

「NWT並列FORTRANに基づく並列評価」

○中村孝\*、吉田正廣\*、福田正大\*  
 中村修一\*\*、村瀬丈夫\*\*、松崎達哉\*\*

Performance evaluation of the NWT with parallel FORTRAN

by

Takashi Nakamura, Masahiro Yoshida and Masahiro Fukuda  
 National Aerospace Laboratory

Shuichi Nakamura, Takeo Murase and Tatsuya Matsuzaki  
 Fujitsu Ltd.

ABSTRACT

The Numerical Wind Tunnel (NWT) is under research and development by NAL-Fujitsu joint activity. It is a parallel computer system of distributed memory architecture composed of vector processors.

In this paper, we show performance on the desk of two application programs that were parallelized by current parallel language-processor. We evaluated the performance using network simulator and evaluation software "VTAP" which is based on VP400 performance analysis.

We present estimations than our machine performs CFD simulation about 100 times faster than Fujitsu VP400.

1. はじめに

スーパーコンピュータ性能の向上によりCFD研究は三次元粘性流解析、さらには燃焼反応、解離/再結合反応を含む流れをも対象とするようになった。また、より現象に忠実な物理モデルと、より正確な数値解を得るための近似法を採用するようになってきた。これに伴って、必要とされる計算機性能もさらに大きくなってきた。

一方、単一の計算機の性能向上は飽和に達しつつあり、現時点での単体の最高速計算機は日立のS3000シリーズで、8GFLOPSである。このため、多くの分野からの要求に応えるには、計算機の並列化が避けては通れない状況になっている。また、昨今Massively Parallel Processorを中心とする並列計算機の話題には事欠かなくなっている。

数値風洞は、CFD研究を対象として富士通VP400の100倍以上の実効性能を実現しようと、航技研と富士通が共同で開発している並列計算機である [1] [2] [3] [4] [5]。

