

特別論説



情報処理最前線

「超並列は役に立つのか」

高性能コンピュータ・システムの動向†

鈴木 則 久††

1. はじめに

現在、コンピュータ業界は大変革のまただ中にある。この大変革は大きく分けると、次のような三つの潮流よりなる。まず、「ダウンサイジング」という言葉に代表されるように、大型機や中型機から、ネットワークで接続されたワークステーションやパソコンへの移行である。二番目は、「閉じたシステムからオープンなシステムへの移行」である。この二つの動きのために、コンピュータを作るのは簡単になり、パソコンでは多くのメーカーのどの機種を買っても大同小異になり、単価も安くなった。その結果、第三の潮流として、利潤はマイクロプロセッサや入出力機器を作る部品会社、ソフトウェア会社、サービス会社へと移行していった。

この大変革の第一の潮流はスーパーチップと呼ばれるマイクロプロセッサの進歩によってもたらされた。スーパーチップの性能の伸びは目を見張らせるものがある。1985年からは毎年倍々で性能が伸びてきている。最近では一チップで、大型機やミニコンピュータよりも性能が良くなってきている。ワークステーションでも、一時代前のベクトル型スーパーコンピュータよりも性能が良いものも出てきている。これらのコンピュータとスーパーチップの性能を比較したのが図-1である。

この比較は、ディスクなどの入出力操作を含まない、プロセッサだけの性能の比較である。実際の大きな商用のアプリケーション・プログラムで、データベース処理やトランザクション処理を行うものでは、この値と違ってくる場合もある。

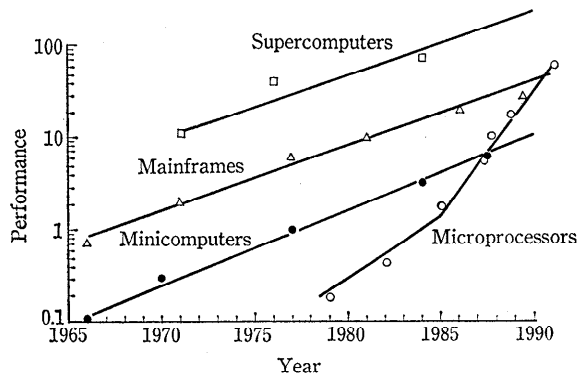


図-1 マイクロプロセッサとコンピュータの性能の伸びの比較 (文献 2) より引用) ©1991 IEEE

しかしながら、単にプロセッサの速度だけでも1990年代前半に、スーパーチップがスーパーコンピュータよりも速くなると予測されている。

一方、並列コンピュータも着々と製品化されてきている。スーパーチップを大量に使った超並列コンピュータも商品化され始めている。

ここでわれわれコンピュータの研究・開発に従事するものにとっての最大の関心事は将来のコンピュータ・システム、特に高性能のコンピュータ・システム、はどうなるであろうかということである。

並列システムは使いにくいと言ってこられた。科学計算などで、並列システムを十分使いこなしているアプリケーション・プログラムはまだ少ない。いわんや超並列システムについてはなおさらである。しかしながら、逐次命令のストリームをそのまま実行するユニプロセッサの速度にもいずれ限界がくるであろう。また、いつになってもユニプロセッサでは十分でない膨大な量の計算というものがあるので超並列システムは必要であるかもしれない。これらの疑問に対しての技術予測を試みるのがこの論説の主題である。

† Trends of High Performance Computer Systems by Norihisa SUZUKI (IBM Tokyo Research Laboratory).
 †† 日本アイ・ビー・エム東京基礎研究所

2. スーパーチップの将来

最初のマイクロプロセッサは、1971年インテルから発売された4004である。その直後に開発された8080は0.1MIPSであった。1992年にDECが開発したアルファというスーパーチップは400MIPSあるといわれているから、データ幅まで入れて考えると、この20年間で1万6000倍に性能が向上したことになる。この間に毎年同じ割合で性能が向上したとすると、10年で130倍になったことになる。

この性能の伸びは、半導体製造技術の進歩と、アーキテクチャとコンパイラ技術の進歩に支えられている。研究段階での半導体製造技術の現状を考えると、少なくとも今後10年間はこの伸びは続くと考えられる。

