPU 使用時で実時間が 6.27 倍に加速された。Poly-(para-phenylene sulfide)($C_6H_4SC_6H_4S$)_x の計算ではクラスタ全体のメモリーに二電子積分を全て置く事が可能となったために、8CPU 使用時に約 20 倍という実時間の高速化が達成され、実時間の短縮という面においても並列処理は非常に有効であることがわかった。

キーワード : Parallel processing, Crystal orbital, Energy band, PC cluster

1 はじめに

Ab initio 結晶軌道法による一次元ポリマーのエネルギーバンド計算は *ab initio* 分子軌道法の無限系への自然な拡張であり、過去にポリアセチレンなどの導電性高分子の電子状態研究において、異性体間の熱力学的安定性や振動数解析などでその有用性が示された [1–3]。ただし *ab initio* 分子軌道法では、計算すべき二電子積分の数は基底関数の数 n の 4 乗にのみ比例するが、*ab initio* 結晶軌道法では、 n の 4 乗だけではなく考慮する隣接セル数 N の 3 乗にも比例し、 $N^3 n^4$ に比例することになる。このことからわかるように CPU 占有時間およびファイルの入出力共に非常に大規模な計算となる。そのため単位セルが小さなポリマー以外の計算はあまりなされていないのが現状である。

そこで我々は、この一連の論文の第一報、第二報において、一次元ポリマーのエネルギーバンド計算の並列化を試み、CPU 占有時間に関しては 16 並列時に約 10 倍程度の性能が得られる事を示し、並列化がエネルギーバンド計算において大変有効である事を示した

[4, 5]。ただしこれらの研究においては、実際に多数の CPU を使用した計算機環境を得る事ができなかったために 2CPU または 4CPU によるシミュレーションに止まり、CPU 占有時間以上に重要である実時間の短縮に関してのデータは不十分なままであった。

前報 [5] では、大規模なポリマーのエネルギーバンド計算として poly-(para-phenylene sulfide) ポリマーについて実際の計算にかかった実時間を比較した。IBM 社製のワークステーション RS6000/590 を 10BaseT イーサネットを介してネットワーク接続し計算に必要であった実時間が計測された。その結果、実時間に関しては、ネットワークの構造や各 CPU の接続形態など、また計算時のネットワークの混雑の程度など、多くの因子が関係する事が示唆された。同一のハブに 2 台の WS を収容した場合には、実時間が 1.74 倍に加速されたが、別のサブネットにある 2 台で実行した場合には、途中のネットワークの状況などに大きく依存するようになり、ほとんど実時間の減少に結び付かない場合があることもわかった。

従って、計算機クラスタを用いた計算には同一の

ハブに接続するなどの専用のクラスターシステムとしてデザインする必要があると考えられる。ただしワークステーション (WS) クラスターで CPU を 10-20 台の単位で専用に確保することは以前と比べて価格が下がってきたとは言え、コスト的に見て難しい。一方で比較的安価な Intel 製 CPU を使用したパーソナルコンピュータ (PC) は近年の性能向上がめざましく、WS との差が縮小してきた。例えば Pentium II 450MHz の Linpack 実測値 (n=100) では浮動小数点演算速度は最高約 90Mflops にも達し、やや古いモデルとの比較ではあるが RS6000/590 の 35Mflops や Cray T932 の単一 CPU での 90Mflops と比較してもむしろ高速な値となっている。また近年、Linux や FreeBSD のような PC で動作するフリーな Unix が普及しはじめたため、PC を事実上 WS のように使用する事が可能になり、WS から最小限のプログラム書き換えで PC への移行が行えるようになった。PC クラスターは分子軌道計算プログラム Gaussian 98 の実行においても有効であることが知られている [6]。

そこで本研究では、Intel 製 CPU 使用のパーソナルコンピュータ (PC) とフリーな UNIX である FreeBSD を使用した PC クラスターによって並列計算専用の環境の構築を行ったので報告する。

2 並列計算環境の構築

本研究を計画した 2000 年当時では、Intel 社製 440BX チップセットによるマザーボード及び Pentium III プロセッサが主流であった。最高速 CPU は Pentium III の 600MHz であったが、CPU としては対価格性能比も考慮して Pentium III の 450MHz を採用した。マザーボードは前報でも使用した Dual CPU マザーボード、ASUS 社製の P2B-D を使用し、8 筐体で 16CPU の構成とした。Figure 1 にここで構築したシステムの外観を示した。デュアル CPU マザーボードを使用したのは主として設置スペースの問題からである。なお当初搭載されていた BIOS ではデュアル CPU での使用時に障害があったために、アップデートユーティリティを使用して BXD112B.AWD にバージョンアップを行った [7]。ディスプレイ・キーボード・マウスはそれぞれ 4 ポートの切り替えスイッチを使用した。カスケード接続により各 1 台とすることも可能であったが、メンテナンスの容易さも考慮して各 2 台とした。



