

## 量子化学計算での非線形最適化問題への並列分散処理の利用

森 和 英\*

### An Application of Parallel Computing on Distributed Memory Computers to the Non-linear Optimization Problem in FVMO Method

KAZUhide MORI\*

(1995年12月15日受付, 1996年1月30日改訂)

並列分散処理をネットワーク環境で実現するソフトウェア PVM3.3を用いて, 量子化学計算, 特に FVMO (Full Variational Molecular Orbital) 法における非線形最適化計算の並列分散処理を行い, その並列化についての問題点 (計算時間, 高速化率など) を検討する。

**Abstract:** The parallel computing of quantum chemistry calculation on distributed memory computers, especially application to FVMO (Full Variational Molecular Orbital) method, in which orbital centers and orbital exponents in the basis functions are determined under the variation principle, is discussed in detail. This parallel computation strategy is implemented on workstation cluster employing the message passing library PVM3.3 (Parallel Virtual Machine). Parallelization strategy, performance and superlinear speedups running on a network of Unix computers through a combined use of the CPU and memory of the different processors are also discussed.

#### 1. 並列分散処理

近年, コンピュータ環境の著しい発展に伴ってこれまで行うことの出来なかった大規模な計算が並列処理を通じて行われるようになってきた<sup>1)</sup>。

並列処理とは, 一定時間若しくは単位時間あたりの計算量を増やし, 計算にかかる時間を短縮させようというものである。この並列処理には一つの命令で多くの計算を行うことで単位時間あたりの計算量を増やし経過時間を削減するベクトル化と, 複数のプロセッサを用いて同時に異なる処理を行い, 1つのプロセッサでの逐次処理に比べ経過時間を短縮する並列化の2つがある。ところが, ベクトル化を実行しようと思えば現在のところスーパーコンピュータに頼らざるを得ない状況で, とても手軽に計算するということは出来ない。一方, 並列化の方は超並列計算機のような専用機もあるがネットワークで結ばれた一連のマシン群があればよく, 並列化ツールの助けを借りれば手軽に並列化して計算を行うことが出来る。よって今後の数値計算はこの並列化を用いた形態が大きく発展すると思われる。

---

\* 情報科学センター 非常勤講師

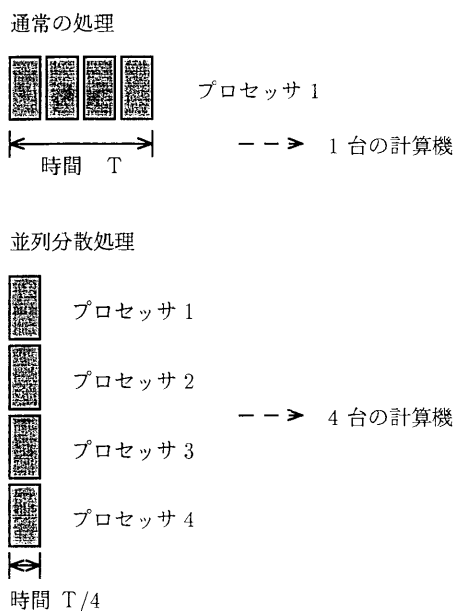


図 1 並列分散処理の概念

さて、並列化を用いた計算といってもハードウェア的に 2 種類がある。1 つは複数のプロセッサからアクセス可能なメモリーを持つ共有メモリー型の計算機を用いる計算であり、もう 1 つは各プロセッサにメモリーを配置し他のプロセッサにあるデータが必要になった場合にはプロセッサ間でデータの転送（通信）を行う分散メモリー型の計算機を用いる計算である。プログラムを書く上では共有メモリー型の計算機を用いる方が扱い易そうだが、計算機にかかるコストは大きくなり手軽に使えるとは言い難い状況にある。一方、分散メモリー型の計算機を用いる計算はデータの転送（通信）を行わねばならないがネットワークで結ばれた計算機を 1 つの並列計算機と見なすことができるので、端末室の計算機を使って手軽に並列化を用いた計算を実行することが出来る。よって、これからの計算はこの並列化を用いて分散メモリー型の計算機、つまりネットワーク環境で計算を行うという形態が頻繁に行われるようになると思われる。このような処理形態を並列処理を分散メモリー型計算機上で行うという意味で並列分散処理と本研究では呼ぶことにする。これまでの概念図を図 1 に示す。

