

## 分子科学計算におけるワークステーションと並列計算機の可能性

長嶋雲兵、青柳 睦\*

お茶の水女子大学、化学技術研究所\*

分子科学計算に広く用いられているSCF計算プログラムをいくつかのワークステーションで実行しその性能を調査した。ディスクI/Oの性能を除けば、ワークステーションはほぼ従来の大型計算機以上の実力を持つ事が判った。また同様のプログラムを用いて並列計算機での性能を調査した。従来より並列計算に向いているといわれている積分の計算ステップのみならず、繰り返し計算のステップにおいても高速化が得られた。CPUの数が大きくなるとCPUの数に比例する性能向上の傾向は得られず、8CPU程度で最大性能が得られた。しかしながらイーサネットで接続されたワークステーションクラスターでは10CPU程度でも性能の改善が見られた。

## Possibility of Workstation and Parallel Computer for Molecular Science

Umpei Nagashima and Mutsumi Aoyagi\*

Ochanomizu University, National Chemical Laboratory for Industry\*

Performance of some workstations and parallel computers are evaluated using a program for molecular science, which calculates Hartree-Fock self-consistent field molecular orbital. Except for I/O performance, recent workstations have almost the same performance of largest mainframes which are equipped at Japanese general purpose computer centers for academic use.

Remarkable improvement of the performance by the parallel computing was suggested from the results by Intel i860/64 and workstation cluster. In the case of i860/64, maximum performance was obtained at 8 CPUs whereas the performance of workstation cluster is almost proportional to the number of CPUs.

## 1. はじめに

1930年代に分子軌道法が確立されて以来、分子科学の一分野である計算化学は1940年代後半に実現された計算機の発達と共に急速に発展してきた。現在では計算機自体の高性能化のみならず数値計算法の改良をはじめ標準的な計算プログラムの流布、またワークステーション(W S)を用いた計算結果の可視化等により、分子軌道法のみならず様々なレベルの計算化学シミュレーション<sup>1)</sup>が身近なものになってきている。分子科学における計算機利用が急速に発達した背景には、従来の物質の性質や化学反応の理論的理解にとどまらず、新しい物質探求を従来より早く、安く、そして安全に行うという化学界全体の強い要請があった。化学は、各種の実験とそのデータの収集・整理、そしてそれらの理論的理解に基づく各種経験則の導出、さらにその経験則から示唆される新たな実験というサイクルをたどりながら発展していく。計算機はこのサイクルの多くの部分に深く関わっており、分子科学の発展と計算機の進歩は独立ではあるが深い関係にある。なかでも1980年代後半から急速に発達してきたW Sの価格性能比の高さは分子科学の研究方法すらも変えようとしている。そのため本稿では分子軌道計算を例に取り、最近のW Sの性能とその可能性を議論する。合わせて最近注目を浴びている並列計算機の可能性を探る。

## 2. 分子軌道法(ハートリーフォック方程式)<sup>1)</sup>

分子科学で取り扱われる分子の性質は電子のシュレディンガーの波動方程式を解く事でえられる。 $H_2^+$ の様な特別な場合を除いてそれは多電子問題となるので、ハートリーフォック近似を用いて解く事になる。ハートリーフォック近似を用いて波動方程式を解くというのは、全系の波動関数( $\Psi$ )をn個のスピン軌道からなる一つのスレーター行列式で表し、

$$|\Psi_0\rangle = |x_1 x_2 x_3 \dots x_n\rangle$$

全電子ハミルトニアン(H)によって記述される、N電子系の基底状態に対する最良の近似となるスピン軌道の組 $\{x_a\}$ を見いだす事に他ならない。変分原理によれば、それは電子エネルギー

$$E_0 = \langle \Psi_0 | H | \Psi_0 \rangle$$

を極小にすることである。規格直交性( $\langle x_a | x_b \rangle = \delta_{ab}$ )という束縛条件の基にEを極小化すると、最良のスピン軌道が満足すべき微積分方程式(ハートリーフォック方程式)

$$[h(1) + \sum J_b(1) - \sum K_b(1)] x_a(1) = \varepsilon_a x_a(1)$$

$$J_b(1) x_a(1) = [\int dx_2 x_b(2) r_{12}^{-1} x_b(2)] x_a(1)$$

$$K_b(1) x_a(1) = [\int dx_2 x_b(2) r_{12}^{-1} x_a(2)] x_b(1)$$

が得られる。左辺の  $[\ ]$  を  $f(1)$  と書くと

$$f x_a = \epsilon_a x_a$$

となり、固有関数としてスピン軌道を、固有値としてスピン軌道エネルギーを持つ固有値方程式を得る。

これは線形固有値方程式の形をしているが、 $f$  がこの固有値方程式の解である関数  $\{x_a\}$  に  $J$  と  $K$  の 2 電子積分を通じて依存しているため、実際には非線形方程式であり、繰り返しの手法で解かねばならない。そのため分子科学の分野ではハートリーフォック方程式の解  $x$  を SCFM と呼ぶ。この方程式を計算機で解く際の困難の多くは分子積分の計算とその保存にあるが本稿では深く触れない。

