

大次元行列計算における S-810 拡張記憶装置 利用の効果

長堀 文子 後 保範 長谷 均 赤間 由香里

(株) 日立製作所 ソフトウェア工場

スーパーコンピュータ S-810 は、汎用計算機に比べ、より効果的に大規模行列計算を行うことができる。ところが、CPU時間が飛躍的に向上すると、従来CPUバウンドであったジョブがこんどは入出力バウンドジョブの性格を帯びてくる。そこで、スーパーコンピュータに効果的に大量の入出力データを供給する装置が必要になる。こうした背景により、拡張記憶装置が開発された。

われわれは、入出力処理を伴う大次元行列計算の一例として、入出力機能付きの連立一次方程式の直接解法ルーチンを開発し、S-810 システムにおいて、拡張記憶装置と磁気ディスク装置を用いていくつかのテスト問題を解いた。

その結果、拡張記憶装置を使用した場合のユース時間は、磁気ディスク装置使用の場合に比べ、約 1/100 ~ 1/200 倍に短縮できた。

Effect of using Extended Storage of the HITAC S-810 system (in japanese)

F. Nagahori, Y. Ushiro, H. Nagatani and Y. Akama

Software Works, Hitachi, Ltd.
549-6 Shinano Machi Totsuka Ku Yokohama-shi
Kanagawa Ken, 244 JAPAN

The supercomputer like HITAC S-810 enables us to compute large scale matrix calculations far more efficiently than ordinary scalar processing computers. Increasing the CPU speed, the proportion of Input / Output processing time to the total execution time is becoming considerably large. So it is necessary for the supercomputer to provide the efficient Input / Output data transferability. The Extended Storage (ES) of S-810 is developed for this purposes.

We developed simultaneous linear equations solvers with Input / Output facilities and these solvers have been used to solve some test problems both with ES and with ordinary magnetic disc device as external data files under S-810 system.

As attractive result, we have recognized that the elapsed-time with using ES is about 1/100 ~ 1/200 that with using disc.

1. はじめに

スーパーコンピュータ HITAC S-810 のベクトル計算の特長については、本研究会でも既に報告²⁾ されている。又、S-810 による高速な行列計算は、MATRIX/HAP という行列計算ライブラリとして提供されており、それに関する文献³⁾ もある。さらに、汎用構造解析プログラム I SAS II のスカイライン法による解法処理を S-810 向きに改良した I SAS II/HAP についても、本研究会での報告⁴⁾ がなされている。このように、ベクトル計算機普及時代の到来により、大規模な行列計算が可能となってきた。

しかし、オペレーティングシステムにより 31 ビットアドレッシングモードがサポートされ、飛躍的に広いユーザ空間が利用可能となったとはいえ、ユーザがジョブの実行時に利用可能なメモリ空間は数メガバイトというのが普通である。主記憶装置上に収納しきれないほど大規模な行列データを扱う場合には、外部記憶装置上のファイルと主記憶装置上の配列との間でデータの転送をしなくてはならない。

いかにベクトル処理による演算速度が高速であっても、それに見合うだけの速さでデータが供給されなければ、プログラムの実行における計算機のユース時間は短縮できない。そこで HITAC S-810 システムは、主記憶装置の外側に一時的データセットとして使用する磁気ディスク装置の代替となり、しかも、磁気ディスク装置よりはるかに高速に主記憶装置とのデータ転送ができる拡張記憶装置を有している。

本報告では、行列計算における拡張記憶装置利用の有効性を述べる。行列計算の代表として、実用的価値の高いガウスの消去法に基づく連立 1 次方程式の直接解法をとりあげた。係数行列を主記憶装置上に納まるデータ量に分割（ブロック化）して、あらかじめ外部ファイル上に記録しておき、これを逐次ブロック単位に読み込んで計算する、入出力機能付きブロック