今回、この並列分散処理を並列化ツールである PVM を用いて FVMO 法の非線形最適化問題に適用した。

## 2. 並列分散処理を実現するツール

並列分散処理をネットワーク環境で実現するツールには様々なものがある。一例を上げれば MPI<sup>2)</sup>, TCGMSG<sup>3)</sup>, PVM<sup>4)</sup> 等である。これらは全て異機種混合ネットワーク環境をひとつの仮想並列計算機として利用できるような仕様になっている。これらのツールの共通の特徴は各プロセッサ上で同時に動作しているプログラムでデータをやりとり (通信) しながら処理を行う、いわゆるメッセージパッシングをサポートしている点と、プログラムを記述する言語として FORTRAN と C サポートしているという 2 点である。

今回は PVM3.3.7 を用いた。PVM は日本語のマニュアル等が存在し、上記の並列化ツールの中では一番取り付きやすいと思われる。PVM では上記のメッセージパッシングを実現するのに PVM デーモン、FORTRAN と C をサポートするために PVM ユーザーインターフェイスライブラリを備えている。

## 3. 並列化プログラミング

並列化のツールを使用してプログラムを記述する方法には 2 通りある。1 つは全てのプロセッサで同じプログラムを実行させる SPMD モデルと、もう 1 つは複数のプログラムを用意し実行させる MIMD モデルである。SPMD モデルではコンパイル等の実行させるまでの手間は楽だが、並列化を用いて実際に計算させた時に本来そのプロセッサでは計算されない変数や配列にまでメモリーが消費されてしまう。これに対し MIMD モデルではコンパイルなどの実行させるまでは面倒であるが、実際に計算を行うプロセッサでは計算されない変数や配列を宣言していないプログラムを実行させることができるのでメモリーの効率的な使用が可能である。

量子化学計算においては一般にメモリーの使用量は大きく、メモリーの効率的な使用に留意しなければならない。本 FVMO 法についてもそれは同様であるため本研究では MIMD モデルを用いた。また、MIMD モデルでもプログラムが 2 種類の場合には特にマスター/スレーブモデルとかホスト/ノードモデルと言うが、今回はこのマスター/スレーブモデルを採用した。

マスター/スレーブモデルを用いるとマスタープログラム、スレーブプログラムと呼ばれる 2 種類のプログラムを記述し、実際の計算においてはある 1 つのプロセッサでマスタープログラムを起動し、他のプロセッサで複数のスレーブプログラムを実行させるという形態を取ることになる。この形態の利点はマスタープログラムからスレーブプログラムをコントロールすることができるという点である。

実行の仕方をごく簡単に説明すると、まず PVM デーモンをある計算機で起動させ、次に使

用する計算機の全てで PVM デーモンを起動させる。ここで、一番最初に PVM デーモンを起動した計算機から他の計算機に対して PVM デーモンを起動させることができるので、全ての計算機にログインする必要は無い。こうして PVM デーモンを起動させた状態でマスタープログラムを実行させる。スレーブプログラムはマスタープログラムの中から各計算機に対して起動させることができる<sup>5)</sup>。