### 3. WS の性能

表 1 にいくつかの WS、メインフレーム、ベクトル型スーパーコンピュータで原子の軌道を求めるプログラム (ATOMSCF) を実行したときの CPU 時間を示す。計算の内容は、コバルト原子の原子軌道の最適化である。この結果は 5 回の測定の平均値であるが、それぞれの計算機で最高性能のコンパイルオプションを用いたものではなく、日常筆者が良く使用しているコンパイルオプションを利用しているため、単純に計算機のハード性能を反映しているものではないこと、また WS の CPU 時間は、極力余分なタスクの無い状況で測定したとはいえ、すべてのデーモン等の実行を止める訳にはいかなかったので、本当の CPU 時間ではなく、むしろ WS を単独で使ったときの時間となっていることを先にお断りしておく。メインフレーム及びベクトル型スーパーコンピュータの時間はシステムの出力による CPU 時間である。残念な事にこのベンチマークプログラムはベクトル計算機向きではなく、ベクトル計算機のス

表 1 いくつかの WS における  
ATOMSCF の CPU 時間

| Computer  | CPU time<br>(sec.) | Ratio |
|-----------|--------------------|-------|
| Sun4/330  | 403.3              | 20.47 |
| Sparc-2   | 254.8              | 12.93 |
| Mips 3230 | 104.9              | 5.31  |
| HP710     | 77.1               | 3.91  |
| HP730     | 68.9               | 3.49  |
| HP750     | 50.9               | 2.58  |
| IBM320E   | 39.7               | 2.01  |
| Cray XMP  | 38.2               | 1.94  |
| IBM320H   | 32.4               | 1.65  |
| H3050R    | 30.0               | 1.52  |
| M680      | 23.8               | 1.21  |
| S820      | 19.7               | 1.00  |

表 2 コンパイラの違いによる  
CPU 時間の変化

| Computer | C1    | C2    | Ratio<br>(C1/C2) |
|----------|-------|-------|------------------|
| HP710    | 77.05 | 30.36 | 2.53             |
| HP730    | 68.91 | 24.65 | 2.74             |
| HP750    | 50.94 | 21.73 | 2.34             |

カラー部分の評価にしかになっていない。

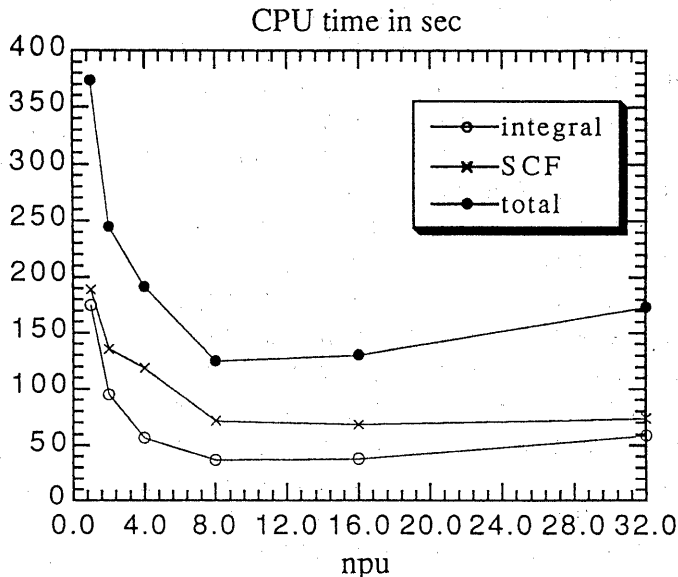
これを見ると、最近のWSは、ほぼ従来の大型メインフレームの1/2から1/3程度の性能を持っていることがわかる。従来の大型メインフレームが数人、また時には数十人で共有されていることを考慮すると、WSを使えばほぼメインフレームと同様の時間で計算結果を得ることができる。大型メインフレームが一台数億円のシステムであり、不特定多数のユーザによって共有されていることを考慮すると、一台数千万円のWSの価格性能比の高さに改めて驚かされる。ただしこれはベンチマークテストであり、ディスクI/Oが無い小さな問題であるので、WSの性能が過大評価されており、実際の大きな問題に対してはWSの方が歩が悪くなる。それはいくつかのファイルをアクセスする際のWSの性能がメインフレームに比べて貧弱であるからである。最近の分子科学の研究室においては、WSの価格性能比を計算手法の上から改良するために、分子積分をディスク上に保存せずに繰り返し計算の度に新たに計算するという方法も一般化してきている。

前回の研究会でも報告があったように、計算機の性能は問題の大きさに対してスケールアップではないという点は計算機の性能評価手法に大きな問題を投げかけている。さらに大きな問題点は、測定する計算機に載っているコンパイラの性能の様な計算機の環境をどの様に評価すべきかも問題である。表2にその例を示す。C1というコンパイラは、標準的に各計算機に装備されているもの、C2は別の計算機に装備されている

