

CFDの推進に必要な計算機性能

三 好 甫*

Computer Requirements for Further CFD Development

by

Hajime MIYOSHI

National Aerospace Laboratory

ABSTRACT

It is well known that the rapid development in computational fluid dynamics (CFD) has followed the development of powerful computers.

In this report, we show firstly that the same situation is also recognized at NAL.

Secondly, we give the requirements for computer power (processing speed and main memory capacity) that is necessary for the CFD to play important roles in the future R&D of aerospace planes and innovative aircraft, and we point out that today's supercomputer architectures are not likely to attain the processing speed required and that more parallelism should be developed.

Finally, we discuss several important items in designing the target computer; bipolar technology-based processing element vs. MOS technology based processing element, shared memory vs. distributed memory, interconnection network's topology, software technologies' cost and running cost of the target computer and so on.

1. はじめに

近年の計算空気力学(以下CFDと略記する)の発展には真に目覚ましいものがある。1983年に第一回が開催され、昨年で7回を数えるに至った本シンポジウム論文集の内容を逐年毎に追って見るだけでもこの発展は一目瞭然である。この発展を支えた要因として、計算法、複雑形状廻りの格子形成法、計算結果の可視化技術等の発展が挙げられるが、最大の要因として、この間になされた計算機性能の大幅な向上を挙げるのに異論をはさむものはないであろう。

現在迄のところ、計算機の性能とCFDの間には、計算機性能の大幅な向上がCFDの研究者および技術者に強い刺激を与え、それによりCFDが発展し、

この発展がまた計算機性能の大幅な向上を促すという良好な関係が持続してきている。この関係を今後とも持続させ、航空宇宙技術の研究開発にCFDを役立たせることは、CFDおよび計算機工学に携わる研究者および技術者にとって重要な課題であろう。

本稿は先ず、航技研においてCFDの発展に大きな刺激を与えた2つの計算機について述べる。次に、CFDの現状認識から、CFDが今後の我が国の航空宇宙技術の研究開発を強力に推進する基盤技術となる為に必要な計算機性能(処理速度と主記憶容量)を明確にする。

最後に、この計算機の近未来における実現可能性と実現において検討すべき重要事項について論ずることとする。

2. 航技研における計算機の歴史 とCFDの発展

航技研の計算機の性能は1960年、Burroughs Datatron-205の設置以来、システム代替毎に向上してきたが、性能向上に一時期を画し、CFDの発展に大きな刺激を与えたのは2つのベクトル計算機、FACOM-230-75APとFACOM-VP 400である。

2.1 FACOM-230-75APの導入とCFDの発展

FACOM-230-75APは1972年末より航技研と富士通の間で検討が開始され、その検討の中で方式上、構成上の多くの変遷を経た後、1977年に完成、航技研に導入された我が国で初めてのベクトル計算機であり、我が国のスーパーコンピュータ開発の引金となった計算機である。そのハードウェア仕様の概略は下記のとおりである。

(i) 汎用大型計算機FACOM-230-75と共に非対称な密結合多重計算機を構成する。

(ii) マシンサイクルタイムは90nsであり、加算(単精度22MFLOPS)、乗算(単精度11MFLOPS)および論理演算の3本のパイプラインをもつ。

(iii) データ形式は、固定小数点および単精度、2倍精度、4倍精度浮動小数点データである。

(iv) 256個の汎用レジスタ、1792語のベクトルレジスタ、2K語*のキャッシュメモリをもつ。

(v) 主記憶は1K語**(1語36bit)で、32ウェイトインターリーブ構成である。

我が国初のベクトル計算機の方式は、現在のベクトル計算機と比較すると、演算パイプラインが並列動作しないことおよびスカラー演算性能が相対的に低いこと等の欠点をもっていたが、それでもCFDプログラムの処理においてピーク処理速度の平均約1/3の実効速度を実現した。これは航技研の以前の計算機の処理速度に比較して20~30倍の高速性の実現であり、航技研のCFD研究の推進に多大の刺激を与えた。

FACOM-230-75AP+FACOM-230-75システムとそれに続く、FACOM-M380+FACOM-M180IIAD×2システムの設置期間(1978~1986年)において航技研のCFDは発展し、遷音速微小擾乱ポテンシャル方程式および完全ポテンシャル方程式を用いた翼型、主翼および主翼-胴体結合体廻りの外部流の数値シミュレーション、航空エンジン内部流の数値シミュレーションおよびレイノルズ平均ナビエ-ストークス方程式(以下ReAvNS方程式と略記)を用いた翼型廻りの数値シミュレーションが盛んに行われた。また、これらの数値シミュレーションの積み重ねおよび数値シミュレーション結果と風洞等試験データの比較検討に基づいて、アプリケーションソフトウェアの開発も大いに進捗した。

FACOM-230-75APにおいて特記すべきは

(i) 連続ベクトルのみならず、ストライド付ベクトルおよびリストベクトルのロード/ストアおよびベクトル演算を可能としたこと

(ii) DOループ中のIF文のベクトル処理を可能とするベクトルギャザー/スキップ命令をもち、Fortranでこの処理を可能としたこと

(iii) 外部記憶(磁気ディスク)とのファイル入出力において、バッファなし、並列入出力を可能としたこと

である。これらの機能は現在のベクトル計算機では標準機能となっているが、当時は先駆的な機能であった*。

2.2 FACOM VP 400の導入とCFDの発展

FACOM-230-75APの成功により、航技研はベクトル計算機の将来性と計算機性能の劇的な向上がもたらすCFDの発展に確信を持った。1978年、航技研はReAvNS方程式による主翼の数値シミュレーションを1時間程度で処理することを目標として、ピーク速度1GFLOPS、主記憶容量128~256MBのベクトル計算機の開発を内外各社に呼びかけた。これに応じて複数社が提案を行い、目

* 1K語=1024語 ** 1M語=1024語

* CRAY社のベクトル計算機が(i)および(ii)の機能を全て備えたのは1983年に発表された新型のCRAY-XMPにおいてである¹⁾。

標性能の実現を目指して航技研との検討を開始した。VP 400はこの検討の中から生まれた計算機である。富士通側との検討において、1980年初頭、航技研はVPシリーズの最上位機種、VP 200をベースとしてCFD用1GFLOPS 計算機の容易な構築が可能であることを提案した。

その提案は

(i) VP 200のスカラ処理系、メモリ系に手を付けず、パイプライン演算器をVP 200の2倍置きVP 200の大容量ベクトルレジスタのデータ転送能力の強化を行えば、少ない開発コストで1GFLOPSの計算機が容易に構築できる。

(ii) この様なベクトル計算機は

- (イ) ベクトル化率が低い
- (ロ) ベクトル長が短い
- (ハ) DOループの入力ベクトル数に対してベクトル演算数が少ない

プログラムに対しては無力であり、VP 200に比較して実効性能は殆ど向上しないが、上記(イ)~(ハ)と正反対の性質をもつプログラムに対しては、(i)の方策は処理能力向上に対して有効である。

(iii) CFD 計算法とCFDプログラムおよびVP 200の構成と各種タイミングの分析から、航技研が提案した計算機はCFDプログラムに対してVP 200の3割ないし8割の性能向上が予測できる。従って、この計算機はCFDプログラムに対しては真にコストパフォーマンスの良い計算機になるであろう。

この提案は富士通側側の受け入れる所となり、80年半ばにはほぼハードウェアの仕様が固まり、

VP 400が誕生した。航技研の提案が正しかったことはVP 400が航技研に導入された段階で実証された。その実例の一部を表に示す。

VP 400はVP 200と共に1986年度末航技研に導入され、それ迄の航技研主力計算機であったFACOM M 380の約20~70倍の性能をCFDプログラムの処理において発揮し、航技研と航空産業界におけるCFDの発展にとって大きな刺激となった。

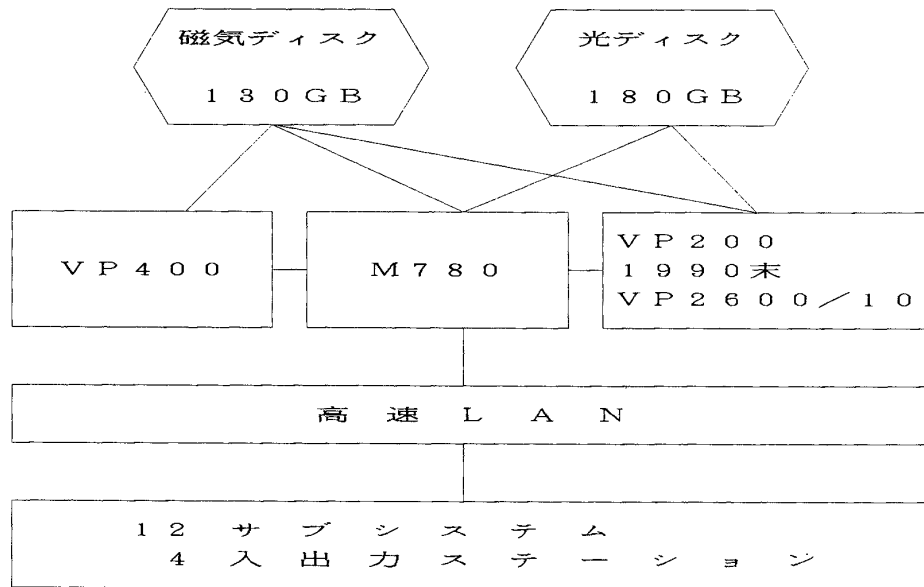
図1にVP 400とVP 200を中心とする航技研NSシステムのブロック図とVP 400ハードウェア仕様的一端を示す。

NSシステムの導入により、航技研のCFDはポテンシャル方程式に基づくものから、ReAvNS方程式に基づくものへと一変した。NSシステム導入後、航技研で行われたCFD研究の代表的なものをいくつか示す。

- (i) 主翼-胴体結合体およびクリーンな全機廻りに対する遷音速流、超音速流および極超音速流の数値シミュレーション
- (ii) 実在気体効果を考慮したクリーンな全機廻りの極超音速流の数値シミュレーション
- (iii) 低速流、超音速流および極超音速流における大規模剥離流れの数値シミュレーション
- (iv) 翼型および主翼廻りの非定常数値シミュレーション、特にフラッタ解析
- (v) ヘリコプターのロータおよびATPプロペラの数値シミュレーション
- (vi) 各種エンジン要素廻りの内部流の数値シミュレーション

表1 VP 200とVP 400のCFDプログラム処理速度比較

計算機 プログラム	VP 200 (単位・秒)	VP 400 (単位・秒)	VP 200 / VP 400
NS LU分解	75.70	38.37	1.97
NS REAL Dy	96.40	85.25	1.13
PNL	133.65	95.64	1.40
ZEBRA FP	34.28	19.64	1.75
NS BLTRI	349.68	212.35	1.65
NS Dyag	86.34	55.40	1.56



VP400 : マシクロック時間 13.5 ns 主記憶 1 GB
 VP200 : マシクロック時間 14 ns 主記憶 128 MB

図1 航技研数値シミュレータ (NS) システム

(vii) 極超音速燃焼の数値シミュレーション

以上は全てReAvNS方程式 (iv)では一部オイラ方程式を用いたものを含む)を用いた数値シミュレーションであるが、この他に

(viii) パネル法による定常および非定常の全機廻りの流れの数値シミュレーション

(ix) 高層大気における2次元および3次元形状のモンテカルロシミュレーション

等も重要な研究である。これらの数値シミュレーション技術に基づいてCFDアプリケーションプログラムが開発され、これらは

(i) スペースプレーン、HOPE等宇宙往還機の研究

(ii) 省エネルギー航空機の研究

(iii) 高バイパスファンエンジンの研究

(iv) 超音速機用エンジンの研究

(v) スクラムエンジンの研究

(vi) YXX等民間輸送機の国際共同開発

等に盛んに使用されている。また、共同研究等により広く航空産業界においても航技研で研究開発された数値シミュレーション技術とCFDアプリケーションソフトウェアが使用されるに至っている。

表2に典型的なCFDプログラムのVP400にお

表2 VP400 CFDプログラム処理速度

項目 プログラム	CPU時間 (秒)	MFLOPS値
NS TVDSD	2268	321.87
NS TVDMD	1874	439.85
NS LU	1446	643.85
REAL G	1161	257.54