並列化したプログラムの流れの概念を図 2 として示す。

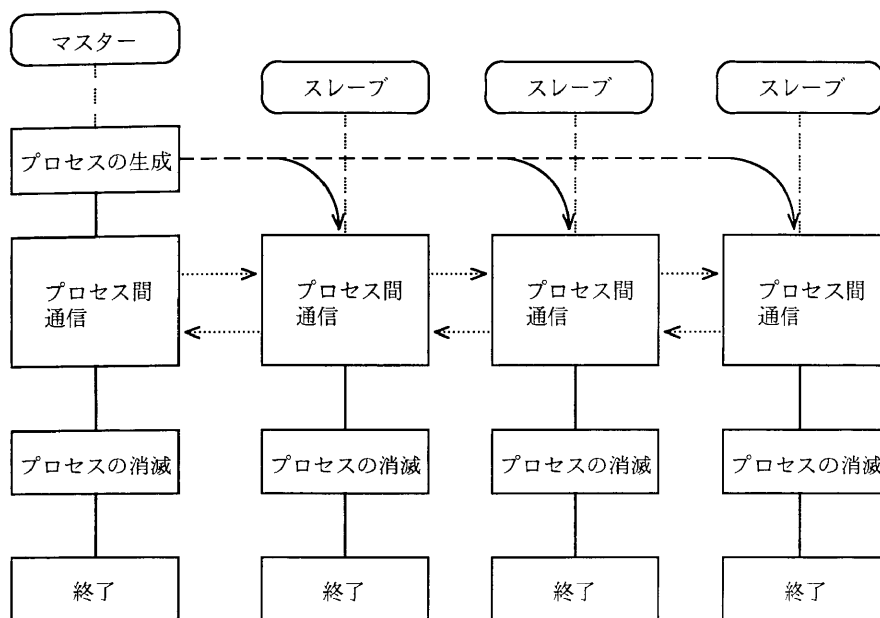


図 2 並列化したプログラムの流れの概念

#### 4. F V M O 法

FVMO (Full Variational Molecular Orbital) 法<sup>6)</sup>は従来の基底関数展開 (重ね合わせの原理) を用いた分子軌道法と異なって、シュレディンガー方程式中に含まれる物理定数以外のパラメータは全て変分パラメータとする。つまり、基底関数として用いる Gauss 型の軌道に対してその係数と軌道指数のみならず軌道中心も最適化の対象として取り扱う。このため、その計算では非線形最適化プロセスは飛躍的に増大し、計算時間もそれに伴い著しく増加する。よって今回この非線形最適化部分のみを並列分散処理の対象とし、各 gradient の計算を各プロセッサに分配するモデルを考え、実行した。FVMO 法の概略を図 3 に示し、FVMO 法における並

Gauss 型関数 :

$$C N(\alpha) e^{\alpha(x-X)^2}$$

(  $N(\alpha)$ : 規格化定数 )