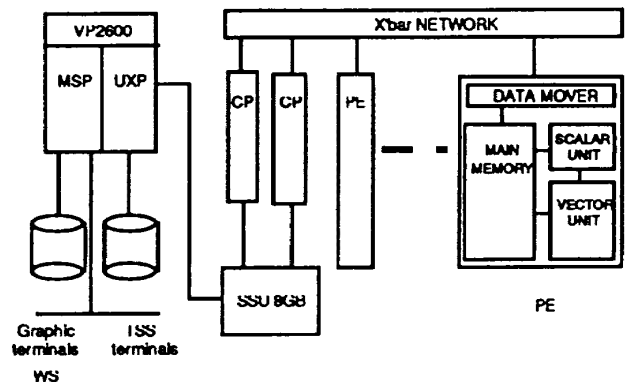


図1 システム構成図

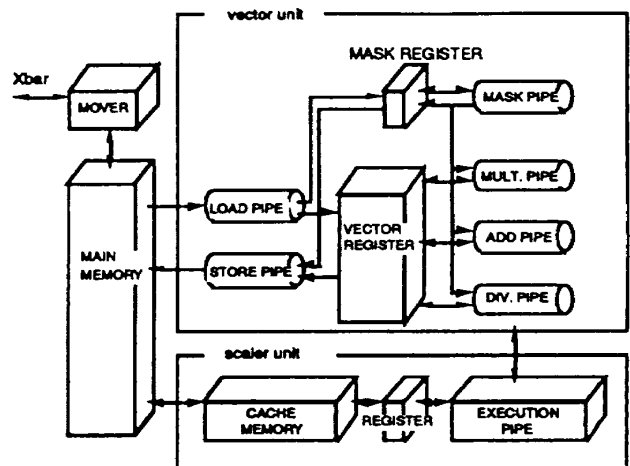


図2 PE構成図

\*航空宇宙技術研究所  
 \*\*富士通株式会社

本報告で、航技研の研究者が実際のシミュレーションに使用しているCFDコードを用いて行った数値風洞の机上での性能評価の結果について述べる。

## 2. 数値風洞のハードウェア概要

数値風洞は256MBの主記憶を持ち処理速度が1.6GFLOPSのベクトル計算機140台をクロスネットワークにより結合した、分散主記憶型の並列計算機である(図1)。要素計算機(Processing Element, PE)の構成を図2に示す。パイプライン多重度は8である。即ち1マシンサイクル毎に8個の演算結果が出力される。演算パイプラインとして加算、乗算、除算の3本、ロード・ストアのパイプラインが各1本ずつ、および論理演算、マスク演算のためのマスクパイプラインが1本ある。またベクトルレジスタの総容量は128KBであり、VP400と同じく16本×1K長×8Bから256本×64長×8Bとその構成を変えることができる。スカラ命令は1マシンサイクルの間に3命令が発行できるLIW方式を採用している。

主記憶は物理的には分散型であるが、OSとFORTRANにより論理的な共有主記憶空間が提供される(以後グローバル空間と言う)。グローバル空間は各PEの主記憶中から構成され、それ以外の主記憶はローカル空間を構成する。グローバル空間として割り付けるデータはユーザが指定できる。ロード/ストアパイプラインが直接にアクセスできるのはローカル空間に割り付けられたデータである。

グローバル空間に割り付けられたデータはプログラム中で任意のPEから任意にアクセスが可能である。実際には、PE<sub>i</sub>がPE<sub>j</sub>にあるデータを参照すると、ムーバと呼ぶハードウェアおよびネットワークを経由してPE<sub>j</sub>にあるデータがPE<sub>i</sub>の作業用ローカル空間に転送される。また自分自身の主記憶上のグローバル空間のデータであってもムーバ経由でアクセスされる。一方、ローカル空間上のデータはそのロ

ーカル空間を所有しているPEからでないとアクセスできない。即ち、PE<sub>i</sub>のローカルデータをPE<sub>j</sub>からはアクセスできない。

クロスネットワークの転送速度は、送信・受信の各々が400MB/秒で同時に動作可能である。ムーバと主記憶間の転送速度は800MB/秒である。

## 3. 並列処理記述言語

数値風洞での並列処理を記述するために、コンパイラ指示行方式を採用した。ユーザが記述しなければならない指示行の基本的なものは、

- (1) 使用するPE台数の指定
  - (2) データのグローバル/ローカル割付
  - (3) アレイデータの分割配置
  - (4) 手続き(主としてDOループ)の分割
  - (5) 同期
- 等である[6]。

## 4. 数値風洞の性能評価

数値風洞の性能評価には、航技研のCFD研究者が使用している三次元ナビエ・ストークスコードを用いて行った。まず、ナビエ・ストークス方程式の数値解を求める方法について、並列化の作業と評価に必要な程度に簡単に説明する。

外力の働かないナビエ・ストークス方程式は一般座標系の保存形式を用いて次のように表される。

$$D_t Q + D_x Q + D_y Q + D_z Q = 0$$

これを定常解を求める場合に近似因子分解法(IAF)により離散化すると、

$$L_x L_y L_z \Delta Q^n = RHS^n$$

$$RHS^n = D_x Q^n + D_y Q^n + D_z Q^n$$

と書ける。ここで $L_x$ 等はオペレータであるが、要素行列が5×5のブロック三重対角行列と考えてよい。RHSは右辺の値で、現時刻の空間差分から構成される。 $L_x$ 等の反転には、ブロック対角行列を直接に解く方法と近似的に対角化して解く方法等がある。また、現在多くのナビエ・ストークス解法はレイノルズ平均化した

ナビエ・ストークス方程式に基づいているため乱流モデルを用いている。

今回は2つのプログラムを用いて評価を行った。最初のプログラムは2つの領域を使用して各々に別の格子を用いたプログラムであり、これを以降MDと呼ぶことにする。もう一つは単一の格子系を用いたプログラムで、これを以降SDと呼ぶ。どちらの格子も構造格子であり、トポロジー的には直方体と同じである。MDでは二つの直方体が相互に交わっている。MDは乱流モデルを使わず、またブロック3重対角行列を直接解く方法を採用している。SDは乱流モデルとして代数モデルを用いており、対角化法を用いてスカラ三重対角行列を反転している。