ける処理速度をMFLOPS値で示す。

航技研における計算機の歴史とCFD研究発展の因果関係を考察することから以下の結論が引き出せる。

(i) CFDの急激な発展を導き出すためには計算機性能の数十倍以上の向上が必要である。

(ii) 新方式の計算機の開発とその利用の推進には強力なリーダーシップが必要である。

(iii) 新方式の計算機の導入にはコストパフォーマンスの劇的向上が伴わなければならない。

(iv) 計算機性能の大幅な向上があって初めて、CFDの航空機、宇宙往還機およびその推進系の研究開発現場における利用に拍車がかかり、開発現場からのフィードバックが新しい数値シミュレーション手法および従来は手がつけられていなかった対象に数値シミュレーション技術を適用しよう

とするCFD技術者および研究者の意欲をかきたてる結果となる。

(v) 計算機性能の大幅な向上には計算機方式の変革が必要である。また、新方式の計算機の効果的な使いこなし技術の習得には多少の労力を必要とし、使用環境条件の整備には多少の年月を必要とするが、これを克服するためには技術者および研究者の努力が必要である。

3. 今後のCFD推進に必要な計算機性能

CFDは今後の我が国における

- (D1) スペースプレーン, HOPE 等宇宙往還機とその推進系
- (D2) HST, SST 等とその推進系
- (D3) 省エネルギー航空機等民間輸送機
- (D4) 航空機等の国際共同研究

等の研究と開発のための中核的な基盤技術となることを期待されるまでに至った。この期待に応えるためにCFDは数年以内に以下に挙げる問題を解決する必要がある。

(R1) 現在クリーンな全機, 或いはそれと同程度の複雑形状の数値シミュレーションには数時間ないし10時間程度の時間がかかっているため, 1研究開発テーマ当たり年間数十ケースしか処理できない。このレベルの数値シミュレーションを用いて航空機等の空力設計に関するパラメトリックスタディを行う為, 1研究開発テーマ当たり年間数千ケースを適切なターンアラウンド時間内で処理する。

(R2) 附属部付完全全機等を対象とする複雑形状廻りの数値シミュレーションを適切なターンアラウンド時間内で処理する。

(R3) 航空宇宙技術の研究開発にとってリアリティのある物体を対象として, 乱流および燃焼現象等の物理現象の高精度計算モデルを確立する。

上記(R1)~(R3)を計算機に対する要求に翻訳するため, 幾つかの前提を置く。先ず数値シミュレーションに必要な格子点数についての前提を示す。

前提1. 現在, 航技研では0.2M点*~3M点程

* M点=百万点

度の格子点を用いてクリーンな全機等の数値シミュレーションを行っているが, これを1M格子点数数値シミュレーションで代表させる。

前提2. 附属物付完全全機の数値シミュレーションには5M点~15M点の格子点数が必要とされている。これを10M格子点数数値シミュレーションで代表させる。

前提3. 種々の物理現象の高精度数値シミュレーションに必要な格子点数は, 対象となる現象毎に異なる。ここではCFDにとって最も重要と思われる乱流のLES(Large Eddy Simulationの略)を主翼等を対象として行う場合で代表させる。文献²⁾で種々の仮定を設けて, LESに必要な格子点数を計算している。これに従えば, 典型的な主翼等のLESに必要な格子点数は最低の場合でも150M点程度となる。そこで(R3)の数値シミュレーションを150M格子点数数値シミュレーションにより代表させる。

次に計算時間に関する前提を示す。

前提5. 1M格子点数数値シミュレーションを1研究開発テーマ当たり年間数千ケース処理するためには, 1ケース当たりの計算時間は, 10分程度でなければならない*

前提6. 10M格子点数数値シミュレーションの計算時間は1時間程度である必要がある。即ち, 一晚(10時間位)でポーラ曲線がかける様にする。

前提7. 150M格子点数数値シミュレーションを一晚で処理できる様にする。

前提8. 航技研の経験では, 1M格子点数数値シミュレーションに要する時間は流れの条件, 形状の複雑度等により変化するが, 概ねVP400で数時間~10時間程である。そこで1M点のシミュレーションに必要な計算時間はVP400で10時間かかると定める。格子点数の数と計算時間の関係は線型ではないが, 計算法の改良等も考慮して線型とす

* 計算時間に対する要求は各機関でなされているが, それにはかなり幅がある。即ち, 最短の要求は2~3秒, 長いもので30分というものである。例えばNASA Ames R.C.は空力設計のための計算時間は10分と言っている³⁾。この計算時間に対する要求には2つの意味がある。1つは早く結果を知りたいということ。今1つは1日数回位は計算を実行したいということである。

る。即ち、1M点の数値シミュレーションの計算時間がVP 400を用いて10時間であれば、10M格子点数シミュレーションの計算時間はVP400で100時間かかるとする。

最後に主記憶量についての前提を示す。

前提9. これ迄の調査ではReAvNS方程式を用いた数値シミュレーションにおいて、大部分のプログラムの1格子点当たりの使用データ量は200バイト~400バイトの間に分布している。この変動はプログラムのプログラミングスタイルによるものである。但し、使用可能主記憶量が相対的に小さい場合には1格子点当たりのデータ量は小さくなる傾向がある。

前提1~9により(R1)~(R3)は

(R1') 主記憶使用量0.2GB~0.4GBの1M格子点数シミュレーションを10分間で処理する。即ち、VP400の60倍の実効CFDプログラム処理速度をもつ計算機が必要である。

(R2') 主記憶量2GB~4GBの10M格子点数シミュレーションを1時間で処理する。即ち、VP400の100倍の実効CFDプログラム処理速度をもつ計算機が必要である。

(R3') 主記憶量30GB~60GBの150M格子点数シミュレーションを10時間で処理する。即ち、VP400の150倍の実効CFDプログラム処理速度をもつ計算機が必要である。

という計算機性能に対する要求となる。

VP400の50~150倍の実効処理速度の計算機の有力な候補として多数台(少なくとも100台程度以上)の計算機からなる並列計算機が挙げられる。こうした並列計算機にとっては、相対的に小さな問題を1ケース10分で処理するという要求は厳しい要求である。そこで(R1')の要求をゆるめて

(R1'') 0.2GB~0.4GBの1M格子点数シミュレーションを10時間で60ケース処理する。即ち、多重マルチジョブで処理する*。

と変更する。同様に(R2')も

(R2'') 2GB~4GBの10M格子点数シミュレ

ーションを10時間で10ケース処理する。と変更する。

(R1''), (R2'')および(R3'')の計算機性能に対するCFDからの要求を標語的に取りまとめると以下の様になる。

CFDが今後の我が国の航空宇宙技術研究開発の中核基盤技術となるためには、ここ数年以内に

(R) 主記憶容量は32GB以上、CFDプログラムの実効処理速度がVP400の100倍以上の計算機が必要である。

(R)の条件を満たす計算機は恐らく現在のスーパーコンピュータとは異なった方式のもとに実現されよう。方式の変更があれば、当然ながら計算機高度利用のためのプログラム技術、デバッグツールとその使用法、プログラムチューニングツールとその使用法等も変化する可能性が大である。また、プログラムの厳密な意味での移行性も損なわれる可能性がある。従って、計算機の方式の変更がしばしばというのは利用者にとって耐え難い事である。これを考慮して、計算機に対する要求として、今1つ

(R') (R)の要求を満足するために採用される計算機の方式は20世紀末には更に10倍の処理速度を達成可能なものであることが望ましいという要求を付け加えることにする。

4. 計算機技術の発展と目標計算機性能の実現の可能性

4.1 目標主記憶容量の実現

図2に最近のメモリ素子の進歩の状況を示す⁷⁾。メモリ素子の進歩は真に順調であり、予測通りに進んでいる。現在1MビットSRAMチップと4MビットDRAMチップが量産に入りつつあり、今後数年のうちには4MビットSRAMチップと16MビットDRAMチップが量産に入るであろう。

一方、最近発売が開始され、近く稼働が予定されている国産スーパーコンピュータFACOM VP2000シリーズおよびNEC SX-3シリーズは共に最大主記憶量が2GBであるが、これらにはそれぞれ1Mビットおよび256KビットSRAMチップが使用されている。

* 10分間で1ケースを処理できる計算機は当然1ジョブストリームの処理形態のもとに10時間で60ケース処理できる。

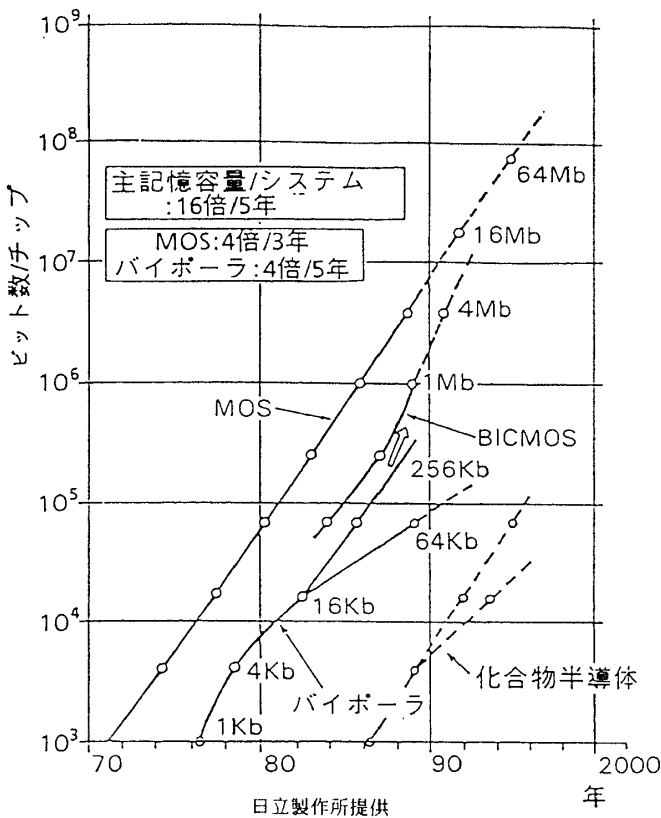


図2 メモリICの進歩

目標計算機の主記憶量は32GBであり、新発表の国産スーパーコンピュータの16倍の主記憶容量である。目標計算機の主記憶に1MビットSRAMチップを使用するとすれば、チップカウントは30万個以上に達する。また、4MビットDRAMチップを使用すればチップカウントは8万個程度に納まる。速度を考慮すれば目標計算機にはSRAMを使用すべきであろうが、1Mビットチップでは少々苦しい。しかしながら現在のスーパーコンピュータの主記憶実装技術を考えれば、これは不可能といった技術ではないであろう。勿論4MビットSRAMチップ或いは16MビットDRAMチップの量産時期迄計算機の実現を待てば32GBの主記憶の実現はより容易となるであろう。

主記憶実装の問題が解決した場合、SRAMを使用するかDRAMを使用するか、或いはSRAMを使用する事に決めた場合でも1Mビットチップを使用するか4Mビットチップを使用するかは素子の速度とコストおよび目標計算機の方式と主記憶階層構造を考慮して、これらが最適になる様になさなければならない。主記憶に使用する

チップの選択は、計算機の計画における最重要事項の1つである。

如何なる種類のチップを採用するにしても、目標計算機の主記憶量のここ数年以内における実現の可能性は大いに高いと考えて良い。

4.2 目標処理速度の実現

4.2.1 現在のベクトル計算機の方式による高速化の限界

図3に最近の商用スーパーコンピュータの主流であるベクトル計算機のピーク処理速度決定の基礎となるパイプラインピッチ時間の推移を示す。1990年以降のものは発表^{(4),(6),(7)}されている推定値である。図4に記憶構造から見た現在のベクトル計算機の方式を示す。また、表3と4に現在およびここ数年以内に発売予定のベクトル計算機の一覧を示す。

図3から明らかな様にパイプラインピッチ時間の短縮傾向は減衰しつつあり、数年以内に期待できるパイプラインピッチ時間は高々1~2nsである。従って、1本のパイプライン演算器のピーク処理速度は1~0.5GFLOPSとなり、乗算器と加算器を平行動作させて得られるベクトル計算機のピーク処理速度は高々2~1GFLOPSにしかならない。従って、なお一層の高速化を図る必要があるが、その為の標準的な高速化手法としては以下の様なものが考えられる。

(i) パイプラインの多重化

物理的に複数本のパイプラインを重ねて論理的に1本に見えるパイプライン演算器を構成する手法である。この手法を用いてパイプラインを4重化すればピーク処理速度は4倍になる。この手法は国産ベクトル計算機に採用されており、FACOM-VP400, FACOM-VP2600, NEC-SX-2, およびNEC-SX-3-14等は4重化されたパイプライン演算器を用いている。

(ii) 演算器の多重化

加算および乗算のパイプライン演算器をそれぞれ複数本並べ、並列動作させる手法である。この手法により、ピーク処理速度は平行動作可能なパイプライン演算器の本数倍だけ向上する。この手