ガウス法と名付けたアルゴリズムを用いた。外部ファイルとしては、磁気ディスク装置と拡張記憶装置とを使用して、その CPU 時間とユース時間の両方を比較した。

2. S-810 システムの拡張記憶装置

2.1 S-810 システムの記憶装置

S-810 システムの構成を記憶装置を中心に概念的にながめてみると、図 2.1 のようになっている。

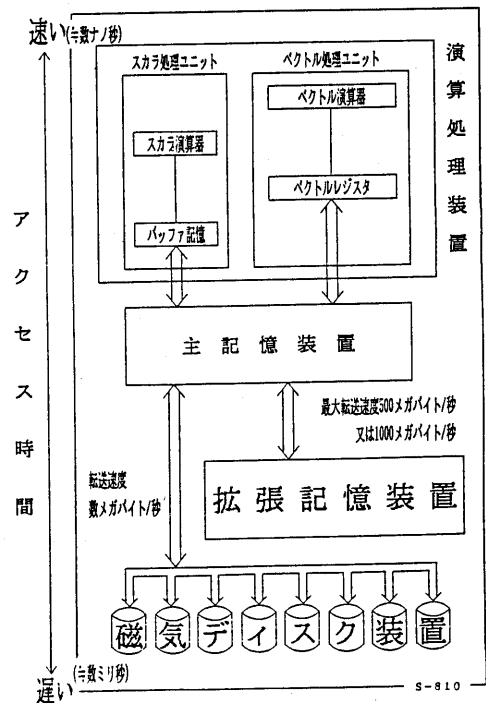


図 2.1 S-810 システムの記憶装置の階層構造

図 2.1 では図の上方に位置する記憶装置からのデータ取り出しは高速で、図の下方に位置する記憶装置からのデータ取り出しは低速であることを示している。

高速なベクトル演算器にベクトルデータを高速

に供給するために、複数セットのベクトルレジスタがあり、さらに主記憶装置からベクトルレジスタへのデータ転送効率を増すために、メモリインタリーブ方式や、複数のデータバスによる並列転送を実現している。メモリインタリーブ方式とは、主記憶装置がバンクとよばれるある単位に分かれていて、異なるバンクに属するデータは、ある時間的なピッチで次々と切れ目なく取り出せる工夫である。

図2.1の中で、データ取り出し時間の点で最も問題となるのは、主記憶装置と磁気ディスク装置との間のデータ転送である。最近の磁気ディスク装置といえども平均的な位置決め時間に十数ミリ秒かかり、半導体ディスク装置でも1ミリ秒を要する。S-810システムでは、主記憶装置との間に複数のチャンネルを置き、これらのチャンネルを並列に利用できるようにして、転送効率をあげている。又、FORTRAN処理系でサポートしている複数ボリュームパラレル入出力機能により、演算処理装置へのデータ供給を高速化している。

しかし、主記憶装置に格納しきれない大容量データを扱う場合に最も有効なのは、拡張記憶装置の利用である。拡張記憶装置は、データ転送速度が500メガバイト/秒のもの、1000メガバイト/秒の2つのタイプがある。H-8598磁気ディスク装置のデータ転送速度が3メガバイト/秒であることと比較すれば、その高速性がよくわかる。

2.2 拡張記憶装置の概要

S-810システムの拡張記憶装置に対する入出力は、FORTRAN言語の従来の入出力文を使って磁気ディスク装置に対する入出力と同様に使用できる。拡張記憶装置は、一時的データセットとして使用したり、主記憶装置上に納まりきれない大規模データの処理で、主記憶装置の延長として利用することができる。主記憶装置の延長という意味は入出力時間

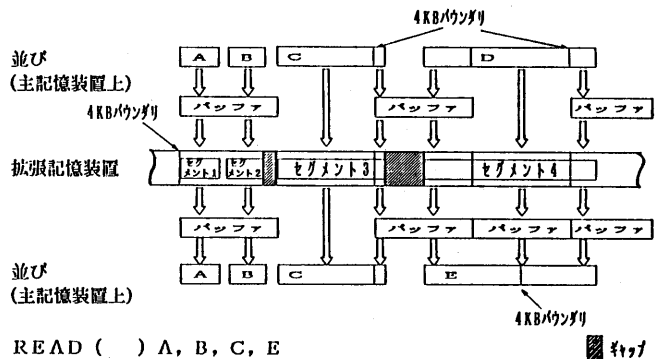
が極めて短いという意味であり、使い方からみれば入出力文を使った外部記憶装置にあたる。