Figure 1. PC クラスターシステムの外観

当時はメモリーも未だ高価であったため、一筐体当たり 256 メガバイトとした。またハードディスクは一筐体当たり 12 ギガバイトの IDE 接続の物を使用した。CD-ROM ドライブおよびフロッピーディスクドライブは省略しても差し支えないが、OS インストールの容易性から今回は装備した。ファストイーサネット (100BaseT) のカード及び 8 ポートのスイッチングハブを使用して各 PC 間では比較的高速な通信ができるようにネットワーク接続した。

オペレーティングシステムは、FreeBSD 3.2-RELEASE を使用した [8]。デュアル CPU 用にカーネルに SMP オプションを付加して再コンパイルを行った。並列計算を行なうためのライブラリとしては、MPI(Message Passing Interface) の実装の一つである LAM(Local Area Multicomputer) version6.2β[9] を用いた。Fortran コンパイラーは egcs-1.1.2 の g77 を使用した [10]。

3 計算方法

ポリマーのエネルギーバンドの計算理論については文献に詳しいので並列計算に必要な部分のみの記述にとどめる [1-3]。一次元ポリマーの計算理論である結晶軌道法は有限の分子系における分子軌道法を無限系のバンド計算に拡張したものである。

Hartree-Fock 方程式は、

$$\varepsilon(\mathbf{k})\mathbf{S}(\mathbf{k})\mathbf{C}(\mathbf{k}) = \mathbf{F}(\mathbf{k})\mathbf{C}(\mathbf{k}) \quad (1)$$

である。 $F_{rs}(\mathbf{k})$, $H_{rs}(\mathbf{k})$, $S_{rs}(\mathbf{k})$ は、 \mathbf{k} をあらわに含まない実空間での各行列要素のフーリエ変換で表すことができ、

$$F_{rs}(\mathbf{k}) = \sum_{j=-N}^N \exp(i\mathbf{k}\mathbf{a}j) F_{rs}^{0j} \quad (2)$$

$$H_{rs}(\mathbf{k}) = \sum_{j=-N}^N \exp(i\mathbf{k}\mathbf{a}j) H_{rs}^{0j} \quad (3)$$

$$S_{rs}(\mathbf{k}) = \sum_{j=-N}^N \exp(i\mathbf{k}\mathbf{a}j) S_{rs}^{0j} \quad (4)$$

ここで、

$$S_{rs}^{0j} = \int \chi_r^0 \chi_s^j d\mathbf{r} \equiv \langle r^0 | s^j \rangle \quad (5)$$

$$H_{rs}^{0j} = -\frac{1}{2} \langle r^0 | \Delta | s^j \rangle - \sum_{h=-N}^N \sum_A^{atom} \langle r^0 | \frac{Z_A}{|\mathbf{r} - h\mathbf{a} - \mathbf{R}_A|} | s^j \rangle \quad (6)$$

$$F_{rs}^{0j} = H_{rs}^{0j} + \sum_{h=-N}^N \sum_{l=-N}^N \sum_t^n \sum_u^n P_{tu}^{hl} \{ 2 \langle r^0 s^j | t^h u^l \rangle - \langle r^0 t^h | s^j u^l \rangle \} \quad (7)$$

$$P_{tu}^{hl} = \frac{a}{\pi} \int_{\mathbf{BZ}}^{occupied} \sum_n \exp\{-i\mathbf{k}\mathbf{a}(h-l)\} C_{tn}^*(\mathbf{k}) C_{un}(\mathbf{k}) d\mathbf{k} \quad (8)$$

これらの関係式からユニットセル当りの全エネルギーは、

$$\frac{E_{total}}{N} = -\frac{1}{2} \sum_{j=-N}^N \sum_r^n \sum_s^n (H_{rs}^{0j} + F_{rs}^{0j}) P_{rs}^{0j} + \frac{1}{2} \sum_{h=-N}^N \sum_A^{atom} \sum_B^{atom} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B - j\mathbf{a}|} \quad (9)$$

ここで、 n は基底関数の数を N は考慮する隣接セル数を表す。式 (7) より容易にわかるように、二電子積分の数は $N^3 n^4$ に比例する。Fock の行列要素を計算するのに必要な電子密度行列は結晶軌道の係数 $C_{tn}(\mathbf{k})$ から計算されるが、結晶軌道の係数は変分方程式を解かないと得ることができない。従って、分子軌道計算の場合と全く同様に SCF 計算が必要である。