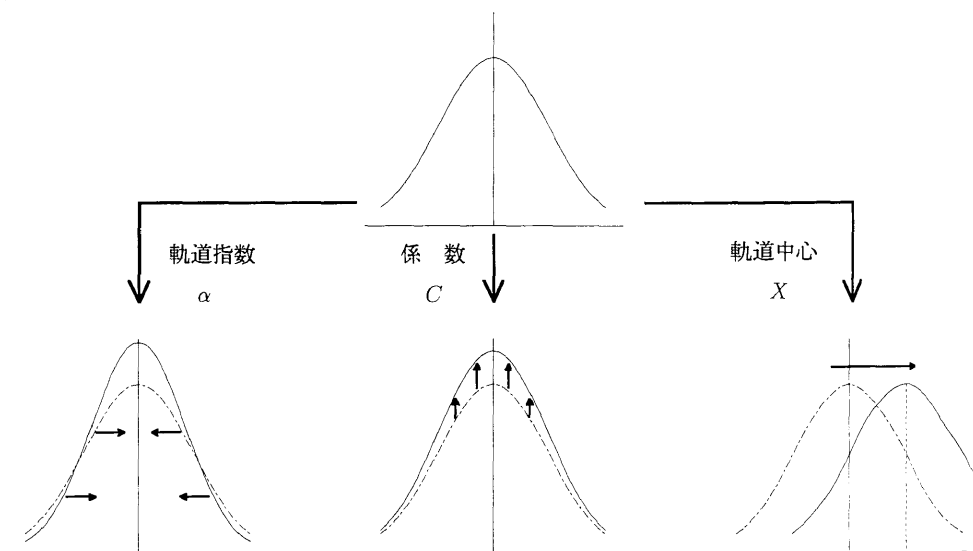


図3 FVMO法の概略

列処理の概念図を図4に示す。

## 5. 結果と考察

計算は  $\text{HeH}^+$  分子に6個の Gauss 型基底関数を置いたもの（これを  $\text{HeH}^+ 6G$  と略す）と He 原子に10個の Gauss 型基底関数を置いたもの（これを  $\text{He } 10G$  と略す）に対して計算機の台数を変化させて行った。ここで各スレーブプログラムは1つの基底関数を受持つ。つまり、 $\text{HeH}^+$  分子の計算では6個のスレーブプログラム、He 原子では10個のスレーブプログラムをそれぞれ起動することになる。このように対象とする原子分子によってスレーブプログラムの数は固定されているため、スレーブプログラムを各計算機に分散するとき計算機の台数によってばらつきが生じ、それが計算時間の短縮に現れる箇所がある。この結果を図5と図6に示す。