2.3 拡張記憶装置の基本性能

(1) バッファレス入出力の効果

S-810システムの拡張記憶装置は、4キロバイト（ページ）単位に管理されている。拡張記憶装置とデータ転送をする入出力並びの先頭が4キロバイト（ページ）バウンダリに調整されていて、しかも入出力並びが4キロバイト以上であるときに、書式なし入出力の場合は、主記憶装置との間で直接データ転送ができる。4キロバイト（ページ）バウンダリに調整されていないときや転送量が4キロバイト以下のときは、図2.2に示すように主記憶装置上に確保したバッファ領域を介して、転送される。

WRITE () A, B, C, D



READ () A, B, C, E

図2.2 書式なし順番探索入出力のデータ転送

バッファを経由しない入出力をバッファレス入出力とよぶことにする。バッファレス入出力の効果を表2.1に転送速度として示す。

表 2. 1 書式なし順番探索入出力による
バッファレス入出力の効果

転送データ量 (メガバイト)	バッファレス効 転送速度 (メガバイト/秒)
8	30.7
64	174.5
256	350.2
1024	468.0
2048	496.6
4096	510.9

- 注 1) 使用した拡張記憶装置は最大転送速度が
500 メガバイト/秒タイプのものである。
- 注 2) 転送速度は数十回～数千回繰り返して実行
した平均値である。

表 2. 1 から分かるように、バッファレス入出力を利用すれば、転送量が 4 メガバイト以上のときには装置の最大転送速度まで達することができる。

このように、バッファレス入出力を利用することにより高速入出力処理が実現できることが分かる。

(2) BACKSPACE文 実行性能

従来の外部記憶装置の使用では、FORTRAN プログラムで順番探索ファイルに対しBACKSPACE文を実行すると、順方向のアクセスに比べ極めて遅かった。プログラマはできるだけBACKSPACE文を避けてコーディングするように工夫してきた。しかし、ガウスの消去法による連立1次方程

式の直接解法の後退代入過程では、LU分解後の行列Uのデータを右下隅から逆方向にアクセスしなくてはならない。このことは、行列データを分割化して順番探索ファイル上に記録する場合には、BACKSPACE文の使用が避けられないことを意味する。そこで、実際に連立1次方程式の解法ルーチンを作成する前に、BACKSPACE文を使って拡張記憶装置上のファイルからデータを読み込む場合のデータ転送速度を実測することにした。

図 2. 3 に示すように、ファイルの終端からBACKSPACE文とREAD文を組合せて、最後の記録から第1番目の記録まで逆に戻りながら読み込む場合と、第1番目の記録から順に読み込む場合との転送時間を比較した。その結果を表 2. 2 に示す。

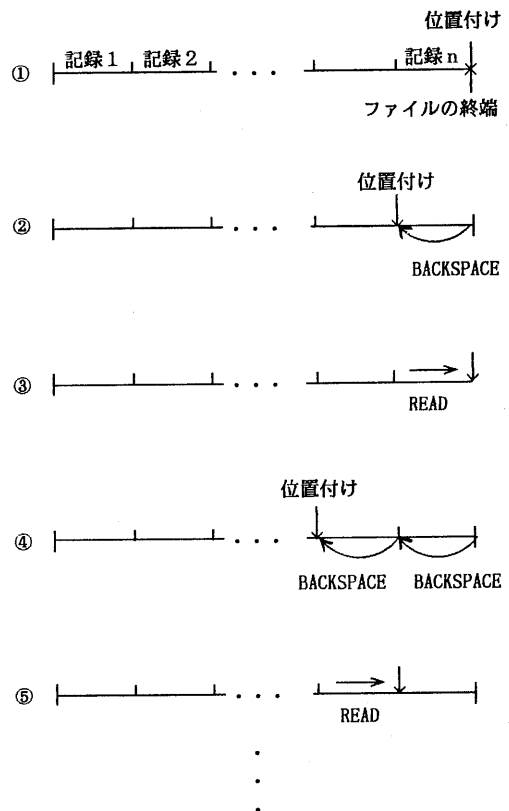


図 2. 3 BACKSPACE文を使った
順番探索ファイルの逆読み

表2.2 BACKSPACE文を利用した順番探索
ファイルの逆読み転送速度

転送データ量 (キロバイト)	① 順読み 転送速度 (メガバイト/秒)	② 逆読み 転送速度 (メガバイト/秒)	速度比 ②/①
8	30.7	19.9	0.65
64	174.5	126.0	0.72
256	350.2	293.2	0.84
1024	468.0	439.5	0.94
2048	496.6	479.9	0.97
4096	510.9	503.1	0.98