Figure 2 に示したように、二電子積分および二電子積分の核座標に関する微分の計算において隣接セル数である N および基底関数の数 n を用いた並列化が行えるため通常分子軌道計算に比べて処理の一層の分散が図れ得る。二電子積分および二電子積分の核座標に関する微分が計算の全体に占める割合はテストに用いた poly-tetrafluoro-ethylene で約 75% であるが、結果として SCF 計算部分も並列処理されるため、並列化される計算量は 90% 以上になる。

```

...
icount=0
DO J= 0,N
  DO K=-N+J,N
    DO L=-N+J,N
      DO R=1,n
        DO S=1,n
          DO T=1,n
            DO U=1,n
              icount=icount+1
              if(mod(icount,nproc).eq.me) then
                Calculate <R(0)S(J) | T(K)U(L)>
                or <R'(0)S(J) | T(K)U(L)>
              endif
            ENDDO
          ENDDO
        ENDDO
      ENDDO
    ENDDO
  ENDDO
ENDDO
...

```

Figure 2. 二電子積分およびその微分の並列計算の概要

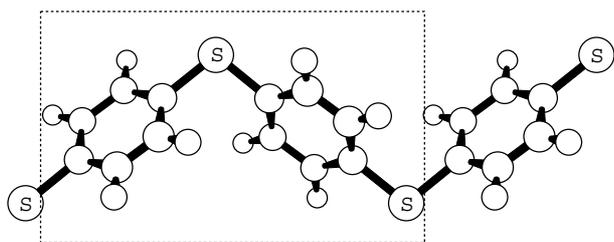


Figure 3. Poly-(para-phenylene sulfide) の構造、破線内がユニットセル

速度比較のために前報と同様に poly-tetrafluoro-ethylene (C_2F_4)_x を対象に選んだ。基底関数系は STO-3G 基底関数系 [11] を用いた ($n = 30$)。隣接セル数 N は 5 として、二電子積分のカットオフ法には Namur cut-off 法を用いた [12]。さらに大規模な計算として poly-(para-phenylene sulfide) ($C_6H_4SC_6H_4S$)_x を対象に選んだ。基底関数系は 3-21G 基底関数系 [13] を使用し ($n = 150$)、隣接セル数 N は 3、二電子積分のカットオフ法には cell wise 法を用いた [12]。波数ベクトルのサンプリングは前者は 41 点後者は 21 点で行ない、Simpson の公式を用いて電子密度の積分を行なった [14]。Figure 3 に poly-(para-phenylene sulfide) の構造を示した。交互に 90 度傾いたベンゼン環が硫黄原子を挟んで連なった構造をしており、電導性があることで知られている [15]。計算時間の計測については各 10 回行い最速値をとった。

4 結果と考察

Table 1 に poly-tetrafluoro-ethylene (C_2F_4)_x の計算において計測された CPU 占有時間および実経過時間(各々秒単位)を示す。CPU 数が 4CPU までは順調に CPU 時間と同様に実時間も加速されていくことがわかり、CPU 時間のみならず実時間の減少にも並列化は大きな効果があることがわかる。8CPU 程度でやや実時間の加速が CPU 時間と分離し始め、8CPU から 16CPU では 6.27 倍が 6.96 倍と僅かな増加にとどまり、CPU 数が増加した効果がほとんど得られていない事がわかる。これは現在のシステムがデュアル CPU マザーボードを使用しているために、ハードディスクやネットワーク通信の入出力が 8CPU までは独立だが 16CPU では各筐体で 2CPU が共用する形になっているのが影響していると考えられる。同一マザーボード上の CPU 間

では共有メモリーを介して通信が行われるため、各々の効果が相殺されるのを期待していたが外部への入出力の負荷の方が高いようである。

Table 1. Poly-tetrafluoro-ethylene の並列計算の速度比較。8CPU までは各 PC で 1CPU に限定。

CPU 数	CPU 時間 (秒)	CPU 時間加速率	実時間 (秒)	実時間加速率
1	182.8	1.00	197.0	1.00
2	93.9	1.95	102.8	1.92
4	45.9	3.98	51.6	3.82
8	24.4	7.49	31.4	6.27
16	15.8	11.60	28.3	6.96