SDの計算手続きを図3により説明する。

- (1) 右辺の対流項を計算する。
- (2) 乱流項を計算する。
- (3) 右辺の粘性項を計算する。
- (4) x, y, z方向の係数の計算と求解をそれぞれ行う。
- (5) 境界条件の計算を行う。

となる。MDは上述の手続きを二つの領域について行い、乱流項の計算がなく、領域間の境界

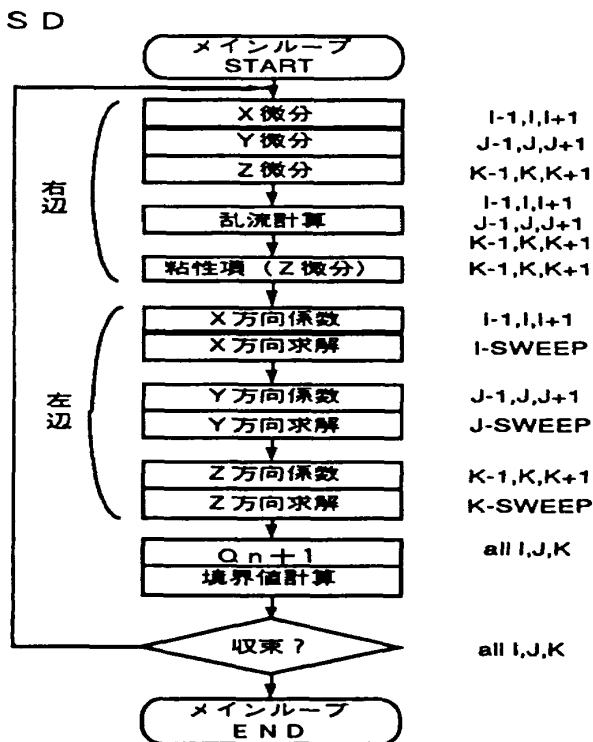


図3 SDプログラムの流れ

条件の計算が増えるだけで、基本的にはSDと同様なので、評価の手続きについてはSDを中心に説明する。

今回の評価では、実際のシミュレーション計算のほとんどの時間を占める繰返し部分のみを用いて評価した。

これらのプログラムの並列化は非常にストレートフォワードな方法で行うことが可能である。そのためまず、アレイデータの分割を行う。なお今後空間座標(x, y, z)と計算機上のアレイデータおよびDOループのインデックス変数(I, J, K)が各々対応しているものとする。アレイデータの分割の前に使用するPE台数を指定する。

140台のPEを使用することを  
IXOCL PROCESSOR PEN(140)  
と表現する。(!XOCLは並列化の指示行であることを意味する)

物理量やメトリック量が格納されているアレイデータをz方向に分割するには、

```
DIMENSION QG(1001,280,140,5),
           QL(1001,280,140,5)
!XOCL GLOBAL QG(:, :, / (PEN), :)
!XOCL LOCAL QL(:, :, / (PEN), :)
```