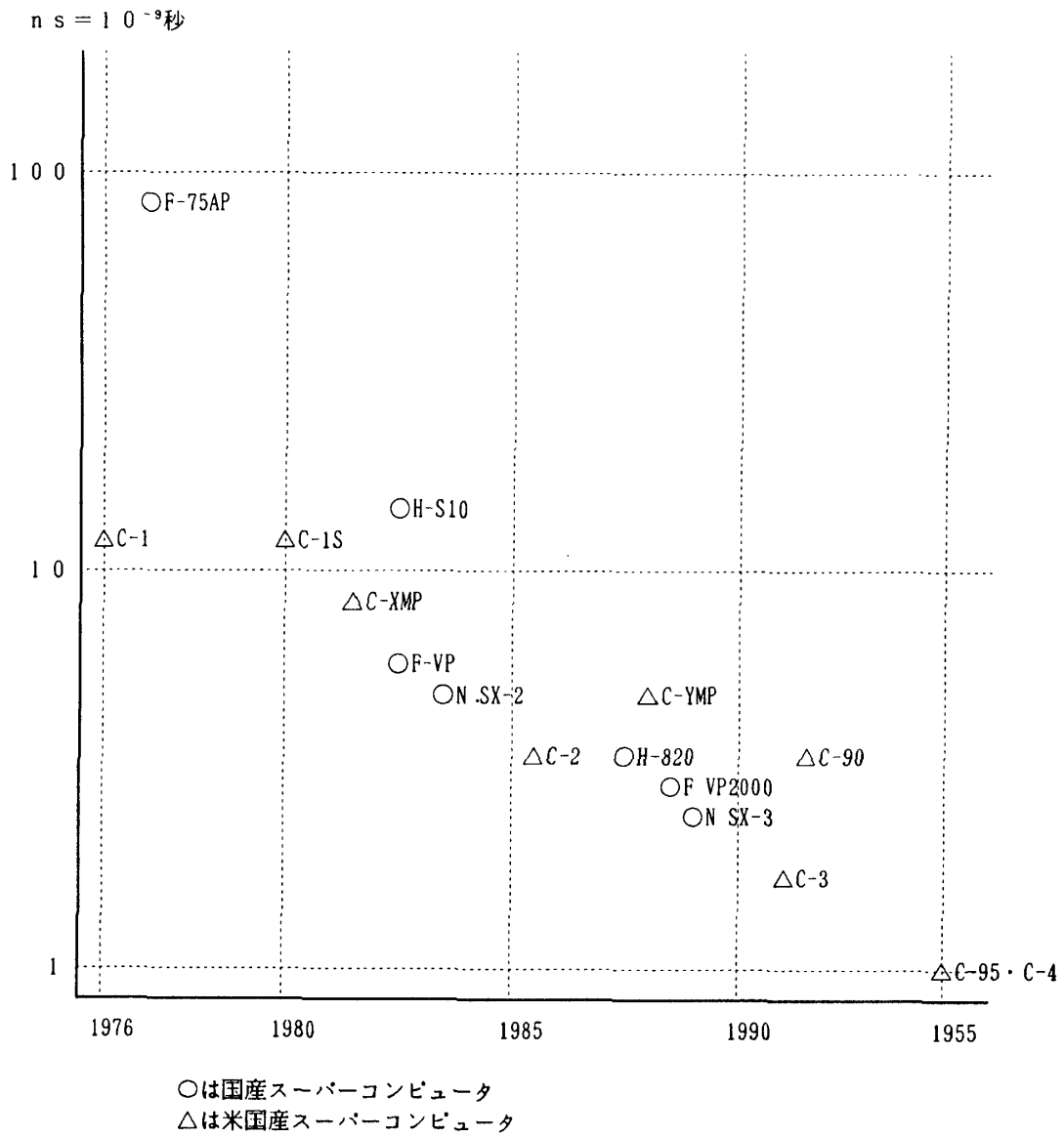
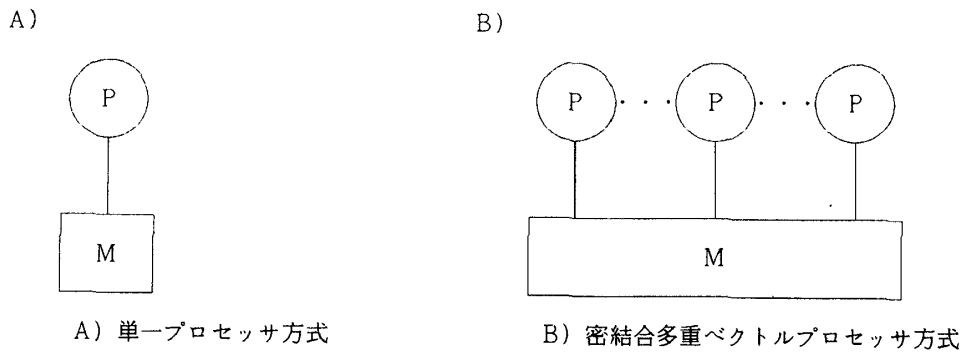


図3 スーパーコンピュータのパイプラインピッチ時間の推移



P：計算機または要素計算機

図4 現在のベクトル計算機の方式

表3 単一ベクトル計算機の諸元

機種名	発表年月	クロック時間 (ns)	ピーク性能 (GFLOPS)	主記憶容量(最大) (MB)
HITAC S820/80	87年 6月	4	3	512
FACOM VP2600	88年12月	3.2	5	2048
NEC SX-3/14	89年 2月	2.9	5.5	2048

表4 多重ベクトル計算機の諸元

機種名	発表年月	台数	クロック時間 (ns)	ピーク性能 (GFLOPS)	主記憶容量(最大) (MB)
NEC SX-3/44	89年 2月	4	2.9	22	2048
CRAY-YMP/832	88年 2月	8	6	2.67	256
CRAY-3	9X年	16	2	16	?
CRAY C90	9X年	16	4	16	?

CRAY-3、CRAY C90の諸元は文献4)による。

法はHITAC-S-810、NEC-SX-3等に採用されている。SX-3-14等では加算および乗算パイプライン演算器をそれぞれ2本もっている。

(iii) ベクトル計算機の多重化

図4のB)に示す様に、主記憶上に複数台のベクトル計算機を並べて1つのベクトル計算機を構成する手法である。この手法により、ピーク処理速度は構成ベクトル計算機の台数倍向上する。現在では2~8台の多重化が行われているが、数年以内に16台構成の多重ベクトル計算機が出現すると考えられている。

しかしながら、上記(i)~(iii)の手法を用いてピーク処理速度を向上させ、それによりベクトル計算機の実効処理速度の向上を図るといふ目論見は以下に示す様な制約がある為、その効果には限界がある。

(i) パイプライン立ち上り時間に起因する制約
パイプライン演算器には良く知られている様に、ベクトル演算命令が発信されてから最初の結果が得られる迄に時間がかかる。これをパイプラインの立ち上り時間という。今、時間の単位をパイプラインピッチ時間 t とし、処理すべきベクトルのベクトル長を l 、立ち上り時間を s 、パイプライ

ンの多重度を m とすると、パイプライン演算器がベクトル長 l のベクトル演算を終了する時間 T は

$$T = (s + l/m) \times t$$

となる。ピーク処理速度は立ち上り時間を無視した性能であるから、ベクトル長 l のベクトル演算をピーク処理速度で処理する時間は

$$T_p = l/m \times t$$

従って、 $E_p = T_p/T$ が多重パイプライン演算器の実効効率となる。パイプライン演算器の実効性能(FLOPS値)はピーク処理速度と E_p の積で示される。表5に $t=1\text{ns}$ 、 $s=10$ とした場合の E_p の値をパイプライン多重度とベクトル長に対して示す。

E_p の表からベクトル長が相対的に短い場合には、パイプライン多重化による実効処理速度の向上には限度があることがわかる。従って、投入されるハードウェア物量の効果的な使用という観点から、採用可能なパイプライン演算器の多重度はCFDプログラムのパイプライン多重化に対する適合性*

* 国産のベクトル計算機ではパイプライン演算器へのデータ投入が途切れない様に、ベクトル演算命令が、例えば5命令が発信され、さらにベクトルレジスタにオペランドが常に存在していれば、実行上のベクトル長は5倍になる。航技研のCFDプログラムの分析から、CFDプログラムの実行ではこの条件を満足する頻度はかなり高いことがわかっている。

表5 パイプライン多重化の効率 E_p

多重度 ベクトル長	1	4	8	16
128	0.928	0.762	0.615	0.444
256	0.962	0.865	0.762	0.615
512	0.981	0.928	0.865	0.762
1024	0.990	0.962	0.928	0.865

を考慮したとしても、4多重或いはせいぜい8多重程度迄であろう。

(ロ) 演算器の使用率の低下による制約

乗算および加算等のパイプライン演算器をそれぞれ複数本並べ、並列実行を可能とすることによりピーク処理速度は向上するが、実際のプログラムにおいて、多数の加算および乗算等のパイプライン演算器が全て動作する頻度は、並行動作可能なパイプライン演算器の本数の増加に伴って低下する。即ち、パイプライン演算器の使用率が低下し、ピーク処理速度の向上に実効処理速度の向上が伴わない結果となる。従って、独立に並行動作する加算および乗算パイプライン演算器の本数は投入するハードウェア物量の効果的な使用という観点から、せいぜい2本程度ということになる。

(イ) ベクトルレジスタのデータ転送能力不足による制約

今、ピーク処理速度 m GFLOPS、主記憶からベクトルレジスタへのデータ転送能力 n GBのベクトル計算機を考えると、ベクトルレジスタのデータ転送能力 P は下記の式を満足する必要がある。

$$P \geq (m \times 3 \times 8 + n) \text{ GB/秒}$$

m が大きく、例えば16, 32, 64となった場合、 P を満足する大容量のベクトルレジスタを構成することは少なくとも数年以内においては不可能といっても良い。現在、最新の国産ベクトル計算機における P の値は高々100GB/秒程度に過ぎない。この制約は、パイプラインの多重化による性能向上手法に対しても、演算器の多重化による性能向上手法に対しても、共に大きな制約となる。

(ニ) 主記憶のデータ転送能力不足による制約

ベクトルレジスタをもつピーク処理速度 m GFL

OPSのベクトル計算機において、主記憶のデータ転送能力の定量的な大きさの下限値に関する定説はない。この下限値はベクトルレジスタの容量と処理すべきプログラムの性格に依存して定まる。そこで、既存のベクトル計算機のピーク処理速度と主記憶データ転送能力（これを M とおく）の関係を借用して、議論を進めることにする。CRAY-YMPは高い主記憶データ転送能力を持ち、これがCRAY-YMPの大きな特徴でもあるが、この計算機におけるピーク処理速度と主記憶データ転送能力 M との関係は

$$M = (m \times 2 \times 8) \text{ GB}$$

である。NEC-SX-2, FACOM-VP 200等ではこの関係は

$$M = (m \times 1 \times 8) \text{ GB}$$

FACOM-VP 400, CRAY-1では

$$M = \left(m \times \frac{1}{2} \times 8 \right) \text{ GB}$$

という関係がある。今、CFDプログラムの実行においてVP 400の100倍の実効処理速度を持つベクトル計算機のピーク処理速度を仮に128GFLOPS（少々不足すると思われるが）としたとする。CRAY-YMP並みのデータ転送能力を要求したとすると、約2048GB/秒、NEC-SX-2並みのデータ転送能力を要求したとすると、1024GB/秒、大容量のベクトルレジスタを持たせて、FACOM-VP 400と同程度迄データ転送能力に対する要求を落としたとしても512GB/秒の主記憶データ転送能力がそれぞれ必要となる。

容量32GB、データ転送能力512GB/秒の主記憶をここ数年の内に構築することは、コストの制約を考えると主記憶素子技術の面から非常に無理

表6 並列計算機の諸元

機種名	台数	ピーク性能 (GFLOPS)	会社名	結合ネットワーク	使用素子技術
コネクションマシン	65536	16	シンキングマシン社	ハイパーキューブ	CMOS
ADENA	256	5	松下電気産業	ADENAキーブ	CMOS
QC DPAX	480	16	筑波大学	隣接結合	CMOS
NCUBE2	8912	18.5	NCUBE社	ハイパーキューブ	CMOS
Touchstone	2048	128	インテル社	隣接結合	CMOS
CRAY-4	64	128?160?	クレイコンピュータ社	?	GaAs?
SSI	64?	64?	SSI社	?	?
CRAY-C95	64?	256?100?	クレイ社	?	?