スーパーチップの速度向上の第一の要因はVLSIのチャンネル幅が指数関数的に減少していることである。この具合を示したのが図-2である。1960年には55ミクロンだったのが、1990年には0.8ミクロンになっている。すなわち30年間に70分の1になっている。

チャンネル幅とスーパーチップの性能の間には、次のような簡単な関係がある。まず、チャンネル幅が n 分の1になったとしよう。すると、スーパーチップ上のトランジスタのスイッチ速度は、 n に比例して速くなる。すなわち $O(n)$ 倍になる。一方、トランジスタの一面の大きさも n 分の1になるので、同じ大きさのチップ上には、 $O(n \times n)$ のトランジスタが乗ることになる。

しかしながら、この20年間には、チャンネル幅は20分の1になっているが、スーパーチップの

速度は4000倍、またデータ幅まで入れれば1万6000倍にもなっていることは前にも述べた。それでは、この能力向上はどこから出てきているのであろうか。

まず、それは単に素子の速度向上からだけで得られたのではないことは明白である。なぜならば、素子の向上だけに頼っていたのではスーパーチップの速度は20倍にしかなっていなかったであろう。

この20年の間にこれだけの性能向上が達成されたのは、スーパーチップの設計の進歩、すなわちアーキテクチャの進歩が大きく寄与している。このアーキテクチャの進歩とは $O(n \times n)$ で増えたトランジスタを使って高速のアルゴリズムを使い、並列性とパイプラインを増やすことを指している。アルゴリズムの分野ではスペースと時間とのトレードオフと言われているが、スペース（すなわち多量のトランジスタ）を使うことによって速度を速くしているのである。

その例としては、一つ一つの演算器を多量のトランジスタを使って速くすることが考えられている。掛算器では、加算器とシフト・レジスタを使って1ビットずつ計算していくやり方と、面積はたくさん使うが高速に計算できるブース・アレイとがある。

また、RISC計算機では、多段のパイプラインを使って、高速化しているが、パイプラインの段数を増やせば増やすだけたくさんのトランジスタが必要になる。また、スーパースカラのアーキテクチャでは、演算器を複数個用意することにより、同時に多命令を実行できるが、これには多くのトランジスタを使うようになっている。

また、もう一つ、スーパーチップ上のトランジスタを使って速くする方法の一つはキャッシュである。キャッシュは、プロセッサと主記憶の間にある小容量だが高速の記憶で、プロセッサのよく参照する主記憶番地を保持することにより、主記憶への直接の参照をできるだけ少なくする目的で作られている。

半導体製造技術の進歩により、チャンネル幅が指数関数的に小さくなり、スーパーチップの性能が指数関数的に進歩していることは前に述べた。しかしながら、同じ半導体で作られているメモリ・チップは容量は指数関数的に伸びている

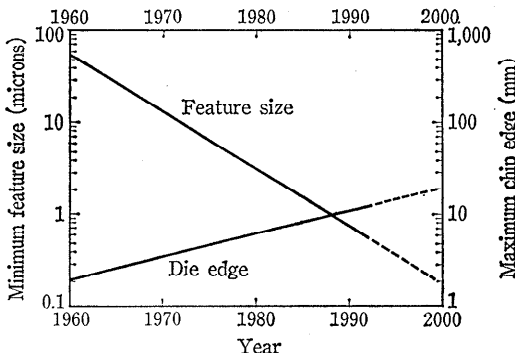


図-2 最小チャンネル幅と最大チップ・サイズの動向
(文献 2) より引用) ©1991 IEEE

ものの、アクセス速度はあまり伸びていない。このためプロセッサとメモリとの速度の差はどんどん広がっている。そのためメモリ・チップを直接プロセッサにつないだのでは、記憶参照命令の実行が遅くなる。そこで、プロセッサと主記憶の間に、容量は小さいが速度の速い記憶を置いて、プロセッサの実行を遅れないようにしている。このキャッシュは最初は一段であったが、プロセッサと主記憶との速度の差が広がるにつれて二段階置くようになっていく。