## ワークステーションクラスター

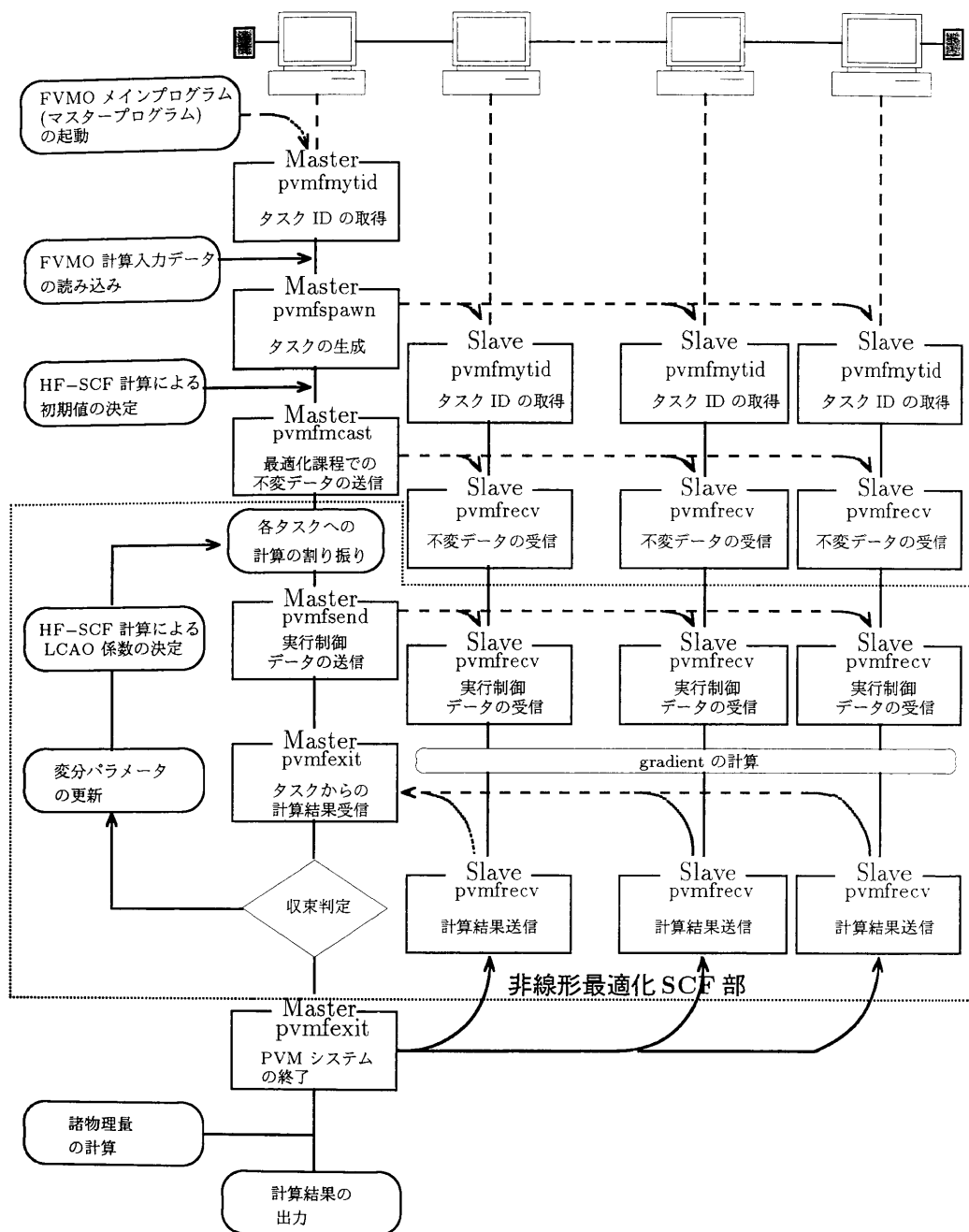


図 4 FVMO 法における並列処理の概念図

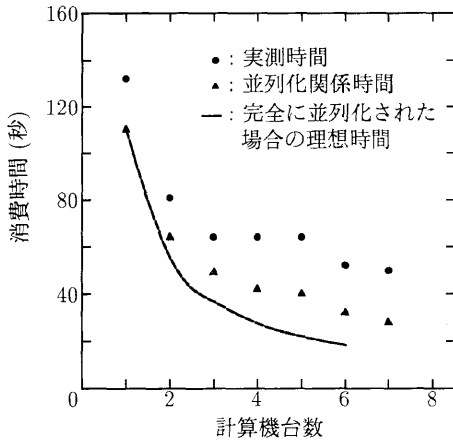


図5 HeH-6G の計算機台数と消費時間の変化

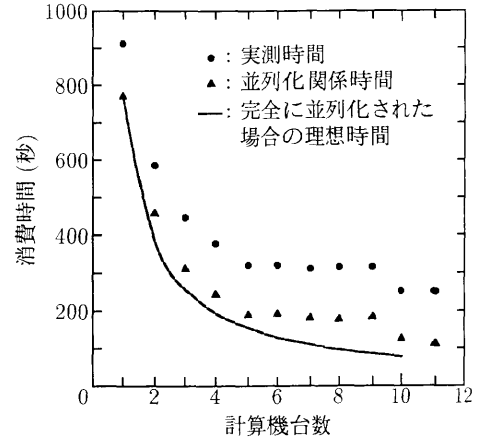


図6 He 10G の計算機台数と消費時間の変化

図5の中で計算機台数が7台の箇所はマスタープログラムとスレーブプログラムを完全に別々の計算機で実行させた結果である。図6中の11台の箇所も同様である。

図5では計算時間が短いために通信時間の全体に対する割合が増大し理想的な曲線から離れているような印象を受ける。計算時間がもう少し長い図6ではこの傾向が幾分弱められ、理想的な曲線からのずれは小さくなっている。この様子は縦軸に計算時間の逆数を取り、これを速度向上率として定義した図で表すと良く分かる。これを図7と図8として示す。