CRAY-4、SSI、CRAY-C95の諸元は文献4)、5)による。
QC DPAXとADENAは文献6)による。

があり、その様な主記憶の実現は著しく困難である。主記憶データ転送能力不足による制約は高速化手法(i)~(iii)全てに対して障壁となる*。

上記(i)~(iii)の手法の他に計算に出現する

$$A=B+C*D, A=(B+C)*D$$

(A, B, C, Dはそれぞれベクトル)

の様な複合演算に目をつけて、複合演算用パイプライン演算器を構成してピーク性能の向上を図る手法もあるが、複合演算の頻度が大でなければこの手法は実効性能を向上させる上で余り効果的とは言えない。

4.2.2 計算機方式の変更

前小節で述べた事から、現在のベクトル計算機の方式を変更せずに目標処理速度を達成することは非常に困難であるとの結論が導かれる。最大の障壁はピーク処理速度を支えるのに十分なデータ転送能力をもつ容量32GBの主記憶の構築の困難性にあり、次に、十分なデータ転送能力を持つ大容量のベクトルレジスタの構築の困難性がある。

また、演算器の多重化およびパイプラインの多重化による目標処理速度の実現も効率の劣化とデ

ータ供給能力の不足理由により、困難であることが示された。

そこで、計算機の方式を変更することにより目標処理速度を達成することを考える必要に迫られる。

最大の障壁が主記憶のデータ転送能力の不足にあることを考慮して、記憶構造の観点から計算機の方式変更を考える。

図5および6に主記憶構造から見た並列計算機の方式を示す。図5においては要素計算機Pは単一計算機であるが、図6においては図5の要素計算機が密結合の多重計算機に変更されているに過ぎない。

CFDプログラム処理においてベクトル処理が非常に有効に働いたというこれ迄の経験*、ベクトル演算器(パイプライン演算器)の物理的効率の良さ**、ベクトルレジスタ方式の優秀性およびベクトルレジスタに関するコンパイラ技術の蓄積があるという理由により、どの方式においても要素計算機はベクトルレジスタ付きのベクトル計算機と定めるのが妥当である。

* ここでは主記憶アクセスにおける干渉については述べなかったが、この干渉の存在は主記憶データ転送能力の不足に起因する実効性能低下を更に悪化させる原因となる。この問題は特にベクトル計算機の多重化による高速化手法において大きな問題となっている⁹⁾。

* 希薄気体のモンテカルロシミュレーションに対してはベクトル処理は連続気体に対する程の高効率性を発揮しないという例外はある。

** ゲート当たりの処理速度に関してベクトル演算器は効率が良い。

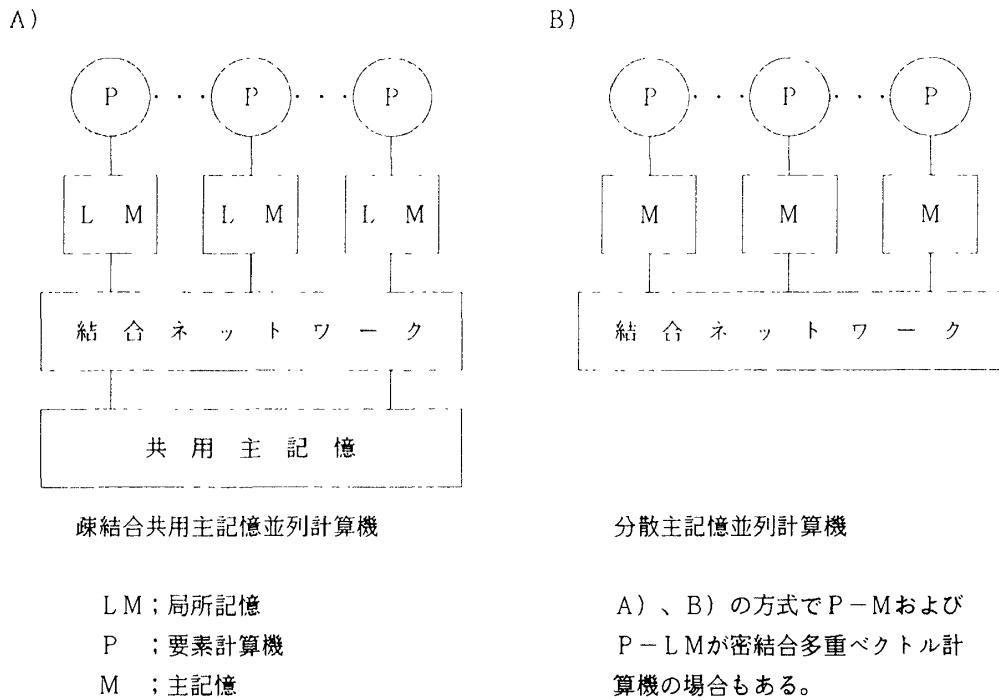


図5 並列計算機方式(その1)

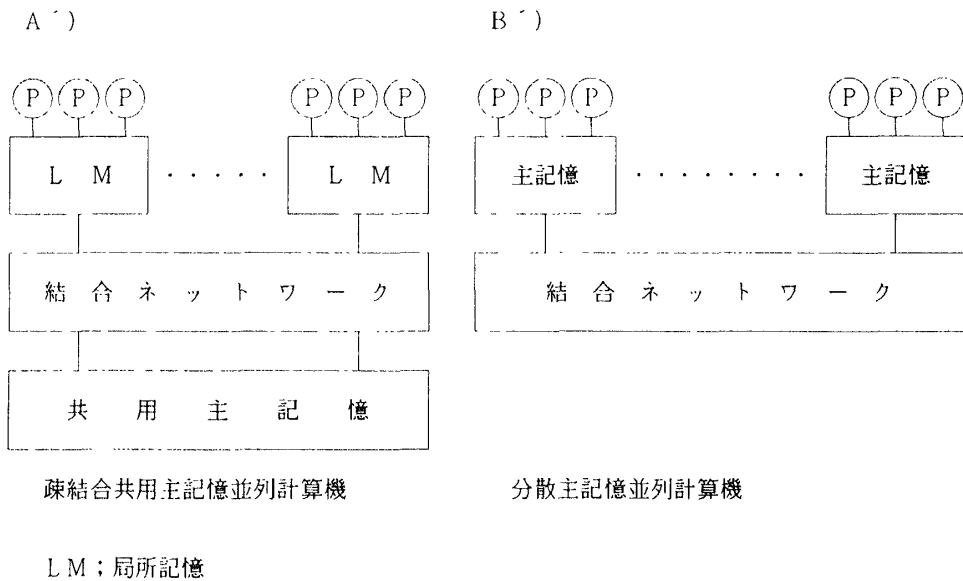


図6 並列計算機方式(その2)

AおよびA'の方式は大容量の高速局所記憶を置くことにより、共用主記憶と結合ネットワークのデータ転送能力に対する厳しい要求を緩和させる事を目指している。

局所記憶LMの容量は要素計算機の処理速度、結合ネットワークおよび主記憶のデータ転送能力等のハードウェアの性能に対するCFDプログラムのデータ転送要求とのバランスにより決められる。

この結果、LMの容量が余り大きくなる場合には、A或いはA'方式よりもB或いはB'方式の方が記憶素子節約の観点からは望ましいことになる。

BおよびB'の方式は局所記憶を廃止し、その代わりに主記憶を各要素計算機に分散配置する。この場合、各要素計算機に分散配置される主記憶の大きさは要素計算機の台数をPとすると $32GB/P + \alpha$ となる。ここで α はプログラムの手続き部分、

OSの常駐部分、他の要素計算機との領域オーバーラップ部分、或いはデータ受信部分等を格納するのに十分な容量である。

BおよびB'の方式は要素計算機に分散配置されている主記憶内にその要素計算機が処置において必要とするデータの殆どが存在する様に出来れば、主記憶転送能力の問題の解決可能性が見えてくる。

〔注意〕

ここでは計算機の方式分類を計算機の記憶構造の観点から行った。計算機の方式分類は他にも種々の観点から行われている。方式分類で有名なものに、計算処理における命令列とデータ列の数に着目した以下の4つの分類がある。

- (イ) 単一命令列, 単一データ列 (SISD) 方式
- (ロ) 単一命令列, 複数データ列 (SIMD) 方式
- (ハ) 複数命令列, 単一データ列 (MISD) 方式
- (ニ) 複数命令列, 複数データ列 (MIMD) 方式

SISD方式の計算機は1つの命令列を1つのデータ列に対して逐次適用して行く方式であり、単一計算機から成る汎用計算機はこの方式に属する。単一ベクトル計算機はベクトル命令を1命令と考え、ベクトルデータを1データと考えればSISD方式の計算機である。SIMD方式の計算機は同じ1つの命令列を多数のデータ列に独立に適用して行く方式であり、例としては、ILLIAC-IV等がある。単一ベクトル計算機はベクトルデータ列を多数の独立に処理可能な要素から成る複数のデータ列と考えれば、ベクトル命令を1命令と考えることによりSIMD方式の計算機と考えて良い。MISD方式の計算機には該当するものがない。しかしパイプライン処理を複数の命令処理と考え、ベクトルデータの要素を1データと考えれば、単一ベクトル計算機はMISD方式の計算機と考えられる。

MIMD方式の計算機は複数の命令列が複数のデータ列を処理する方式である。多重ベクトル計算機を含む多重計算機と言われているものは全てこの方式の計算機である。

SIMD方式の並列計算機は特殊な用途に限れば非常に魅力があるが、CFDへの適用を考えて、本稿で念頭に置く並列計算機の方式はMIMD方式とする。

計算機の方式分類は計算機の駆動方式の観点に立てば

- (イ) コントロール駆動方式
- (ロ) データ駆動方式
- (ハ) 要求駆動方式

の3つに分類できる。現存する殆どの計算機は(イ)の方式である。(ロ)および(ハ)の方式の計算機も最近研究が進み、試作機レベルではかなりな性能のものがある。例えば通産省大型プロジェクトの1つの成果であるSIGMA-Iは数値積分で170MFL OPSの性能を發揮している。

しかしながら筆者はCFDの様に規則性のある応用に対して(ロ、或いは(ハ)の方式が(イ)の方式に優るといふ根拠は見出し難いとする。従って、本稿で考える計算機の方式はコントロール駆動方式の計算機である。

4.2.3 計算機結合ネットワーク

並列計算機の結合ネットワークはそのトポロジー* (幾何学的形状)に限って見ても多数のものが提案されている。結合ネットワークのトポロジーは更に静的な結合ネットワークと動的な結合ネットワークの2つに分類される。静的というのは結合が固定されているものをいい、動的というのは結合が処理の実行時の要求に基づいて随時変更可能であるものをいう。

図7に静的結合ネットワークの例を示す。図中の○印は要素計算機を示す。これらの結合ネットワークは主としてBおよびB'方式の並列計算機に対して提案されているものである。

図中の完全結合においては、全ての要素計算機が直接結合されているので、全ての要素計算機対間におけるデータの転送は直接行われ、データの転送に関する競合がない。また、全ての要素計算機対間の距離は1である。以上述べたことから完全結合は最も望ましい結合であることが明らかである。完全結合の問題点は要素計算機の台数が増加すると、これを実現するための物量が急激に増加し、その実現が困難になることである。

* 結合ネットワークの選択において重要な要因にはトポロジーの他にネットワークの同期方式、制御方式および情報交換方式がある。

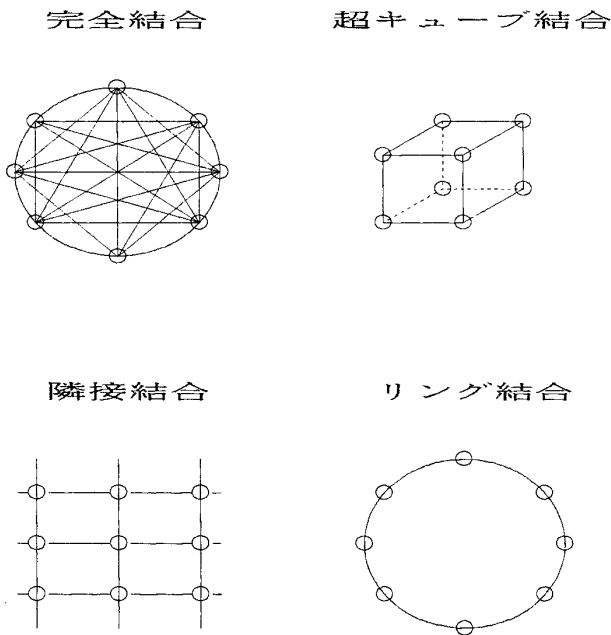


図7 並列計算機の静的結合ネットワークの例

超キューブ結合は 2^n 台の要素計算機を n 次元超立方体の頂点に配置する結合ネットワークである。要素計算機対間のデータ転送に関する距離は最短のものが1、最長のものが n となり、要素計算機対間の相対的位置関係により、データ転送に関する距離が異なる。また、或る要素計算機対間のデータ転送が他の要素計算機の配置されている超立方体の頂点を經由して行われる場合、データの転送は頂点に配置されている要素計算機の仲介を必要とするという不都合が生ずる。これを避ける為には、要素計算機とは別に各頂点にデータ転送制御用装置を配置する必要がある。それでも他の要素計算機対間のデータ転送との競合が立方体の頂点で発生することを避けることはできない。以上述べた超キューブ結合における問題点は隣接結合においてもリング結合においても同様に避けることはできない。