マイクロプロセッサのキャッシュにはもう一つ効果がある。プロセッサはどんどん速くなり、最近では 100 MHz, 200 MHz でも動かせるようになった。しかし、信号が一度チップから外に出ると、速度が遅くなり、そんなに速く信号を伝播させることができなくなる。そこで、チップ内では、超高速で動かし、チップから信号が出たら、もっと制御可能な 50 MHz 以下の低い周波数で動かすようにすることが考えられた。こうすれば実装も簡単になる。そこでキャッシュを使い、プロセッサ側からのアクセスは高周波数で、外部バスからのアクセスは低周波数とするようなスーパーチップができていく。これもキャッシュがあるからできるわけである。

さらにスーパーチップ上にたくさんトランジスタを乗せることができるようになると、並列処理をおこなうようになると考えられる。逐次実行型の直列の命令列は、そのままにして、二つ以上の命令を同時に実行して、高速化するアーキテクチャにスーパースカラというものがある。しかしながらスーパースカラでは、4つから8つの演算器をもうけて、4から8命令を同時実行するのが限界で、これ以上の並列度をアーキテクチャにもたせても、直列命令列からはそんなに多くの並列度を取り出すことができないだろうという実験結果が出ている。そうなると、複数の命令列を並列に実行することが考えられてくる。このようにして、スーパーチップの性能は、チャンネル幅の縮小に対して、ほぼ $O(n \times n \times n)$ に比例して向上してきた。図-3に、実際に市販されているマイクロプロセッサの各年ごとの最高クロック速度と、その上のトランジスタ数を示す。

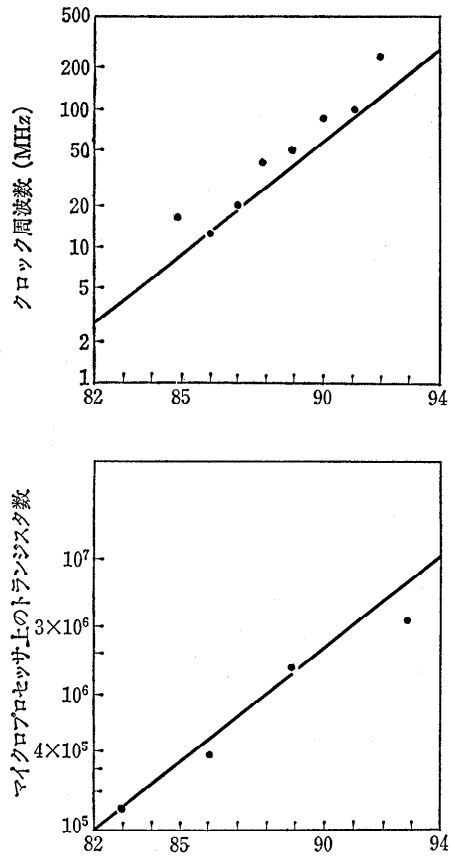


図-3 実際に市販されているマイクロプロセッサの最高クロック周波数と、トランジスタ数

3. 通信技術の進歩

1990年代に革新的に伸びるだろうと考えられているもう一つの技術は通信である。現在は、いろいろなところで使われている電子による信号の伝送が、今後 10 年間であらゆるところで光による伝送によって置き換えられていく。コンピュータの内部では、チップの間の通信、ボードの間の通信などが光になる。コンピュータの外では、LAN が光になる。また建物の外に目をやると、大都市内の通信、あるいは大陸横断の通信などがどんどん光にとって変わられる。

また、光通信自身の伝送容量も上がる。現在は 1 ギガ・ビットの通信が実用化されているが、10 年後には、10 ギガ・ビットの通信が実用化されるであろうと考えられている。通信媒体が電気から光に変わることと、光通信の伝送容量が増えることを合わせて、通信能力はこの 10 年で 1000 倍