コンパイラである。これもそれぞれの計算機で最高性能が出るコンパイルオプションではなく、筆者が標準的に使用している物を利用したので両者の比較はただコンパイラの性能による違いの一例を示しているにすぎないが、このように自分のプログラムの性能がコンパイラによってほぼ2倍変わるとなると、コンパイラと計算機とそして自分のプログラムの整合性を見た上でWSを手にいれることを考えなければならない。

ともあれ最近のWSは従来の大型のメインフレームに十分対抗できる性

図1 i860におけるCPU時間(sec.)とプロセッサ台数の関係



能を持ってきたといえる。特に価格性能比では圧倒的にWSの方が高いといえよう。

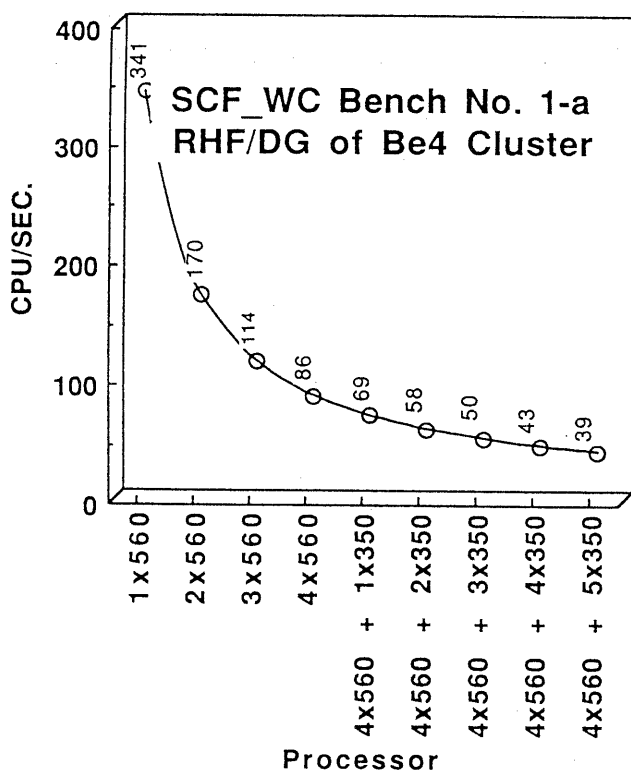
### 3. 並列計算の可能性

分子軌道法の計算の最も困難な計算ステップは、分子積分の計算であるが、この部分は並列性が極めて高く、並列計算機での性能向上が極めて高いと予想される。そこで今回たまたまであるがインテルのi860/64をほぼ一週間自由に使用する機会を持ったので、分子軌道計算のプログラムを一部移植し性能評価を行った。図1にその結果を示す。計算された分子は1、2、4トリアジン(C<sub>3</sub>H<sub>3</sub>N<sub>3</sub>)という分子であり、基底関数の数は66である。計算方法は、分子積分を従来の計算方法のようにディスクにしまうのではなく、繰り返しの度に毎回計算するという方法である。完全にプログラムがi860むけにコンバートされていないとはいえ、プロセッサの数が8までは、分子積分計算、フォック行列の対角化とともに並列化され、ほぼ1/nの挙動を示すが、nが16または32では

8の場合に比べ性能が低下している。これはまだ詳細な検討が終わっていないが、ロードバランスが崩れたためであろうと考えられる。もう少し系が大きくなればさらに並列性が高くなるので、大きな並列度でも性能向上が見られよう。分子科学の分野における並列計算機の重要性は極めて大きいといえる。

価格的にみてi860やCM2といった並列計算機がメインフレームであるとすれば、WSをネットワークで多数つないだワークステーションクラスターはWSに対応づけられよう。現在ワークステーションクラスターに関する研究は各所で精力的に進められており、PDSとしてすでにDongarraらのPVMとうが入手可能である。図2にベリリウムクラスターの計算を4台のIBM560と5台のIBM350で構築したワークステーションクラスターを用いて行った結果を示す。9台

図2 ワークステーションクラスターのCPU時間とプロセッサ台数



では 1 台でかかった CPU 時間の  $1/10$  の速度が実現されている。この図からは 10 台程度までは  $1/n$  に綺麗に乗っていることが読み取れよう。

大きな並列計算機は、その価格が数億円するのでそうたやすく手に入れることはできないし、並列計算機を持つ共同利用のセンターもまだ無いが、数十台程度のワークステーションは数千万円程度で手に入れることができ、10 数台のネットワークで接続されているワークステーションは現在でもそう大きな困難無く利用できる。控えめにみても大きな並列計算機がたやすく利用できるようになるまではワークステーションクラスターの持つ意義は大きい。

#### 4. まとめ

分子科学の分野において、最近の高性能 WS と並列計算機の持つ可能性は非常に高い物がある。現在はまだ個々の WS で従来のメインフレームで行われていた計算が実行され始めたにすぎないが、将来、並列計算環境が整備されるにつれ寄り一層の発展が予想される。

#### 参考文献

1) 分子軌道法の入門書としては藤永茂「入門分子軌道法」(講談社サイエントフィック、東京、1990)が、その上級書としてはザボ・オストランド「新しい量子化学」上下、大野公男、坂井健男、望月祐志訳(東大出版、東京、1988)が好著である。