但し、要素計算機の台数が増加する場合、結合を実現するために必要な物量の増加は完全結合に比較してはるかに小さい。

隣接結合は2次元の格子点上に要素計算機を配置した結合である。1つの要素計算機は東西南北の4台の要素計算機とだけ直接結合されている。 $n \times m$ 個の要素計算機をこの結合ネットワークに

より結合した場合、要素計算機対間のデータ転送に関する距離は最短の場合1であり、最長の場合 $n \times m - 2$ である。

リング結合はリングの上に要素計算機を配置するもので、1つの要素計算機は左右の要素計算機とだけ直接結合されている。

隣接結合においてもリング結合と同様に東西および南北の両端をそれぞれ結合してトーラス状にする方が単なる隣接結合より一般に用いられている。

図8に動的結合ネットワークの例を2つ示す。図における結合はAまたはA'方式の並列計算機に対するものであるが、図中の $M_1 \dots M_m$ を $P_1 \dots P_n$ に取り替えればBまたはB'方式の並列計算機に対する結合ネットワークになる。

クロスバー結合はAまたはA'方式の並列計算機の場合、任意の要素計算機-メモリモジュール対間のデータ転送はデータ転送の実行時に直接結合されて実行される(BまたはB'方式の並列計算機の場合は任意の要素計算機対間のデータ転送)。また、データ転送が1対1で行われる場合、異なる対間のデータ転送との競合は生じない(多対1の場合には競合が生ずる)。この点に関してクロスバー結合は最も望ましい動的結合といえる。クロスバー結合の問題点は要素計算機の台数が増加するとクロスバー結合を構築するための物量が急激に増加して、その実現が困難となることである。

図8に多段結合ネットワークの一例であるオメガネットワークを示す。多段結合ネットワークは他にも間接超キューブネットワーク、バタフライネットワーク等多数のものが提案されている。多段ネットワークを構築するために必要な物量は要素計算機の台数を N (メモリモジュールの数も便宜的に N としておく) とすると $O(N \log N)$ である。これはクロスバー結合の構築に必要な物量が $O(N^2)$ であるのと比較して格段に改良されている。その代わりに、要素計算機とメモリモジュール間のデータ転送に関する距離はクロスバー結合の場合が1であったのに対して、 $O(\log N)$ となる。また、要素計算機-メモリモジュール対のデータ転送は他の要素計算機-メモリモジュール対のデータ転送

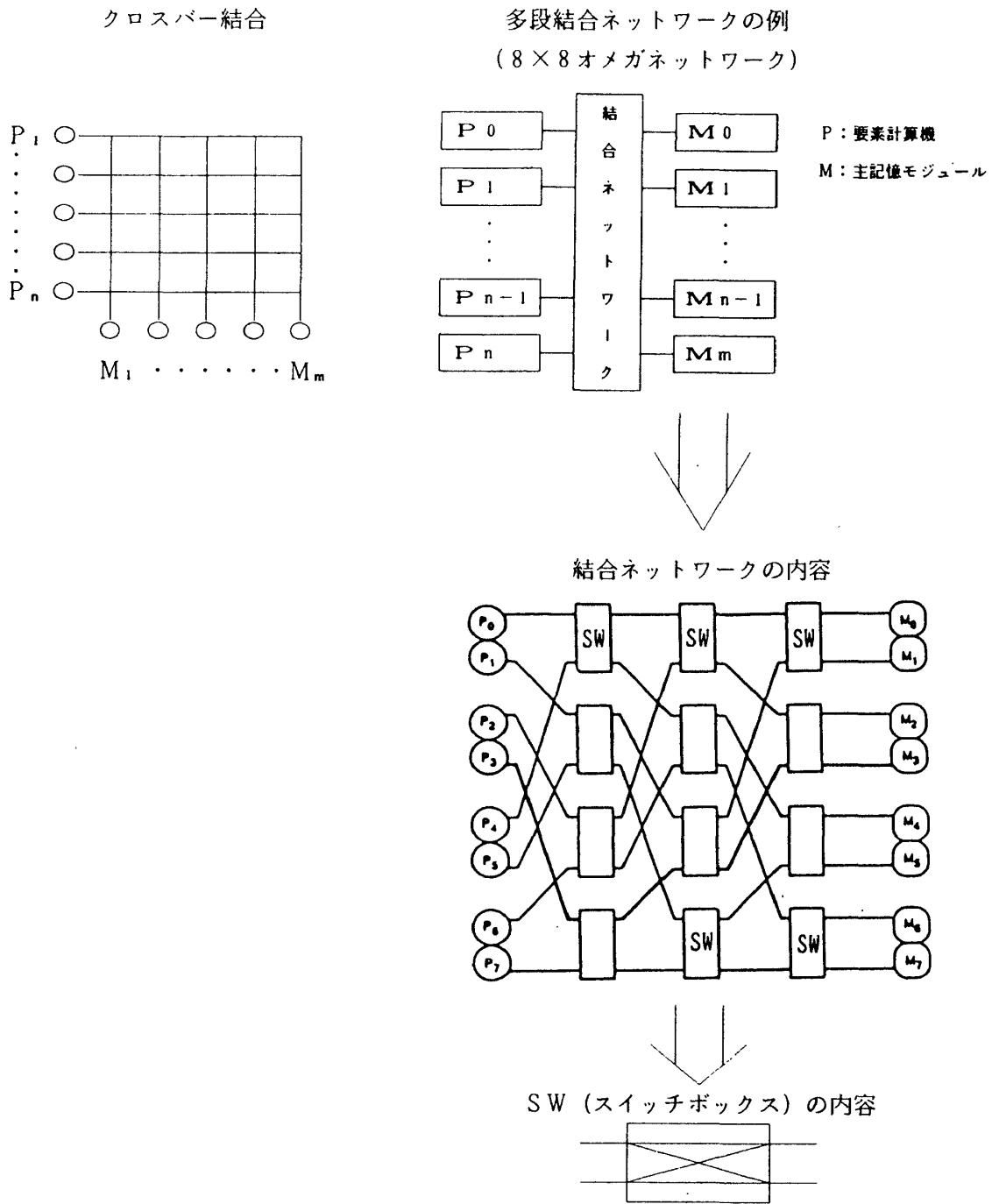


図8 並列計算機の動的結合ネットワークの例

とスイッチボックスにおいて競合する可能性は常に存在する。この競合は要素計算機-メモリモジュール対の対応が1対1であっても発生する。これを極力回避する為には結合ネットワークに冗長度を持たせる必要がある。

CFDプログラムの高速処理を目的とする並列計算機の構築において結合ネットワークのトポロジーの選択は重要である。選択の基準としては

(イ) 並列計算機の要素対 (BおよびB'方式にお

いては要素計算機と要素計算機, AおよびA'方式においては要素計算機とメモリモジュール) におけるデータ転送の距離が短いこと (アクセスタイムが短い) が望ましい。

(ロ) CFDプログラム処理においてデータ転送が発生する全ての要素対において, データ転送に関する距離が等しいことが望ましい。

(ハ) 要素対間のデータ転送において競合が発生しないことが望ましい。競合が発生した場合, そ

のうちの1つだけが実行され、他は待たされることになる。

(ニ) データ転送レートが大きいことが望ましい。

(ホ) 物量が少ないことが望ましい。これは結合ネットワークの実装可能性(構築可能性)とコストに関係する。

が挙げられる。(イ)～(ニ)は並列計算機の効率、従って実効処理速度に関係する。(ホ)は並列計算機の実現可能性に関係する。残念なことに(イ)～(ニ)を満足する結合ネットワーク、この節で述べた例では完全結合およびクロスバー結合は要素計算機の台数が増加すると(ホ)を満足しない。台数が増加しても(ホ)を満足できる結合ネットワークは(イ)～(ニ)の全て、或いはいずれかを十分に満足させることは出来ない。

以上述べた事から並列計算機の結合ネットワークのトポロジーの選択は要素計算機の台数等、並列計算機の構築における他の重要事項とのバランスの上で決定されなければならないことが明らかとなる。

要素計算機の台数と結合ネットワークのトポロジーの選択が決定すれば、結合ネットワークの同期方式、制御方式および情報交換方式、データ転送経路におけるバッファリング方式、耐故障性の方式等はそれに従って最適に決定することが可能となろう。

並列計算機の結合ネットワークは計算機工学の研究に格好の話題を提供しており、この分野における論文は山程あるが、1例として米国IEEEが編集した論文集¹⁵⁾を挙げておく。

4.2.4 計算機方式変更の影響

計算機の目標処理速度の実現のため、計算機の記憶構造から見た方式の変更が必須となったが、この変更はプログラム記述言語とコンパイラ技術に大きな影響を及ぼす。

AまたはA'方式の並列計算機であっても、BまたはB'方式の並列計算機であっても、プログラムの処理の進み方は、これ迄、航技研の主力計算機であった単一計算機方式の計算機の場合とは異なる。

この様な並列計算機を効率良く使用し、性能を

発揮させる為には、少なくとも以下の措置が必要となる。

[AおよびA'方式の並列計算機の場合]

A1. 計算処理される手続きの分割と分割された手続きの各要素計算機による並列処理を可能にする。

A2. 並列処理される手続きの間の同期をとる。

A3. 局所記憶と共用主記憶の間のデータ転送のスケジューリングと実行。

A4. A3のデータ転送と計算処理をオーバーラップさせる。

[BおよびB'方式の並列計算機の場合]

B1. A1と同じ。

B2. A2と同じ。

B3. アレイデータの分割と各要素計算機に分散配置されている主記憶へのデータ割付け。

B4. 分散配置されている主記憶間のデータ転送のスケジューリングと実行。

B5. B4のデータ転送と計算処理をオーバーラップさせる。

等である。A1およびB1においては各要素計算機の計算処理負荷が出来るだけ均等になる様に分割されなければならない。B3においてはB1で行われた手続きの分割と出来るだけ整合性が取れる様にデータ配置を行う必要がある。A3およびB4においてはデータ転送の頻度とデータ転送量が出来ただけ小さくなる様にデータ転送を行う必要がある。A4およびB5においてはオーバーラップ部分が最大となる様にデータ転送をスケジュールする必要がある。

A1～A4およびB1～B5の措置をコンパイラが全て引き受ける場合には、プログラム記述言語は現在のままで何の拡張も必要ではない。

しかしながら、ハードウェアの性能を最大限に発揮させるという条件のもとでコンパイラがこれら全てを行う為にはベクトル計算機の自動ベクトル化技術とはレベルの異なるはるかに高いレベルのインテリジェンスがコンパイラに必要となる為、ここ数年のうちにこれら全てをコンパイラが引き受けることは困難である。この理由で、これらの措置のうちの或る部分はプログラマが引き受けざ

るをえない。これを前提とすれば、プログラムは

(イ) Fortranを拡張する。

(ロ) A1～A4またはB1～B5に関するコンパイラへの指示文を作る。

(ハ) A1～A4またはB1～B5の措置を可能とする記述を含んだ高級言語を作り、プリプロセッサによりFortranプログラムを作成する。

等のうちいずれか1つの手法を採用することにより、並列計算機の効率的な使用を可能とする様に記述されなければならない。

4.2.5 計算機方式の変更による目標処理速度実現

主記憶データ転送能力不足という呪縛から解放されることにより、目標処理速度の達成の展望が開けてくる。

目標性能の達成に必要なピーク性能が256GFLOPS必要であるかどうか、また、128GFLOPSでは不足かどうかということは計算機の方式と構成、主記憶、演算器および結合ネットワーク等の計算機の要素の性能、或いはコンパイラ等基本ソフトウェアの性能等に依存して定まる。以下では、便宜的にピーク処理速度128GFLOPSの計算機により目標処理速度の達成が可能であると仮定して議論を進める。ピーク処理速度が200GFLOPS必要である場合でも類比は成立する。