表2.2から分かるように転送量が2メガバイト以上であればBACKSPACEの影響はほとんど無視できる。

3. 行列計算での拡張記憶装置の利用例とその効果

行列計算の代表として、ガウスの消去法に基づく連立1次方程式の直接解法を取り上げた。係数行列データを限られた容量の主記憶装置上の配列に入りきるだけの量に分割（ブロック化）し、分割単位ごとに外部ファイル上に記録しておき、これをプログラムで確保した配列上に1ブロックずつ読み込んで、参照又は更新し、LU分解を進める。得られた分解結果は、外部ファイル上に記録として書きだしていき、代入処理で1ブロックずつ読み込み参照する。これを入出力

付きブロックガウス法と名付けた。

入出力付きブロックガウス法は、図書館情報大学の村田健郎教授が日立製作所の技師長時代に、大次元行列を仮想記憶計算機上で扱う場合に、メモリスワップを少なくするアルゴリズムとして御提案された‘たてブロックガウス法’⁹⁾の思想を継承している。

‘たてブロックガウス法’では仮想記憶装置と主記憶装置の間の入出力が行われているが、この入出力（メモリスワップ）は、オペレーティングシステムによってなされ、プログラムは入出力を意識していない。一方、入出力機能付きブロックガウス法は拡張記憶装置、又は磁気ディスク装置といった外部記憶装置と主記憶装置上のデータ転送を、プログラム内部で入出力文の実行という形で行われる。

入出力機能付きブロックガウス法では、主記憶装置上に2ブロック分のデータが必要である。一方は参照され、他方は更新されるデータブロックである。

入出力機能付きブロックガウス法のプログラムとして、

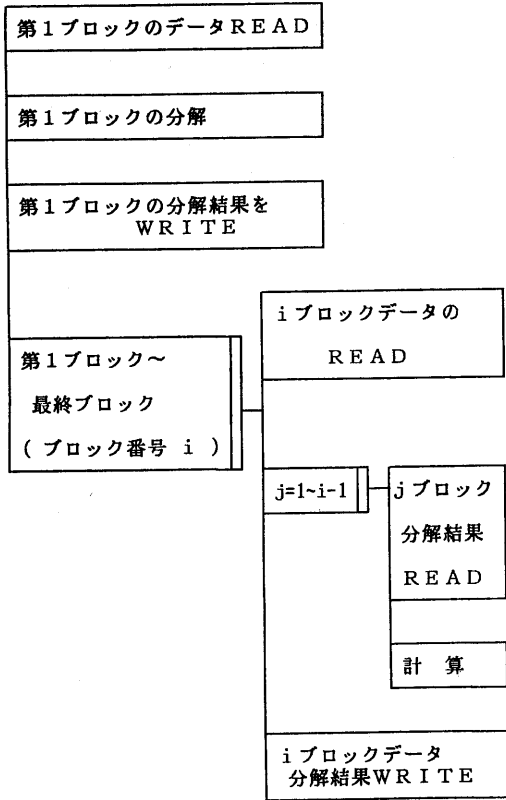
- (1) 実対称密行列
- (2) 実非対称密行列
- (3) 実対称帯行列
- (4) 実非対称帯行列

の4種類の係数行列用に計4本を用意した。拡張記憶装置を利用する場合には、2.3(2)で示したように、代入処理の逆読みによる性能低下の心配はないことが分かったが、磁気ディスク装置用としてもできるだけ速いものとするために、アクセス法は直接探索入出力を採用することにした。