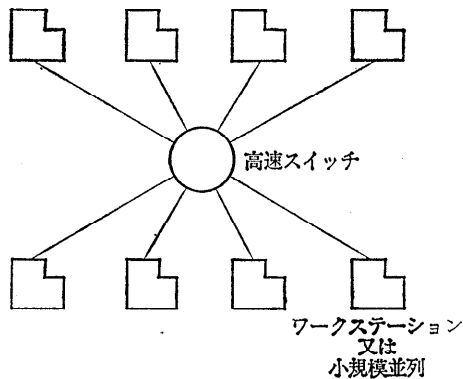


図-4 高速スイッチでコンピュータを接続して作られるクラスタ・システム

になると予測される。

これらの通信技術の中で、高性能計算機システムにとくに影響が大きいのは、距離が高々 100 m までの複数のコンピュータを接続する、スイッチ型のネットワークである。これは、図-4 のような形をとる。このようなスイッチ型のネットワークのprotocolsとしては、標準化が望まれるが、現在広くサポートされているものとしては、ATM スイッチと FCS スイッチとがある。

まず、高速ネットワークとして、なぜイーサネット型やトークンリング型の LAN 接続でなく、スイッチが選ばれるかという点、それは、アダプタのコストに比べて高性能が得られるからである。たとえば、100 MHz のネットワークに接続する ATM スイッチ用のアダプタの平均性能と、FDDI 用のアダプタの平均性能を比べてみよう。ATM スイッチでは、接続しているコンピュータの数がいくら増えても、100 MHz でデータを送り出すことができる。一方 FDDI の場合、リング上に 10 台ワークステーションがあり、同時にすべてのワークステーションがデータを送り出すとしたら、1 台ではその 10 分の 1、すなわち 10 MHz に対応するデータしか送り出すことができない。しかしながら、アダプタ・カードは、両方とも 100 MHz に対応するように作らなければならないので製作費は同程度である。

10 MHz のアダプタ・カードと 100 MHz のアダプタ・カードではあまり製作費が変わらないかもしれないが、100 MHz のアダプタ・カードと 1 GHz のアダプタ・カードでは製作費が大幅に違うので、スイッチのほうが経済的であるわけ

ある。

また、スイッチ型の利点はノード間の遅延が短いことである。ノード間の遅延が 2 マイクロ秒という超高性能の ATM スイッチの試作に成功したという報告がある³⁾。2 マイクロ秒というと、50 MHz のスーパーチップでいうと 100 命令である。これは、バス結合の共有記憶型並列コンピュータでのバスを通してのプロセッサとプロセッサの通信速度と同じ程度である。そこで、安くて速いワークステーションやパーソナル・コンピュータ、あるいは小規模並列コンピュータを高速スイッチで結合すると、性能的には、超並列コンピュータとあまり変わらないものができる可能性がある。一方リングでは、アダプタを通るたびに遅延があるので、ノード間の遅延はリング上のノードの数に比例する。

しかしながら現在のワークステーションそのままでは、いくら高速のネットワークで接続しても性能は上がらない。それは、ワークステーション側のネットワーク接続用アダプタ・カードやオペレーティング・システムに問題があるからである。データをアプリケーション・プログラムからネットワークまで送るまで、また送られてきたデータをアプリケーション・プログラムに渡すまでの時間が長く、特に高速通信では、それが主要部分になるからである。それを解決するために、アダプタ・カードやオペレーティング・システムでのオーバーヘッドを減らす研究がなされてきた。

その効果を示したのが図-5 である。この図では普通の UNIX の UDP プロトコルと、パケットのコピーをなくすような実現法によりオーバーヘッドを減らしたときの性能を比較している。現在のままでは、100 MHz 以上のスイッチを使っても性能は全然上がらないことが分かる。

4. 超並列の将来

さて、それでは超並列コンピュータの将来はどうであろうか。

超並列コンピュータを、スーパーチップを 100 以上（現在は数万が最高である）集めた並列コンピュータと定義しよう。アーキテクチャとしてはメッセージ送信型の MIMD が主流であるが、そのほか SIMD アーキテクチャ、分散記憶方式の共有記憶型などがある。超並列コンピュータが今