と表現する。このように表現するとQGはグローバル空間に、QLはローカル空間に配置される。

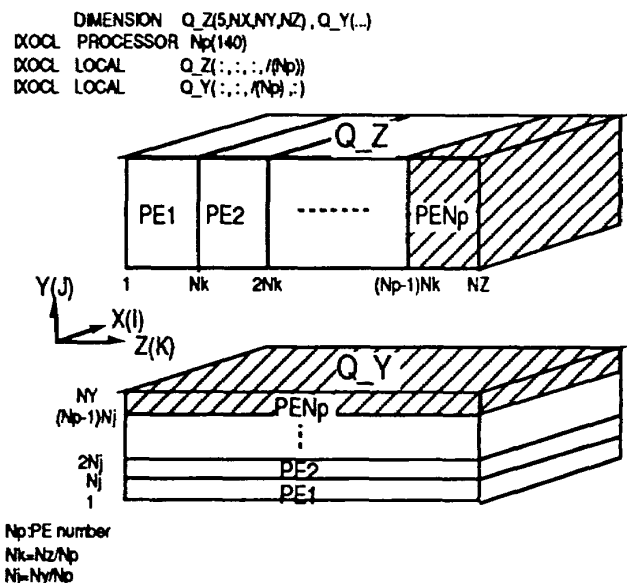


図4 配列の分割

後述するアレイデータの分割方向を替えた場合の配置は図4のQ\_\_Yの様に分割される。ここで、Q\_\_ZはZ方向に分割してPEに配置し、Q\_\_YはY方向で分割してPEに配置している。

次に、この分割に整合するように手続きを分割する。最も単純な行列の和で説明すると以下のようになる。行列A、B、Cをローカル空間に配置し、その分割を

```

          DIMENSION A(N1,N2,N3),
                    B(N1,N2,N3),
                    C(N1,N2,N3)
!XOCL   LOCAL   A(:, :, / (PEN)),
                    B(:, :, / (PEN)),
                    C(:, :, / (PEN))

```

とする。行列の和  $A = B + C$  は

```

!XOCL   SPREAD DO / (PEN)
          DO 10 K=1,N3
          DO 10 J=1,N2
          DO 10 I=1,N1
          A(I,J,K)=B(I,J,K)+C(I,J,K)
          10 CONTINUE
!XOCL   END SPREAD

```

と表現すればよい。これにより、DO 10 K=... のループは、PE1ではKは1からN3/PENまで、PE2のKはN3/PEN+1からN3/PEN\*2までというように各PEでN3個の処理を分割して並列実行される。CFDコードはこのような単純なDOループだけではないが、基本的にはこのようなパターンで表現することができる。

こうすることにより各PEは、自分自身のグローバル空間又はローカル空間にあるデータだけを使用する処理部分では、他のPEとは独立に計算を進めることが可能である。SDの例(一般的なナビエ・ストークスコードも同様)では、z微分を含まないRHSの計算、 $L_x$ 、 $L_y$ の反転、 $\Delta Q$ の加算の部分がこれに相当する。

他の部分では、自身のPE以外にあるPEのデータを必要(z微分を含む項のRHSの計算)としたり、z方向に依存性のある計算(乱流モデルの計算、 $L_z$ の反転)を行っている。

これらの部分を処理するには二通りの方法が考えられる。一つは図4に示すように、データ分割の方法をこれまでのz方向の分割からy方

向の分割に変えて、一斉にデータ転送を行ってから計算する方法である。この方法は決ったパターンで処理するために一斉に同期をとって行うことが可能である。また、クロスネットワークの特性を利用した転送が行え、ネットワーク上でデータ同志がぶつかりあうことが少なく効率的な転送が可能となる。

もう一つの方法は、データ分割はそのまま、各々のPEの中で処理を進め、となりのPEのデータが必要となる都度データを転送する方法である。z微分を含む項ではz方向の僅かのデータが必要なだけでデータ転送量も少ないが、z方向に依存性のある計算の実行では、これはいわゆるウェーブフロント法になる。 $L_z$ の反転等でこの方法を用いるとプログラミングはかなり技巧的になる。

後者の方法は、並列計算の技術としては興味があるが、不確定要素による性能低下が生じる可能性があり、また性能評価を困難なものとする。各PE毎にばらつきが生じると同期のためのオーバーヘッドが大きくなり並列の実効性能もそれほど上がらないことが予想される。数値風洞のPEは強力なベクトル計算機であることとクロスネットワークの転送能力が高いことを考慮すれば、各PEが一斉にデータを交換し、そのデータに対して演算を施す前者の方が優れているといえる。即ち、IAFのような陰解法を用いて三次元ナビエ・ストークス方程式の定常解を効率的に求めるために、航技研が数値風洞にこのようなハードウェアを採用した理由の一つである。これらのことによりMD、SDの並列化にあたっては前者の分割、転送方法を採用した。