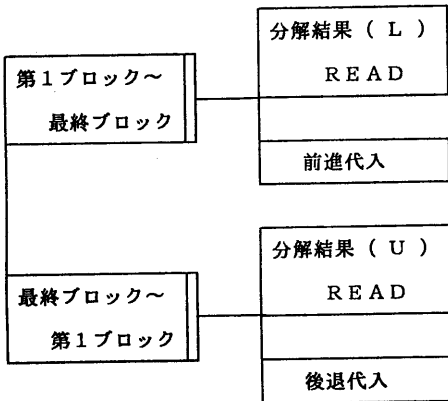
計算の流れを次に示す。

3.1 実測効果

LU分解処理



代入処理



S-810モデル20の下で最大転送速度 500 メガバイト/秒 タイプの拡張記憶装置と H-8598型磁気ディスク装置とを使って、入出力機能付きブロックガウス法プログラムの実行時間を測定した。

拡張記憶装置の利用効果を見るために、磁気ディスク装置使用の場合のユース時間、CPU時間それぞれを比較して、拡張記憶装置使用の場合が何倍速くなっているかを表3.1に示す。なお、実行はいずれもベクトルモードである。

表3.1 入出力機能付きブロックガウス法のプログラムにおける拡張記憶装置利用の効果

係数行列	分解処理		代入処理		主記憶容量 メガバイト
	ユース時間	CPU時間	ユース時間	CPU時間	
	高速化倍率	高速化倍率	高速化倍率	高速化倍率	
対称密 2000次元 ブロック数7	8.3 倍	1.3 倍	168 倍	7.4 倍	5
非対称密 2000次元 ブロック数11	11.6	1.4	192	8.3	6
対称帯 5000次元 帯幅250 ブロック数5	15.2	1.6	91	4.4	4
非対称帯 5000次元 帯幅501 ブロック数5	8.1	1.1	128	5.8	10

表3.1から分かるように、代入処理のユース時間は拡張記憶装置利用効果が、91~192倍と極めて顕著である。これはS-810のベクトル処理のような演算の高速な計算機では、通常の磁気ディスク装置を使ったのでは代入処理が入出力バウンドになってしまうことを表している。これに対し、分解処理は演算の絶対量が多く、CPUバウンド性が強いいため、拡張記憶装置利用の高速化効果は10倍前後という結果になっている。

例えば、線形非定常解析問題の計算では、連立1次方程式のLU分解は1回だけ計算し、代入処理を数十回あるいは100回以上繰り返すというようなことが多いため、代入処理のユース時間が100倍違うのでは、拡張記憶装置を使用したい。

4. む す び

S-810システムの拡張記憶装置の高速入出力機能の有用性を、入出力機能付きブロックガウス法プログラムを利用して確かめた。

スーパーコンピュータの普及に伴い演算速度の飛躍的向上がもたらされ、より大規模な計算が実用化されつつある。しかし、入出力処理を伴う計算の場合、CPUバウンドであった計算が、CPU時間の大幅短縮により、今度は入出力バウンドの性格を帯びてくる。そのような計算の場合、通常は磁気ディスク装置の入出力に比べ、100倍以上も速いデータ転送能力を持つ拡張記憶装置の利用がCPU時間と入出力時間のアンバランスを解消する。

拡張記憶装置は標準装備ではない。S-810システムにおいては、是非とも拡張記憶装置を組み込みたい。

今回報告した入出力機能付きブロックガウス法は、有限要素法や境界要素法のプログラムの解法に適用可能であり、実用的価値も高いと考えている。

《 参 考 文 献 》

- 1) 唐木, 情報処理学会数値解析研究会 12-2 ('85.2.22) 「スーパーコンピュータ S-810の応用機能」
- 2) 唐木, 情報処理学会数値解析研究会 8-3 ('84.2.24) 「スーパーコンピュータ S-810 東大システムにおける数学ライブラリの性能」
- 3) 後, 外, 日立評論 65,8 ('83.8) 「スーパーコンピュータ "HITAC S-810" による行列計算」
- 4) 原野, 外, 情報処理学会数値解析研究会 15-2 ('85.2.22) 「スーパーコンピュータ S-810 向け構造解析プログラム ISAS II/HAP」
- 5) HITAC マニュアル プログラムプロダクト 数値計算副プログラムライブラリ MSL II 機能編 -第1分冊- 行列計算