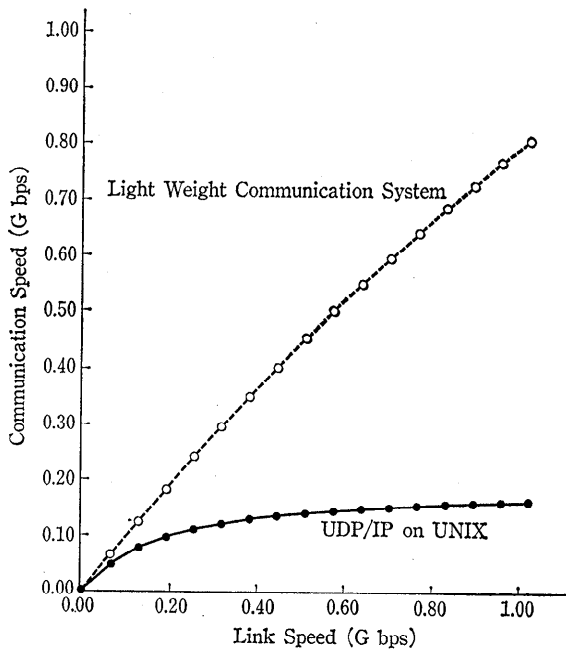


図-5 高速通信回線で接続したときの、通常の UNIX での実行通信速度と、改良した通信機構での通信速度 (文献 5) より引用。

後広く使われるかどうか考えるのには、いろいろなファクタを考える必要がある。重要なファクタとしては、一つは、代替技術の進展状況、次はエコロジ的考察、三番目には政府からの資金のファクタがあげられる。エコロジ的考察は 5. で行う。

まず代替技術としては、スーパーチップの伸びが今後も素晴らしいので、一チップあるいは小規模の共有記憶型並列処理のワークステーションが考えられる。もう一つは、これらのワークステーションを高性能スイッチで接続したクラスタが考えられる。これらのシステムで超並列コンピュータのやる計算を十分カバーできるので、超並列コンピュータを導入する必要がないのではなかろうかということである。

まず、一つには、スーパーチップに代表される直列命令流を実行するチップの速さは指数関数的以上の速さで速くなってきている。このチップのアーキテクチャとしては、スーパスカラとか VLIW というような形をとっていて、内部的には並列実行するコンピュータではあるが、実行する命令流はいまだ直列である。すでに、プロトタイプとしては 200 MHz, 400 MHz で動くチップも出てきている。1 ギガ・フロップス, 1 ギップス

のチップも、数年先には出てくるであろう。このような超高性能のスーパーチップを複数個スヌープキャッシュ¹⁾を使って接続した共有記憶型小規模並列計算機の商品計画は続々発表されている。これらの計算機は 10 ギガ・フロップス出ると発表されている。

一方、現在、普通に使われているベクトル型スーパーコンピュータは 1 ギガ・フロップス程度である。だから、数年後のスーパーチップは、現在の最も広く使われているスーパーコンピュータと同程度の計算ができるということになる。西暦 2000 年には、1 チップで数ギップス、共有記憶型小規模並列計算機で数 10 ギップスのワークステーションあるいはサーバが安く使えるようになるだろうが、果たしてこれ以上の演算を望むアプリケーションがどれほどあるだ

ろうか。

現在のコンピュータの性能を比較したデータとして、最も広くデータを集めたものにドンガラ・ベンチマークがある⁴⁾。この表を見て分かることは、超並列コンピュータでは、そのピーク性能を出すには、データ量が莫大に多いときに限られている。要するに、超並列コンピュータが性能を発揮するのは、大量のデータを分割して計算する場合に限られるのである。このベンチマークの中の典型的なものを表-1 に示す。

ここで取り上げた三つのコンピュータはそれぞれ特徴のあるコンピュータで、ドンガラ・ベンチマークで計られた最高性能のものである。すなわち NEC SX-3/44 はベクトル型スーパーコンピュータ、Intel Delta は MIMD 型超並列スーパーコンピュータ、Thinking Machines CM-200 は SIMD 型超並列スーパーコンピュータである。この表で分かることは、どの形のスーパーコンピュータでも、プロセッサ数が多いほど、ピーク性能に近づくには問題の大きさが大きくなってはならず、CM-200 では、28000 次元の連立方程式を動かしてもやっとピークの 45% の性能しか出ていない。これから言えることは、ピーク性能が同じならば、