以上をまとめると、並列化したSDのプログラムの骨格は図5のように表せる。このような並列化を行うと、処理に要する時間の見積りに不確定要素のはいる余地が少なくなる。即ち、各計算処理部分(図5の微分項、行列解法の部分)は各々のPEで閉じて実行が可能なので、通常のベクトル計算機の性能評価方法を採用することができる。次にデータ転送部分は、各並

列計算部分が終了後に同期をとって転送を行うことにすれば、転送に必要な時間の見積り誤差は小さくなる。ただし、数値風洞は非同期転送機能を備えており、事前にデータ転

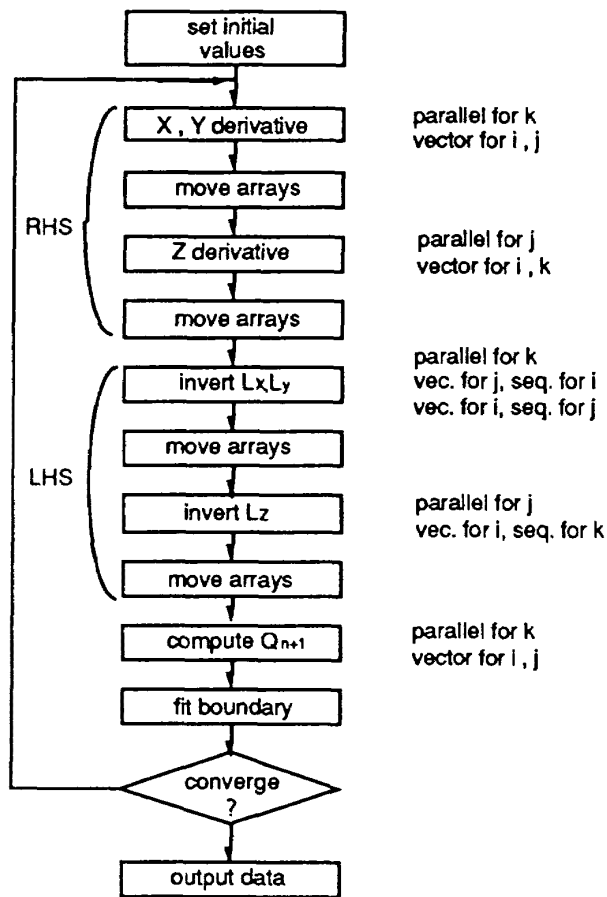


図5 並列プログラムの流れ

送しておくことも可能である。この場合、主記憶上でデータ転送と演算との間で衝突が起こる可能性がある。その影響はさほど大きくないと考えられるが、これを正確に評価することは不可能に近い。

MDの並列化もSDの並列化と同様に行うことが可能である。

SD、MDの評価にあたってはMDの評価データを基に行ったので、まずMDの評価について説明する。

140台のPEを最大限に利用するという観点から、格子数を各々  $501 \times 140 \times 280$ 、 $401 \times 140 \times 140$ とした。また並列化の方法の説明からわかるように、計算部分は各PEを独立したベクトル計算機として評価した。この評価にはベクトル計算機をシミュレーションするソフ

トウェア「VTAP」<sup>[7]</sup>を用いて行った。VTAPによる評価は以下の手順により行う。ベクトル化の対象となる全てのDOループに着目する。VP400用にコンパイルされたDOループのオブジェクトプログラムをVTAPの入力とする。VTAPのパラメータとしてPEのハードウェアデータを設定する。これによりVTAPで与えられたベクトル長に応じた実行時間を求める。この作業をMDの測定対象の全てのDOループについて行う。