現在のベクトル計算機は1台で5GFLOPS程度のピーク処理速度を達成している。ここ数年の内に1台のピーク処理速度が8GFLOPS程度のベクトル計算機は、実現可能であろう。

従って、5GFLOPSのベクトル計算機であれば25台、8GFLOPSのベクトル計算機であれば16台を、図4または5に示した方式のいずれかにより結合してピーク処理速度128GFLOPS程度の計算機を構成することは可能である。この様に現在のベクトル計算機の延長線上に構築する目標計算機を仮にE型計算機と呼ぶことにする。

もう一方の極は、現在急激に発展しつつあるマイクロプロセッサ技術をベースに目標計算機を構築しようという方向である。これをR型計算機と呼ぶことにする。

現在、米国Intel社のi860マイクロプロセッサ

は64ビット浮動小数点演算に関して約60MFLOPSの処理速度を達成している。現在でも、i860クラスの1チップ計算機2000台を、図7或いは8に示した方式のもとに結合すればピーク処理速度128GFLOPS程度の計算機は構築可能である。

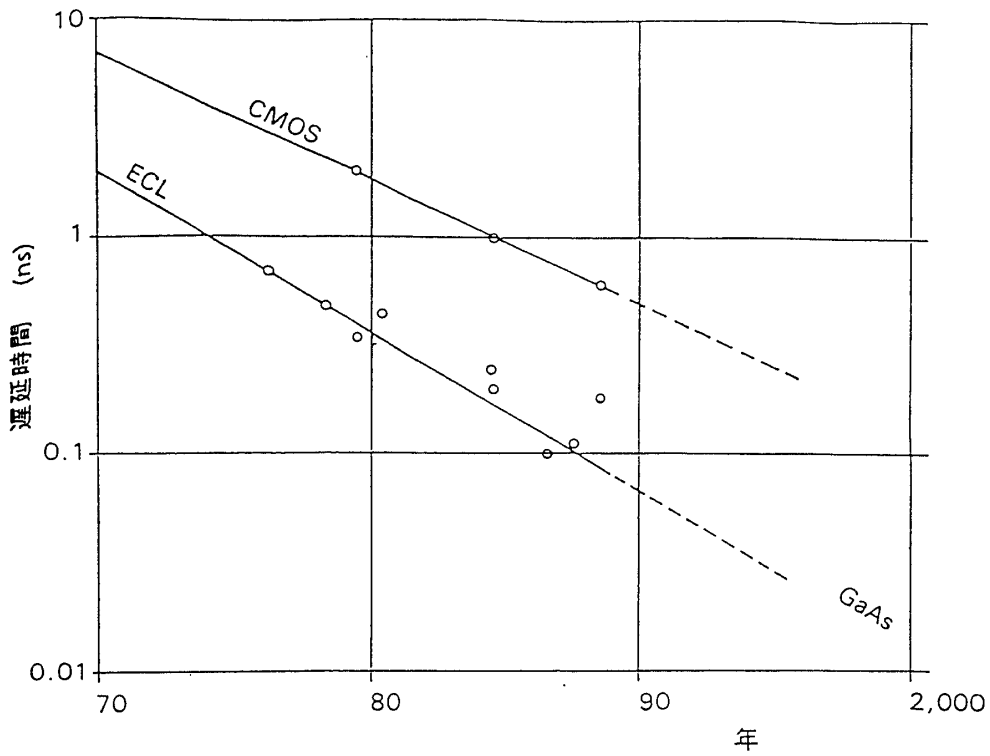
E型計算機とR型計算機の間接型ともいふべき計算機も構築可能である。図9¹⁰⁾および図10¹¹⁾はそれぞれ論理LSIの密度と速度の発展の状況を示したものである。図によれば現在、密度、数十万ゲート、スイッチングディレイ時間数百ピコ秒のCMOS論理素子が入手可能であり、数年内には百万ゲートのCMOS論理素子も入手可能となろう。また、図には表示されていないがバイポーラ素子とCMOS素子の中間型ともいふべきBICMOS素子も急速に発展しており、10万ゲートクラスの素子が入手可能になりつつある。

この様な論理素子を数チップないし数十チップ実装することにより、1GFLOPS級のコンパクトな計算機を作り、この計算機を図4或いは図5の方式により100～200台結合することにより目標計算機を構築することが可能となる。この型の結合をR'型と呼ぶことにする。

表6に現在発表されている並列計算機の一部を示す。これらは全てE型またはR型であり、その方式はBまたはB'に属するものと思われる。R'型に属する物はない様である(方式等について詳細な発表のないものについては推測の域をでないが)。

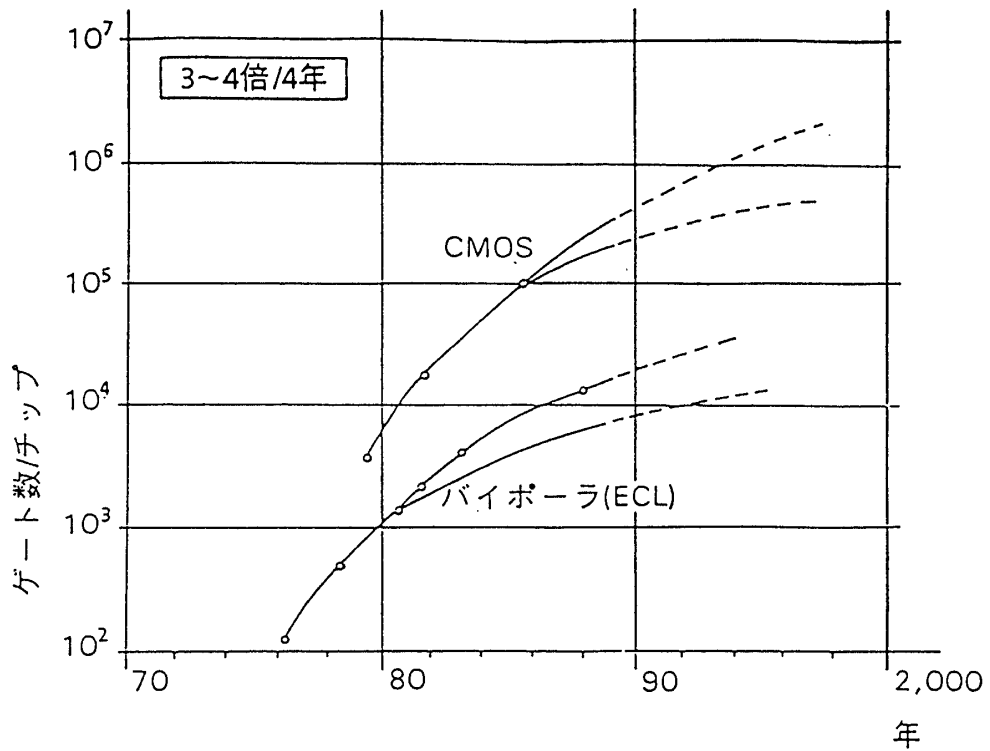
E型計算機の問題点は偏えにかかってコスト、ランニングコスト(消費電力)、設置面積、床面荷重等ハードウェア面の実現可能性にあると考えて良い。消費電力に関しては、コスト面の問題が解決(大量生産と大量消費以外に解決策はない)すればGaAs等の新素子の採用を考える必要があると考えられる。

R型計算機の問題点はE型計算機の場合と全く異なり、台数が多数になることに起因する結合ネットワークの実効データ転送能力の問題、要素計算機の効率的使用の問題等ソフト面にあると考えられる。これらはプログラム記述言語、コンパイラ、ユーザのプログラム技術および計算法にも関



日立製作所提供

図9 論理ICの速度の進歩



日立製作所提供

図10 論理ICの進歩

係するので、この問題の解決には計算機の作製者とユーザの緊密な連携が必要である。

R'計算機はE型計算機の問題点とR型計算機の問題点の両方がかかえているといえるが、問題点の厳しさは緩和されていると考えて良い。

文献¹³⁾では2000年には1チップの上に5千万個のトランジスタを積み込み、マシクロック時間4ns、スカラ処理速度750 MIPS、ベクトル処理速度1GFLOPSのマイクロプロセッサが実現できると予想している。こうした傾向はほぼ確実と考えられるので、その場合R'計算機はR型計算機に収束する可能性が大である。

以上の検討から、目標処理速度の実現にはかなり厳しい問題点はあるものの展望が開けていると考えて良い。

以上の議論では専らベクトル処理速度のみを考えてきた。単一計算機のスカラ処理速度の向上についても、最近、並列処理機能の導入により劇的な変化のきざしが見えて来ており、1マシクロック時間で1命令の処理*という大きな壁を破り、1マシクロック時間内に数命令処理する方式が実現され始めてきている。スカラ処理速度の向上も素子技術の向上による処理速度向上を数倍上廻る可能性が大となってきた**。

次節では目標計算機の実現に関する問題点についてもう一步議論を進めることにする。

5. 目標計算機の構築における重要検討項目

5.1 コスト

V.L.PetersonはNASA Ames R.C.のNAS計画に関する学会において¹²⁾、数値シミュレーションを航空機等の設計に使用するためには1ケースを10分間以内に100ドルの費用で計算する必要があると述べている。目標計算機の処理速度はクリーンな全機のReAvNS方程式に基づく数値シミュレーションを10分以下(5~6分)で実行する事を可能としているが、コスト面ではどの程度にな

るであろうか。これは単に経済的な問題ではなく、目標計算機の使用素子等を選択する場合の選択原理にかかわる重大問題である。使用素子の選択は計算機がE型かR型か或いはR'型を決めることであり、これにより要素計算機の台数、結合ネットワークの選択の方向が定まる結果となる。

10分間で100ドルというのはNASAの主要な風洞による風試の費用との比較から計算した費用と思われる¹²⁾。計算機を年間を通じて連続運転に近い状況で運用すればその稼動時間は8400時間程度に達する。計算機の償却を5年で行う(計算機の性能は5年もたつと陳腐化する)とすると10分間100ドルで計算を行う計算機のコストが計算できる。この費用は奇妙に航技研が近年設置または設置を計画している多くの大型試験研究設備のコストと符号するとともに、商用の超大型といわれるスーパーコンピュータのコストの上限値ともほぼ一致している。

現在のR型計算機のコストは1MFLOPS当たり500ドル程度と言われているが、このコストはここ数年以内に1/4程度に低下するという推定がある。この推定が正しいとすれば10分間でクリーンな全機の数値シミュレーションを100ドルのコストで実現することはかなり現実味を帯びてくる。E型計算機に関しては、S.CRAY氏はSupercomputing 88 conferenceでCRAY-3はCRAY-2の12倍のコストパフォーマンスで実現すると発言している。E型計算機のコストパフォーマンスも今後劇的に向上する可能性はあるが、コスト面ではR'型およびR型計算機の方が有利である可能性が高い。

5.2 消費電力量

計算機のランニングコストは人件費を除けば殆ど計算機および計算機の冷却に使用する電力料金で定まる。我が国の様に電気料金の高い所では特にそうであることは言うまでもない。

消費電力量の大小は殆ど使用素子により決まってしまう*。

* 現在のスカラ計算機の方式では実際には1命令の処理に数マシクロック時間がかかっている。

** この状況については文献¹⁴⁾を参照されたい。

* 消費電力量は冷却技術にも依存するが、素子の高速化の為の液体窒素、或いは液体ヘリウム冷却を考えなければ冷却技術も殆ど使用素子で定まってしまう。

現在の商用超大型スーパーコンピュータの消費電力は最も少ない消費電力のものでも80KW程度である。従って、現在の技術の延長線上にE型計算機を構築することは消費電力量の面から考えられない。E型計算機の構築においては素子技術の点で革新的技術開発が行われることが必要と考えられる。