表-1 超並列コンピュータの性能

コンピュータ	プロセッサ数	r_{\max} Gflops	n_{\max} order	$n_{1/2}$ order	r_{peak} Gflops
NEC SX-3/44 (2.9 ns)	4	20.0	6144	832	22
Intel Delta (40 MHz)	512	13.9	25000	7500	20
Thinking Machines CM-200 (10 MHz)	2048	9.0	28672	11264	20

この性能表を得るには、線形連立方程式を解いた。

r_{\max} はこの計算機で動かした最も大きな問題で到達した性能。

n_{\max} は、最も大きな問題のサイズ。

$n_{1/2}$ は、 r_{\max} の半分の性能を得るのに必要な問題のサイズ。

r_{peak} は理論的最高速度。

プロセッサ数の少ないコンピュータのほうが一般的に性能が出るということである。また、超並列コンピュータでは、扱うデータの数が大変大きいときにのみ、その性能を発揮できる。こういう用途はあまり多くない。

しかしながら、大量のデータを扱い、またいくらかでも速いコンピュータが必要であるという用途もある。たとえば、天気予報のために、天候のシミュレーションを行うという用途がある。シミュレーションではメッシュを細かくすればするほど精度が上がるが、それだけデータ量も増え、計算能力への要求は増える。

しかしながら、それらははたして必要な計算であるかどうかという疑問が残る。スーパーコライダについて考えてみよう。クオーク理論は素粒子論の基礎として広く認められるようになり、クオーク理論を最終的に証明するには、いまだ見つけられていないトップ・クオークを探せばよい。そのためにスーパーコライダを作ろうという提案がある。しかしながら、トップ・クオークが見つからないからといって人類の日常生活には一切関係ない。トップ・クオークの発見は科学的に人類の知識に偉大な貢献をすることに疑いはないが、資金のないときには、スーパーコライダの建設というものは延期をしても構わないような種類のようなものである。

同様に超並列コンピュータは圧倒的に政府機関で、主に物理計算やシミュレーションのために購入されている。これらは、多くが科学においての新しい知識の獲得のために使われているが、その目的はスーパーコライダと同じく、延期しても構わない種類のものである。

これらの考察から分かるように、超並列コンピュータが必要な分野は非常に狭いことが分かったが、これらの分野についても、代替機がある。

それは前にも述べた、ワークステーションあるいは小規模並列コンピュータのクラスタである。クラスタは今すでに使われていて、アメリカのリバモア研究所では、ベクトル型スーパーコンピュータの代わりに使われている。しかしながら、現在のクラスタはイーサネットなど既存のネットワークを使っているため、そこがボトルネックになっていて、システム全体の性能もあまり良くない。

しかしながら、将来のクラスタは、現在のクラスタよりもはるかに性能が良くなり、超並列コンピュータとの差は縮まっていくと考えられる。

その根拠としては、今後 10 年間では、通信路の性能は、光ケーブルが普及し、その光通信でもコーヒーレント光やソリトン通信などが実用化し、今よりも 1000 倍ほど伸びると考えられる。一方プロセッサの性能は数 10 倍から 100 倍程度の伸びしか見込まれていない。そうすると、クラスタでも、ワークステーションの速度が 10 倍になったとき、通信網の性能は 10 倍以上になっている。ところで、クラスタでの計算は、プロセッサ上での演算と、通信に費やされる時間とに分けられる。バランスの良いアーキテクチャとは、この両方が同じくらいの時間がかかるシステムである。いまプロセッサでの時間と通信時間が同じシステムがあったとしよう。ここでプロセッサの速度が 10 倍になれば、システムの全体の性能は、ほとんど通信路の速度で決まり、せいぜい 2 倍にしかならない。しかし、同時に通信にかかる速度が 10 分の 1 になればシステム全体の性能も 10 分の 1 になる。