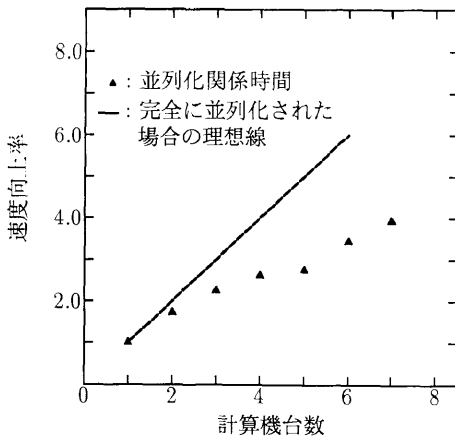


図7 HeH-6G の計算機台数と速度向上率

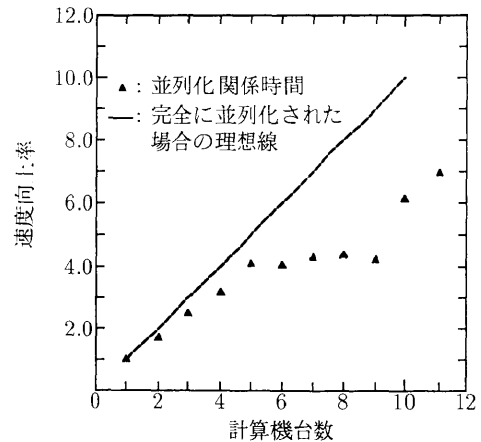


図8 He 10G の計算機台数と速度向上率

理想線からのずれはほとんどが通信にかかった時間である。また、逆数をとったため図7と図8中の計算機台数の多い所では理想線からのずれを過大評価してしまう。ここで図7中の計算機台数が3~5台、図8中の計算機台数が5~9台のところはスレーブプログラムの分配に対して

得をしなかった所である。よって並列化関係時間が横這いになっている。

以上のことより並列化の効率を高めるには次の3点に関する考察が必要であると結論される。①通信にかかる時間の減少, ②全体の計算量の均等な配分, ③なるべく計算規模の大きな問題への適用, である。

ここで, ①の通信にかかる時間はさらに通信量と通信回数に依存する。どちらも減らすように出来れば一番良いのだが, ハードウェアの性能の向上により, 通信量の問題は改善される可能性があるのでできれば通信回数を減らすようにすると良いと思われる。②については, 今回のように計算量を均等に分配することが出来る場合は問題ないが計算量が見積もれないもしくは不均等にしか分けられないときは循環に分配するという方法を使うと良いと思われる。最後の③については, 計算の規模が大きい場合通信にかかる時間が計算量に比すると相対的に減少するのでより理想線に近づき並列化による恩恵を多く受けることが出来る。

## 6. 参 考 文 献

- 1) 並列処理の文献は数多く存在するので比較的最近の文献のみ示す。  
T. R. Furlani, and H. F. King, J. Comp. Chem., **16**, 91 (1995),  
A. T. Wong, and R. J. Harrison, J. Comp. Chem., **16**, 1291 (1995),  
D. E. Bernholdt, and R. J. Harrison, J. Chem. Phys. **102**, 9582 (1995)  
日本語で書かれた文献としては,  
大西樽平, 固体物理, **30**, 789 (1995).  
等がある。
- 2) いろいろな派生型がある。今一つだけ示す。  
<ftp://info.mcs.anl.gov/pub/mpi/mpich.tar.gz> 等
- 3) 他にいろいろなツールも入っているが以下のファイルを展開すると得られる。  
<ftp://ftp.pnl.gov/pub/global/global2.0.tar.gz>
- 4) 国内にも多くのサイトに存在する。今, 最新バージョンを示す。  
<ftp://ftp.sra.co.jp/pub/misc/netlib/pvm3/pvm3.3.10.tar.gz>
- 5) UNIX MAGAZINE 4月号, 1995年, (株式会社アスキー)。
- 6) 小林, 種田, 後藤, 森, 宗像, 鈴木 (賢), 相賀, 笹金, 伊藤, 日本化学会第69春季年会, 1E137, (1995)  
鈴木 (賢), 末原, 種田, 森, 笹金, 相賀, 伊藤, 日本化学会第69春季年会, 1E138, (1995)  
種田, 小林, 後藤, 森, 伊藤, 日本化学会第69春季年会, 1E139, (1995)  
後藤, 森, 宗像, 笹金, 相賀, 小林, 伊藤, 日本化学会第69春季年会, 1E140, (1995)  
高田, 森, 宗像, 大江, 伊藤, 日本物理学会第50回年会, 28aPS-58, (1995)  
後藤, 森, 宗像, 高田, 大江, 伊藤, 分子構造総合討論会, 3P0056, (1995)  
種田, 森, 伊藤, 分子構造総合討論会, 4A02, (1995)  
大江, 鈴木 (一), 森, 立川, 笹金, 井口, 伊藤, 分子構造総合討論会, 4Pb37, (1995)  
宗像, 森, 大江, 伊藤, 分子構造総合討論会, 4Pb38, (1995)