消費電力の面ではR型計算機の方がはるかに有利である。R'型計算機はE型とR型の中間にある。

5.3 並列計算機の効率

4.2.2で示した様にCFDプログラムの処理速度がVP400の100倍の計算機はベクトルレジスタ付のベクトル計算機を要素計算機とする並列計算機により実現するのが適切である。すると、CFDプログラムの目標処理速度と並列計算機の台数、要素計算機1台のピーク処理速度との間に下式が成立しなければならない。

$$\text{CFDプログラムの目標処理速度} \\ = (P \times R) \times (N \times E_N)$$

P ; 要素計算機のピーク速度

R ; 要素計算機のCFDプログラム処理速度
/ 要素計算機のピーク速度

N ; 要素計算機の台数

E_N ; 要素計算機の実効平均的効率

R は、これ迄のベクトル計算機のCFDプログラム処理の経験から0.3~0.6程度と考えるのが妥当である。

R の値の大小は要素計算機の優劣を示すものではなく、計算機のコスト等を含めた設計思想を表現する数値であると考えるのが妥当である。一般的に、同一素子技術のもとで4.2.1で説明した多重化方式によりピーク処理速度を向上させれば R は相対的に小さくなる傾向がある。 E_N の値は要素計算機の台数、結合ネットワークのデータ転送能力に関する特性、要素計算機の実効処理速度と要素計算機の記憶容量と性能(AおよびA'方式では局所記憶容量、BおよびB'方式では分散配置されている記憶容量)および共用主記憶(AおよびA'方式)の性能とのバランス等に依存して定まる。 E_N の値は決して1にならず、1以下であるが、

その主な理由は

(1) プログラムは並列処理できない逐次処理部分が必要含まれており、逐次処理の場合、1台の要素計算機を除いて他の要素計算機は全てアイドルとなる。

(2) 並列処理に伴うOSのオーバヘッドが発生する。

(3) 並列処理過程では各要素計算機で処理されている処理間で同期をとる必要があるが、この為の同期オーバヘッドが発生する。また、プログラムの同期待ち合わせの為、要素計算機に待ちによるアイドル時間が発生する可能性がある。

(4) データ転送の為のオーバヘッドと待ちによるアイドル時間が発生する可能性がある。

(1)および(2)は主として計算法およびソフトウェア技術(プログラム記述言語を用いてのプログラム技術とコンパイラ技術およびプログラムの作成にかかわるOS技術。以下、ソフトウェア技術と言えどもこの意味である)に関する。これについて論ずることは本稿の考察の範囲を越えているが、並列計算機を使用する上で重要な事項であるので簡単に考察しておく。

5.3.1 プログラム逐次処理部分

次式は並列計算機の簡略化モデルに基づいて性能向上率を算出するアムダールの式である。

$$S_N = \frac{1}{1 - \alpha + \alpha/N} \quad (1)$$

α はプログラムの並列化率(単一ベクトル計算機の場合のプログラムのベクトル化率に対応するもの)、 N は要素計算機の台数である。

$$TE_N = \frac{S_N}{N} \quad (2)$$

TE_N は N 台の要素計算機を使用してプログラムの処理を行った時の要素計算機の理論平均的効率であって、一般的に $TE_N > E_N$ となる。その原因は主として以下5.3.2~5.3.4に示す事項にある。

(1)によれば、 α が0.95であれば S_N は要素計算機の台数を如何に増加させても20倍以上の性能向上は望めないことになり、並列計算機による処理速度向上には悲観的にならざるを得ない様であるが、この結論は誤っている。問題規模が大きくな

れば α は1に近づく。これは簡単な問題については証明されている¹⁶⁾。また、プログラムとしてFLO52を使用し、翼型廻りの流れを4種類の格子点(128×32, 256×32, 256×64, 256×128)の上で並列計算機 iPSC/2 (1台~16台) および NCUBE (1台~512台) を用いて行った数値実験でもこの事実は示されている¹⁷⁾。即ち、問題規模が大となれば並列化率 α は大となる。この事は逆に、相対的に並列化率の小さな問題を大規模な並列計算機を使用して処理するのは得策ではないことを示すものである。即ち、並列化率の小さい問題は並列計算機の一部を使用して処理すべきである。

本稿3において示したR1~R3のCFDプログラムについてはどれも並列化率は十分に高いと考えて良い。

プログラムによっては並列計算機の外部との入出力部分が逐次処理の内の大きな部分を占めることがある。この場合、外部との入出力を各要素計算毎に並列実行すれば α は大となる。これを可能にする為には並列計算機の外部記憶アクセス法にこの機能を附加する必要がある。

一般に α を大とするのはプログラムの使用する計算法およびプログラム技術の責任である。

この為、

- (1) 並列計算機に適した計算法の採用
- (2) コンパイラのプログラムの並列化可能部分の発見能力の向上
- (3) プログラムによる並列化部分の指示機能の強化

等が必要になる。(2)と(3)は互いに相補的な関係にある。

5.3.2 並列化の為のオーバヘッド

並列計算機ではプログラムは逐次処理→並列処理→逐次処理→並列処理……と処理される。処理モードの切り換え時にOSのオーバヘッドが発生する。このオーバヘッドを縮小する為には

- (1) 1回のオーバヘッドを極力小さくする。
- (2) 処理モードの切り換え頻度を少なくすることが必要である。(1)はOS性能の問題である。
- (2)については、処理モード切り換えの頻度を減らす為には、計算結果に影響がなければ逐次処理部

分も各要素計算機で独立に並列実行できる様にプログラム記述言語とコンパイラ技術で解決策を取ることが必要である。

5.3.3 同期

先ず、同期処理の為に以下の手段を用意する。

(1) 各要素計算機上で処理される処理間の共有変数を共用主記憶上に配置し、これの書き込みと読み出しにより行う等の手段(AおよびA'方式。B'方式も局所記憶上で同じ手段)を用いる。

(1') 要素計算機間のメッセージの送受信を行う(BおよびB'方式)および

(2) 要素計算機間の通信線とレジスタ或いは私有記憶の一部を用いた同期の為にハードウェア機構を用いる。

同期は(1)または(1')と(2)の組み合わせにより実現する。但し、要素計算機の台数が多くなると(2)を用意することは物量の点で困難になるので(1)または(1')のみの手段で同期を実現する場合もある。

同期処理は単一計算機では必要のない処理であるので並列計算機の効率 E_N を低下させる原因となる。この低下を軽減する為には、

- (イ) 同期の頻度を減少させる。これは主としてソフトウェア技術で解決すべき事項である。
- (ロ) 同期待ち合わせにより要素計算機がアイドルになる事を防止する。これは、各要素計算機に割当てられた処理量の負荷バランスの問題であり、(イ)と同じくソフトウェア技術で解決すべき事項である。

(イ) 同期処理手続きを効率的にする。これは、主として並列計算機OSで解決すべき事項である。

(ロ) (1)または(1')を高速に実行できるハードウェアを構築する。これは主として結合ネットワークを経て行われる要素計算機と共用主記憶(AおよびA'方式)、或いは要素計算機間(BおよびB'方式)のアクセスを高速化し、且つ競合発生を可能な限り防止する事により達成される。

結合ネットワークの高速化および競合の防止の為に並列計算機の構築において物量を投入することと並列計算機を構成する要素計算機の効率はト

レードオフの関係にあり、並列計算機構築における重要な選択肢の一つである。

5.3.4 データ転送とCFD計算法

本題に入る前に以下の考察に必要となる2つの概念の説明をしておく。

〔(通信量/計算量)比〕

今、AおよびA'方式の計算機の場合、要素計算機の局所記憶LMを要素計算機の私有記憶ということにする。また、BおよびB'方式の計算機の場合、要素計算機に分散配置されている主記憶を要素計算機の私有記憶ということにする。

並列計算機では計算の途中で何度も私有記憶と共用主記憶(AおよびA'方式)、私有記憶と私有記憶(BおよびB'方式)との間でデータの転送を行う必要があるが、データ転送とデータ転送の間に処理される計算量と転送されるデータ量の比を通信量/計算量比(以下CCと略記する)という。CCは出来るだけ小さい方がよい。CCを小さくするのはソフトウェア技術の責任である。

〔要素計算機の豊穡度*〕

要素計算機のピーク処理速度Pと私有記憶の容量PMの比PFを要素計算機の豊穡度ということにする。コスト、主記憶、実装技術面の制約がなければ、豊穡度は出来るだけ大きくとる方が並列計算機の効率上望ましい。

$$PF = \frac{PM}{P}$$

方式の如何を問わず、要素計算機がCFDプログラムの割当てられた一連の処理を実行する際、一連の処理に必要なデータが全て要素計算機の私有記憶に常に存在しているとは限らない。この場合、私有記憶と共用主記憶の間、或いは要素計算機の私有主記憶の間に必要データの転送が発生する。

要素計算機の処理速度に対してデータ転送が相対的に低速であり、処理に対してデータ転送が遅延し、要素計算機の必要とするデータが不在であれば、要素計算機はアイドル状態になり、要素計算機の効率 E_N は低下する。この場合は、計算処理に必要な主記憶量が不足し、これをバックアッ

プする為に磁気ディスク等の低速記憶装置を主記憶として使用する場合と相似である。この様な場合、通常の計算機においても効率は大幅に低下する。従って、データ転送遅延による効率の低下を防止することは並列計算機の構築における最重要課題の1つである。

データ転送のパターンは並列計算機の方式がAおよびA'の場合とBおよびB'の場合とでは異なる。また、CFDが使用する計算スキームとその解法によっても異なる。

〔並列計算機の方式とデータ転送〕

(1) AおよびA'方式;この方式では各要素計算機は処理に必要なデータを共用主記憶から私有記憶に読み出し、他の要素計算機が必要とするデータでそれが共用主記憶に存在せず、私有記憶にある場合、それを私有記憶から共用主記憶に書き出すというデータ転送が生ずる。

私有記憶容量が大きければ処理に必要なデータを長期的に渡って保持することが可能となる為、共用主記憶から読み出す必要がなくなり、データ転送の頻度は大幅に減少する。また、長期間におけるデータ転送のスケジューリングが可能になる為、種々の意味で最適化の自由度が大となる。

(2) BおよびB'方式;この方式では全ての主記憶は要素計算機に分散配置されているので、各要素計算機の処理に必要なデータが私有記憶になれば、それは必ず他の要素計算機の私有記憶に存在するので、そこから読み出す。或いは、他の要素計算機が処理に必要なデータがその要素計算機の私有記憶になく、自分の私有記憶に存在すれば、そのデータを他の要素計算機の私有記憶に書き出すというデータ転送が生ずる。この方式では、処理の各要素計算機への分割と割付がデータの各要素計算機への分割と割付にうまく対応していれば、データ転送の頻度は大幅に減少する。

〔CFD計算法とデータ転送の関係〕

目標とする並列計算機は全てのCFD計算法を同程度の超高速性で処理する必要はないが*、陽解法

* この概念は公認のものではない。ここで便宜的に作った概念である。

* これは現在のベクトル計算機でも同じである。単一ベクトル計算機では非構造体格子を用いた計算は構造体格子を用いた計算よりも低速で処理されているのが実情である。

といった特定の計算法のみを超高速処理の対象とするのは、CFDで用いられている計算法の現状から判断すると誤っていると言える。目標とする並列計算機はベクトル計算機を要素計算機とするので、少なくとも使用する要素計算機の台数×100程度の並列性を持ったCFD計算法についてはVP400の100倍以上の超高速性で処理をする必要がある。

これは3次元の計算空間で考えると2方向の並列度（面パラレル処理可能）を持ち、1方向の処理は逐次処理である様な計算法も超高速で処理できることを意味している。CFD計算法をデータ転送の観点から考察する場合、以下の(1)～(3)が重要である。

(1) 差分スキームの形（5点差分，3点差分等）；多数の格子点を使用するスキームを採用すればデータ転送の転送量，頻度ともに一般に増加する。

(2) 構造体格子または非構造体格子の採用；構造体格子ではデータ転送は一般に連続或いは一定ストライド付ベクトルの転送となる。各要素計算機から発生するデータ転送のパターンも規則的であり，結合ネットワーク上での競合を回避するスケジューリングの可能性も高い。非構造格子の場合，データ転送は一般にリストベクトルの転送となり，各要素計算機から発生するデータ転送は不規則でスケジューリングによる結合ネットワーク上での競合回避の可能性は余りないと考えて良い。競合は結合ネットワークの性能低下の原因となり，効率の低下をもたらす。

(3) 解法；

(イ) 陽解法では一般的に各要素計算機に割り当てられた計算空間の部分領域 D （手続きの分割とその要素計算機への割付け）の処理に必要なデータは図11に示す様に D と D の近傍 D_b の和集合 D_c 上のデータに局所化されている。

AおよびA'方式の場合，少々困難な所は有るがソフトウェア技術を高度化して，私有記憶へのアレイ変数割付けの最適化，データ転送の最適スケジューリングを行えば，結合ネットワークの負荷を軽減できる。BおよびB'方式の場合，陽解法は結合ネットワークに対する負荷が最も小さく，