それを確かめるために、同一マザーボード上の 2CPU を同時に使用するように設定して計測した結果を Table 2 に示す。ここで例えば 4CPU の場合は 2 台の PC にある 4CPU を使用している。Table 1 では、4CPU の場合は 4 台の PC を使用して計測した。CPU 数が 2 の場合のみ加速率が理論最高値である 2 倍以上になっているが、これは同一マザーボード上に CPU があり SMP 構成として動作しているため、CPU 間の通信もメモリーのみを介して行われるなど、並列化によるオーバーヘッドがほとんど生じない上に、カーネルの動作も多少は異なってくるためと思われる。また時間計測のルーチンの誤差も考えられる。

Table 2. Poly-tetrafluoro-ethylene の並列計算の速度比較。2CPU 以上では各 PC で 2CPU 使用。

CPU 数	CPU 時間 (秒)	CPU 時間加速率	実時間 (秒)	実時間加速率
1	182.8	1.00	197.0	1.00
2	88.8	2.06	97.0	2.03
4	51.3	3.56	57.9	3.40
8	28.4	6.44	34.2	5.76
16	15.8	11.60	28.3	6.96

CPU 数が 4 または 8 の場合は、同じマザーボード上の CPU を使用しなかった Table 1 に比べて明らかに CPU 時間・実時間共に増加しており加速率は落ちている。従って、Table 1 における 16CPU 使用時にパフォーマンスが上がらなかったのは、デュアル CPU マザーボードの使用が原因であったと考えられよう。共有メモリー型の計算機をさらにクラスター構成とするのは避けるべきと思われる。現在ではラックマウントタイプの筐体の PC が発売されているため、省スペース

スを目的としてデュアル CPU マザーボードを使用する必要性も薄れてきていることを付け加えておく。

Table 3 に poly-(para-phenylene sulfide) ($C_6H_4SC_6H_4S$)_x の計算において計測された CPU 占有時間および実経過時間 (各々秒単位) を示す。CPU 数が 2 および 4 まででは、CPU 時間・実時間共に順調に加速されており計算のみならずファイルの入出力も自然に並列化されており、実時間は CPU 時間の約 3 倍程度となっている事がわかる。これはローカルにハードディスクやメモリーを持つクラスター型の並列計算機に特徴的なことで、共有メモリー型や高速なスカラ計算機においては CPU 時間の短縮と実時間の短縮は必ずしも一致しない。特にスカラ計算機で高速な CPU に取り替えてもハードディスクが遅いため実時間の短縮に至らないという事はよくある事例である。

Table 3. Poly-(para-phenylene sulfide) の並列計算の速度比較

CPU 数	CPU 時間 (秒)	CPU 時間加速率	実時間 (秒)	実時間加速率
1	5330.7	1.00	15341.0	1.00
2	2741.5	1.94	7787.9	1.97
4	1477.3	3.61	4088.4	3.75
8	641.6	8.31	784.8	19.55
16	462.2	11.53	764.2	20.07

CPU 数が 8 になったところで、CPU 時間・実時間共に非連続的に加速されている事がわかる。特に実時間は 8CPU 時に約 20 倍という加速率が得られている。これは二電子積分のファイル容量が約 2 ギガバイトで、一台あたり 256 メガバイトのメモリーを搭載しているため、ちょうど 8 台使用時に全てがメモリー上にバッファされファイル入出力がほぼ無くなったためと考えられる。OS として UNIX を使用する事により、有効にメモリーが使用され、プログラムの書き換え無しに実時間の減少に結び付いたことがわかる。また 8CPU 使用時には CPU 時間も加速率は理論値である 8 倍より大きくなっているが、これもファイル入出力に関するオーバーヘッドが無くなったためと考えられる。この例でも、8CPU 使用時から 16CPU 使用時への実時間の短縮は CPU 使用時間の短縮に比べて残念ながら小さい。前に述べたのと同じ事情に依ると思われる。なお、Table 1 と Table 2 の poly-tetrafluoro-ethylene の計算ではファイルが小さいため、最初から全てのファイルがメモリー上にバッファされている事を付け加

えておく。

以上のように CPU 使用時間のみならず実時間の減少という点でも並列処理が有効であることが示された。

5 結論

本研究では Pentium III デュアル CPU 仕様のマザーボードを使用したパーソナルコンピューター 8 台 (CPU 数は 16) からなる PC を 100BaseT イーサネット・スイッチングハブを使用してクラスター構成としたものを並列計算機として、一次元ポリマーのエネルギーバンド計算の並列処理計算を行った。CPU 時間ではなく、より重要な実時間がどの程度短縮されるかを計測するのを目的とした。