次に転送に要する時間見積りについては、ソフトウェアによるネットワークシミュレータを作り、リード/ライトの典型的なパターンについて、転送時間の見積りを行った。この結果、今回採用した並列化で現れるデータ転送のパターンでは約90%の転送効率を得られた。今回評価を行うにあたっては、主記憶、ネットワークの衝突などでこの値よりも若干悪くなることを想定して、転送効率を80%として評価した。またバケットなどの生成に必要なソフトウェア走行時間も当然考慮に入れている。またMDにおける、二つの領域間の境界値を求める場合に現れる間接アドレス参照のパターンではネットワーク効率として、シミュレータによる評価値8%を採用した。

以上の結果MDでは、1回のイタレーションの間で、

- (1) 計算時間：1,800m秒
- (2) 転送時間：1,410m秒
- (3) 非同期転送可能な時間：185m秒
- (4) 浮動小数点演算数：153GFlop
- (5) 性能：50.7GFLOPS

の実効性能予測値が得られた。この値は元の問題規模である  $123 \times 29 \times 30$ 、 $71 \times 36 \times 20$ の格子数でVP400により実行した場合の性能、434MFLOPSの117倍である。但し転送時間の内、間接アドレス参照の部分が804m秒(57%)ほどある。

SDの評価をするにあたっては、MDの各ループをVTAPで計測して得られたPEのVP400に対する相対性能値を採用した。即ちM

DのほとんどのループでPEがVP400よりも1.4倍以上速い性能値が得られた。このため、SDの各ループもPEがVP400よりも1.4倍速いとして評価を行った。またクロスバネットワークの転送効率もMDと同じく0.8にした。この結果、SDに対しては、1回のイタレーションの間で、

- (1) 計算時間：1213m秒
- (2) 転送時間：195m秒
- (3) 非同期転送可能な時間：2m秒
- (4) 浮動小数点演算量：87.2GFLOP
- (5) 性能：62.0GFLOPS

の性能が得られた。この値は元の問題規模である $191 \times 33 \times 24$ の格子数で、VP400により実行した場合の性能、391MFLOPSの158倍である。

## 6. まとめ

CFDコードを用いてベクトル並列計算機である数値風洞の性能評価を行った。対象としたプログラムは三次元ナビエ・ストークスコードの内、単一の格子(SD)と、2つの格子(MD)を用いた二種類である。

数値風洞では、データの分割とループの分割実行を指示する非常に単純でストレートフォワードな並列化で実行可能であり、計算の実行とデータの転送を明確に区別した並列化が可能である。

この結果、机上評価ではVP400の158倍と117倍の実効性能が得られた。若干楽観的な評価かも知れないが、数値風洞の当初の目的である、「CFDコードでVP400の実効速度の100倍以上」は達成できると確信している。

なお、本報告は平成3年2月より開始された航技研・富士通の共同研究「数値風洞の開発研究」に基づいている。

## 参考文献

- 1) 岡田信、高村守幸：“CFD向け並列計算機のソフトウェア”、航空宇宙技術研究

所特別資料、SP-13, P109, 1990

- 2) 三好甫、吉岡義朗、池田正幸、高村守幸：“数値風洞のハードウェア”、航空宇宙技術研究所特別資料、SP-16, P99, 1991
- 3) 福田正大、末松和代、土屋雅子、大空瞭、工内隆、坂本喜則：“数値風洞のオペレーティングシステム”、航空宇宙技術研究所特別資料、SP-16, P107, 1991
- 4) 福田正大、中村孝、吉田正廣、岡田信、中村修一：“数値風洞の言語処理ソフトウェア”、航空宇宙技術研究所特別資料、SP-16, P115, 1991
- 5) 三好甫：“航技研超高速数値風洞(UHSNWT)の構想”、航空宇宙技術研究所報告、TR-1108, 1991
- 6) 吉田正廣、中村孝、福田正大、岡田信、中村修一：“NWT向け並列Fortranプログラミングについて”、航空宇宙技術研究所特別資料、SP-, 1992, to appear
- 7) 三好甫、島間晴夫、軽部行洋：“ベクトル計算機のベクトル処理に対するソフトウェアシミュレータVTAP”、航空宇宙技術研究所報告、TR-1094, 1991