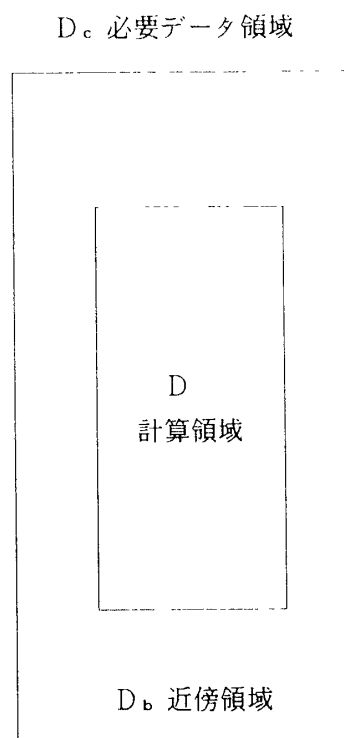


図11 陽解法の計算領域と必要データ領域

隣接結合ネットワーク*の様な簡単な結合ネットワークでも並列計算機は高い効率を実現できる。

(ロ) 陰解法；陰解法は多種多様である。図12は現在CFDで最も良く使用され解法に対する信頼性の高いIAF法における行列の逆転手順を示したものである。また，図13には収束が速いと言われているSymmetric ADIスイープ法における行列の逆転手順を示したものである。陰解法においても剰余項の計算において発生するデータ転送は陽解法と同じである。問題は行列の逆転手順にある。AおよびA'方式の場合，図12および13に示す様に行列逆転におけるスイープ方向が変わる所で要素計算機に割り当てられた計算領域が変化する。その都度，局所記憶の大部分の入れ替えの為のデータ転送が発生する。このデータ転送は結合ネットワークに対して大きな負荷となり，競合も多発する可能性が高く，データ待ち合わせに起因する並列計算機の効率低下をもたらす。BおよびB'方式においてもこの状況は全く同じである。一般にBおよびB'方式の私有記憶量はAおよびA'方式のそれ

* 上下左右の4方向結合だけではなく，これに斜め上方および下方を加えた8方向結合程度を考えるのも良い結合ネットワークである。

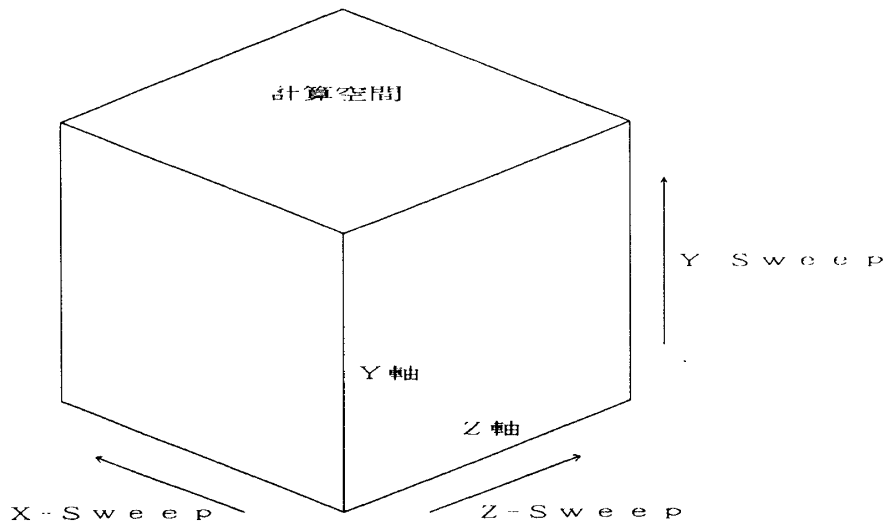


図12 IAF (ADI) 法

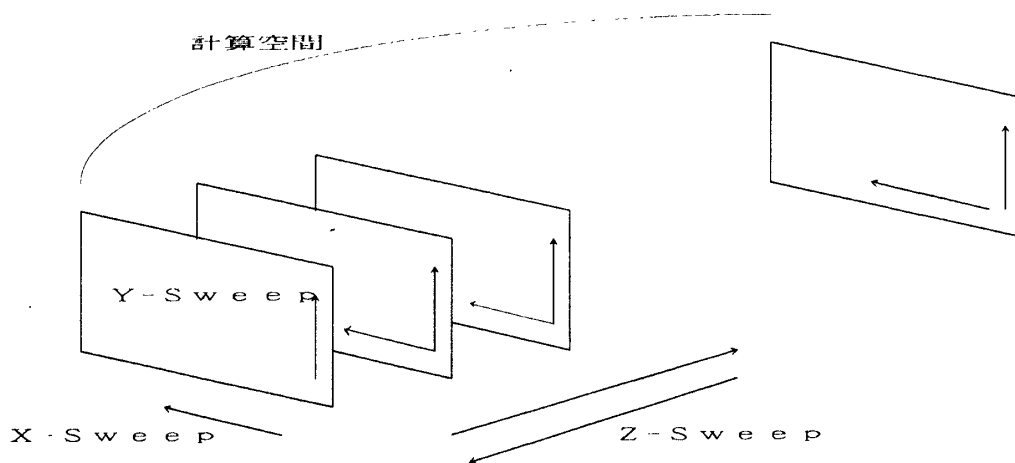


図13 Symmetric ADI Sweep

よりはるかに大きい。従って、スイープ方向の変更におけるデータの配置換えは全要素計算機を巻き込んだ大量のデータが結合ネットワーク上を移動する為、このデータ転送による並列計算機の効率低下はAおよびA'方式の場合よりはるかに大きな問題となる。これを防止する為には大量のデータ転送に対して強力な結合ネットワークを検討することが必要である。

並列計算機の効率に重大な影響を及ぼすデータ転送と並列計算機の方式およびCFD計算法の関係の概略は以上のとおりである。

データ転送の起因する並列計算機の効率低下の防止には、

(1) データ転送の頻度とデータ転送量を減少さ

せる、その結果データ転送とデータ転速の間の計算量が増えてCCが減少する。この為には、

(イ) 私有記憶におけるデータの割付け、およびデータ転送スケジューリングの最適化を行う(ソフトウェア技術)。

(ロ) 実装技術とコストの許す限りPFを大とする(ハードウェア)。これはCCを小さくする為、間接的であるが有効な措置である。

(ハ) BおよびB'方式においては要素計算機に割当てられる手続き(計算領域)と割付けデータ領域を可能な限り適合させる(ソフトウェア技術)。

(2) データ転送と計算処理をオーバーラップさせ、データ転送時間が露出しない様にする為、

(イ) データ転送と計算処理を並列実行できるハードウェアを用意する(ハードウェア)。

(ロ) データ転送と計算処理の最適スケジューリングを行う(ソフトウェア技術)。

(3) データ転送手続きの高速化を行う為にOSとコンパイラの高速化を図る(ソフトウェア技術)。

(4) データ転送のパイプライン化を図る(ハードウェアおよびプログラム技術)。

上記(1)~(4)の措置を取ることを前提とした上で、アクセスタイムが短く、転送レートが大であり且つ競合の少ない結合ネットワークを選択することは並列計算機の効率向上における最重要課題の1つである。しかしながら前にも述べたとおり、強力な結合ネットワークの構築は要素計算機の台数が増大すると困難となる。その場合は効率の低下を台数の増加により補う(並列計算機のピーク性能の向上)か、台数を減少させて(並列計算機のピーク性能の低下)効率を向上させるか、いずれかをコスト、ランニングコストおよび将来性等の観点から選択しなければならない。

5.3.5 目標計算機の構築

4.および5.に述べたことから、32GB以上の主記憶容量を持ちCFDプログラムの処理速度がVP400の100倍以上の計算機の構築には多種多様な選択肢が存在することが明らかとなった。これらの選択は実現可能性、将来性、コスト、ランニングコスト、および使い易さの観点からなされなければならない。

最も重要な選択分岐は4.2.5に述べた素子技術の選択(E型、R'型およびR型)である。素子技術の選択により要素計算機の実現可能性のあるピーク処理速度の範囲、豊穡度の範囲および目標処理速度を実現する為の要素計算機の台数の範囲等が定まってくる。台数の範囲が定まれば構築可能な結合ネットワークのトポロジー、アクセスタイム、データ転送能力等の範囲も定まってくる。

方式を主記憶共用方式(AまたはA'方式)とするか、或いは分散主記憶方式(BまたはB'方式)とするかの選択はソフトウェア技術の観点からは主記憶共用方式の方が有利であることは明らかである。従って、要素計算機のピーク処理速度と豊

穡度の実現可能性の範囲が定まることを前提として、共用主記憶の結合ネットワークを介在させた実効データ転送性能等が要素計算機の必要実効処理速度を支えるのに充分であれば共用主記憶方式を選択すべきことは明らかである。これが可能でない場合ソフトウェア技術上困難なことが予想できてもBまたはB'方式を採用する必要が生ずる。AおよびA'方式またはBおよびB'方式の選択はソフトウェア技術の容易さと目標処理速度の実現およびソフトウェア技術の実現可能性の観点からなされる必要がある。

使用素子技術と方式が定めれば、そこから並列計算機の効率を考慮しつつ、要素計算機の命令仕様、制御方式等、私有記憶の制御方式、管理方式等、結合ネットワークの制御方式、同期方式、情報交換選択方式等およびソフトウェア仕様とその実現方式等を定めて行くことになる。これらの選択についての考察は与えられた紙数との関係もあり本稿では省略する。

以上述べた過程は理念的なものであって現実には多くの手もどりのある繰返し過程であることは言うまでもない事であろう。

ここで強調しておくべきことは並列計算機の構築においては検討の一区切り毎に効率の観点からの実効処理速度の推定の為に検討のレベルに対応したシミュレーションを行うことが不可欠であるということである。

6. おわりに

本稿では今後のCFDの推進に必要な計算機性能を明確にし、この性能を実現する為には方式の変革が必要であることを指摘し、方式の変革による目標性能の実現可能性、目標計算機の構築における重要検討項目等について考察した。

与えられた紙数に対して考察すべきことが多岐に渡ることから、各項目に対する議論が上滑りせざるを得なかった。しかしながら事の本質についての考察には大きな誤りはないと信じている。

並列計算機が計算機性能の大幅な向上によって期待の星であると言われ始めてから可成りな年月が経過している。現在迄の所、計算機の実用現場

で、この期待に応えた並列計算機は皆無であると言っても言い過ぎにはならない。その理由は多岐に渡り複雑である。

航技術は我が国で初めてのベクトル計算機の開発と我が国におけるGFLOPS級計算機の実現にリーダーシップを取ることが出来たという誇るべき経歴を持っている。航技研がリーダーシップを取り、世界に先駆けて超高速の実用並列計算機の実現に成功し、その利用技術に大きな成果を挙げていることが出来ればその経歴には更に花が添えられる事になろう。航技研の研究者の尚一層の努力を祈念する次第である。

最後に筆者の読み難い原稿をワープロ原稿にして頂いた林明美さんに感謝の意を表明して稿を閉じることにする。

文 献

- 1) W. Gentsch, K. W. Neves, H. Yoshihara; AGARDograph No. 311, 1988.
- 2) D. R. Chapman; Computational Aerodynamics Development and Outlook AIAA paper 79-0129 1979. 1.
- 3) Request for proposal for System Design Study for NAS, 1980. 4 NASA Ames R.C;
- 4) IEEE-USA Scientific Supercomputer Subcommittee; Super Computer Hardware 1989.
- 5) Buisness Week; 1990. 4. 30, pp 38~44.
- 6) 日経エレクトロニクス; 1989. 7. 28, pp. 123 ~145.
- 7) 日経コンピュータ; 1989. 10. 23, p. 61.
- 8) 日立製作所; private communication.
- 9) D. H. Bailey; Vector Computer Memory Bank Contention, IEEE Transaction on computer vol. c-36 no.3, 1987. 3, pp. 293 ~ 298.
- 10) 日立製作所; private communication.
- 11) 日立製作所; private communication.
- 12) V.L. Peterson: NASA CP 2032, 1987. 2, pp. 5 ~30.
- 13) P.P. Gelsinger; Microprocessors circa 2000, IEEE Spectrum 1989. 10.
- 14) S. Weiss; Scalar Supercomputer Architecture, Proceedings IEEE vol. 77 no. 12, 1989.
- 15) C. Wu & T. Feng ed; Tutorial Interconnection networks for Parallel and Distributed Processing, IEEE computer Society, 1984.
- 16) A.L. De Cegama; Parallel Processing Architectures and VLSI Hardware vol. 1, pp. 5 ~8, 1989, Prentice Hall.
- 17) E. Barszcz, T.F. Chan, D.C. Jespersen, R.S. Tuminaro; Performance of an Euler Code on Hypercubes, Parallel CFD Implementations and Results using MIMD Computers 1989.