これらの議論から、今後は、クラスタでは、ボトルネックは通信からプロセッサのほうに移っていき、基本的に超並列コンピュータと同じ問題になってくるわけである。

今後、本当に大きな問題において、なんらかの

形の専用超並列コンピュータがいるか、クラスタのほうがコスト・パフォーマンスが良いかは明確な回答は得られていないし、大きな研究問題である。

5. 社会的制約

現在は情報技術に対して社会は大変好意的である。情報産業の作り出す雇用や便利さに対する期待が大きい。しかしながら、今後は情報技術の作り出すネガティブな側面にも一般世論の目は向いていくであろう。すでに、情報産業が作り出す中古紙が、エコロジの観点から大問題になっている。またプライバシーの侵害をどう防ぐかが重大な問題になっている。今後大きな問題となると思われるのは、電力消費量である。

進歩のそれぞれの時点において、適当な電力消費量で、解ける範囲の問題というのはある。それよりはるかに大きな問題を解こうとすれば、特別に多大な投資をしてコンピュータを設計して、多大の電力を消費すればできるであろう。しかしながらそれが人類の望んでいることであろうか。現在の技術進歩をみていけば、5年あるいは10年すれば同じ問題をはるかに安く解けるのである。そうすると、超並列コンピュータを作るよりも既存のワークステーションをたくさん使うクラスタのほうが、エコロジ的には許容されやすい。いわゆるビッグ・サイエンスも1980年代のバブルとともに終焉するのであるか。

6. おわりに

1990年代に伸びる技術はマイクロプロセッサと通信技術である。特に通信技術はこの十年間で1000倍にもなり、マイクロプロセッサの性能の伸びよりもはるかに大きい。ちなみにマイクロプロセッサの性能は数十倍になると考えられている。この意味するところは大きい。今まで通信がボトルネックであると思われていたシステムも、将来は可能であるということである。この進歩で一番有望であると思われるコンピュータ・システムはクラスタである。クラスタは、一番安いパソコンやワークステーションを超高性能のネットワーク

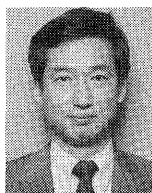
でつなげたものである。こういうシステムは現在でも使われているが現在はネットワークがボトルネックであるといわれているが、将来はそういう問題もなくなっていく。

このようにマイクロプロセッサが超高性能になり安いパソコンの中に入り、またマイクロプロセッサを数十台つなげた中規模並列コンピュータが簡単に手に入り使われるようになると、ベクトル型のスーパーコンピュータや超並列コンピュータの使用用途は特別に大きな問題を解くときに限られてくる。そのような問題でもクラスタが代替機としていつも比較検討されなければならないであろう。

参 考 文 献

- 1) 鈴木則久:「CPU チップの主導権」別冊日経サイエンス「スーパーパソコン」Vol. 104 (Oct. 1992).
- 2) Hennesy, J.L. and Jouppi, N.P.: Computer Technology and Architecture: An Evolving Interaction, *Computer*, Vol. 24, No. 9, pp. 18-29, IEEE Computer Society (Sep. 1991).
- 3) Anderson, T.E. et al.: High Speed Switch Scheduling for Local Area Networks, *ASPLOS-V Proceedings*, pp. 98-110, ACM (Oct. 1992).
- 4) Dongarra, J.J. et al.: Performance of Various Computers Using Standard Linear Equations Software, *Computer Architecture News*, Vol. 20, No. 3, pp. 22-44, SIGARCH ACM (June 1992).
- 5) Tago, K. et al.: A Proposal for an Operating System Designed for Cluster Servers, *J. of Information Processing*, Vol. 14, No. 4 (1991).

(平成5年4月27日受付)



鈴木 則久 (正会員)

1946年生まれ。1969年東京大学物理工学科卒業。1971年東京大学大学院工学系電子工学専門課程修士卒業。1975年スタンフォード大学大学院コンピュータ・サイエンス学科博士課程修了。その後カーネギー・メロン大学、ゼロックス・パロ・アルト研究所、東京大学で、プログラム言語、プログラムの検証、コンピュータ・アーキテクチャ、並列処理、データベース、オブジェクト指向などの研究に従事。現在日本アイ・ビー・エム東京基礎研究所所長。Ph. D.