Poly-tetrafluoro-ethylene (C_2F_4)_x の計算では 8CPU 使用時に実時間が 6.27 倍に加速された。ただし 16CPU 使用時には 6.96 倍に止まった。デュアル CPU マザーボードを使用したため I/O 待ちや通信待ちが生じたためと思われる。

Poly-(para-phenylene sulfide) ($C_6H_4SC_6H_4S$)_x の計算ではクラスター全体のメモリーに二電子積分を全て置く事が可能となったために、8CPU 使用時に約 20 倍という実時間の高速化が達成され、実時間の短縮という面においても並列処理は非常に有効であることがわかった。

今後はさらに多数の PC を用いて並列処理を行った場合に、どの程度までの実時間の短縮につながるかは興味深いが多数つなぐとネットワーク待ち時間が増大する問題が出てくる。そうなると 1000BaseT やミリネットのようなより高速な通信インターフェースが必要かどうか興味深く、今後の課題であろう。また本研究を行った時点に比べてメモリーの価格が下落してきているため、なるべく多くのメモリーを搭載することにより、本研究で述べたような実時間の短縮がより低い並列度でも可能となると予想され興味深い。

参考文献

- [1] M. Kertesz, *Adv. Quantum Chem.*, **15**, 161 (1982).
- [2] P. Otto, E. Clementi, and J. Ladik, *J. Chem. Phys.*, **78**, 4547 (1983).
- [3] H. Teramae, *J. Chem. Phys.*, **85**, 990 (1986).
- [4] H. Teramae, *J. Chem. Software*, **4**, 73 (1998).

- [5] H. Teramae, *J. Chem. Software*, **6**, 75 (2000).
- [6] 坂本章, 北嶋一欽, 田隅三生, *CACS FORUM*, **20**, 135 (2000).
- [7] BIOSアップデートユーティリティ並びにデータは <http://www.asus.com.tw/> よりダウンロードした。
- [8] <http://www.freebsd.org/>
- [9] Gregory D. Burns, Raja B. Daoud, James R. Vaigl, Supercomputing Symposium '94 (Toronto, Canada, June 1994); Greg Burns, Raja Daoud, MPI Developers Conference, University of Notre Dame (June 1995)
- [10] <http://gcc.gnu.org/>
<ftp://go.cygnum.com/pub/ftp.cygnum.com/egcs/releases/egcs-1.1.2/>
- [11] W. J. Hehre, R. Ditchfield, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.*, **52**, 2769 (1970).
- [12] H. Teramae, *Theoret. Chim. Acta*, **94**, 311 (1996).
- [13] J. S. Binkley, J. A. Pople, W. J. Hehre, *J. Am. Chem. Soc.*, **102**, 939 (1980).
M. S. Gordon, J. S. Binkley, J. A. Pople, W. J. Pietro, W. J. Hehre, *J. Am. Chem. Soc.*, **106**, 2797 (1984).
- [14] P. J. Davis and P. Rabinowitz, *Method of Numerical Integration*, Academic Press, New York (1975), p.45.
- [15] M. Knupfer, J. Fink, E. Zojer, G. Leising, U. Scherf, K. Mullen, *Phys. Rev.*, **B57**, 4202 (1998).

Study on Parallel Processing of *ab initio* Crystal Orbital Calculations Using PC Cluster

Hiroyuki TERAMAE* and Kazushige OHTAWARA

ATR Adaptive Communications Research Laboratories
2-2-2 Hikaridai, Seika-cho Soraku-gun Kyoto 619-0288, Japan

*e-mail: teramae@atr.co.jp

We study the parallel processing environment using the MPI/LAM message passenger. We perform *ab initio* crystal orbital calculations of one-dimensional polymers using parallel processing. The personal computer cluster of eight dual CPU motherboards with Intel Pentium III 450MHz (total 16 CPU) is connected to a 100BaseT ethernet switch. The calculation on poly-tetrafluoro-ethylene (C₂F₄)_x shows that the wall clock by 8 CPU is 6.27 times faster than that by the single CPU. The wall clock by 16 CPU, however, remains just 6.96 times faster than that by single CPU, which should be attributed to the inefficiency of the dual CPU motherboard on the I/O wait time and the network wait time. The calculation on poly-(para-phenylene sulfide) (C₆H₄SC₆H₄S)_x shows that the wall clock by 8 CPU becomes about 20 times faster than that by the single CPU. This is enabled by the fact that the total system memory becomes greater than the temporary file of two electron integrals, which eliminates the access to the hard disk and I/O wait time. These results show the parallel processing of the electronic structure calculations on the one-dimensional polymers to be very effective in terms of the wall clock.

Keywords: Parallel processing, Crystal orbital, Energy band, PC cluster