

21世紀に生き残る計算機センターを目指して

姫野, 龍太郎
理化学研究所情報基盤研究部情報環境室

<https://doi.org/10.15017/1470465>

出版情報 : 九州大学情報基盤センター広報 : 全国共同利用版. 2 (3), pp.181-193, 2002-11. 九州大学
情報基盤センター
バージョン :
権利関係 :

21世紀に生き残る計算機センターを目指して

姫野龍太郎[†]

1. はじめに

PC(パソコン)の急激な性能向上と普及で、計算機センターの役割は大きく変化してきている。また、スーパーSINETなどの高速なネットワークと、世界最速のスーパーコンピュータ：地球シミュレータの両方が、わが国で来年早々稼動を始める。さらには国立の研究所や大学の独立行政法人化、特殊法人の近々の廃止など、計算機センターを取り巻く情勢は嵐のような変化を迎えている。私は理化学研究所の計算機センターを預かる身である。このような情勢の中で、計算機センターはどのような方向へ行けばいいのか、私見をまとめてみた。

最初にこれまでの計算機センターの役割を振り返ってみる。そして、スーパーコンピュータとPCの計算性能の変化を調べてみた。さらに、そのPCで構成したクラスターの性能が現在どの程度の実効性能かを示す。その次に現在、理化学研究所のスーパーコンピュータのユーザー統計を挙げ、そこから今後の運営の指針を導く。この過程では一般小売の業態や放送事業なども参考にした。最後に、ここで導いた運営指針のために日本全国の計算機センターに協力を呼びかけたい。

2. 計算機センターの役割の変化

2.1. 黎明期からワークステーションの登場まで

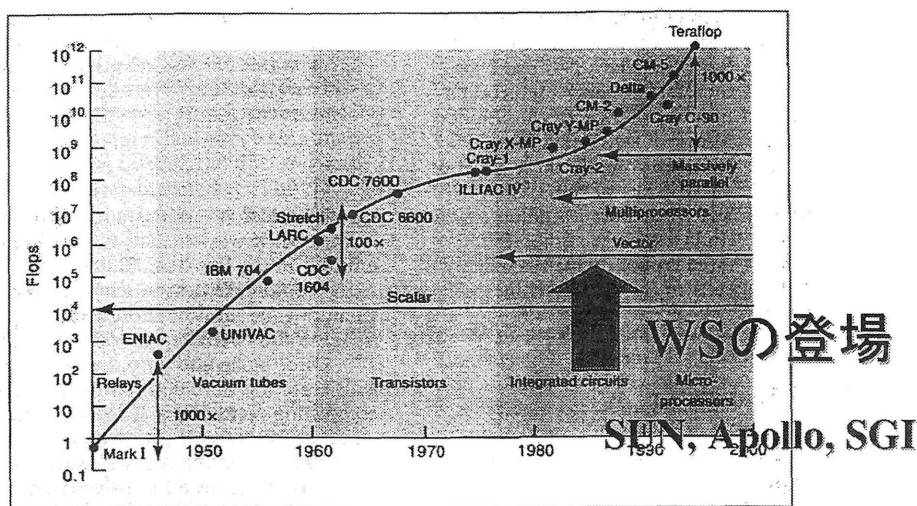
そもそも大型計算機センターは何のために設立されたのだろうか。それはコンピュータが非常に高価で、しかも非常に大きく、既設のビルにはとても納まらず、また大容量の空調機も必要であった時代のことである。運用にも特別な知識が必要だったので、専門の担当者において、みんなで共同して使うしか方法がなかった。当然ユーザーは使えることで満足し、数日待った結果がプログラムの一文字のパンチミスだったということも、当時は他に全く手段がないために、納得していた。紙テープやパンチカード、リール式の磁気テープ、ラインプリンターが計算機センターでは使われていた。紙製のカードを入れるジュラルミン製のケースやプリントアウトされたプリンター用紙の分厚い束、オー

[†] 理化学研究所 情報基盤研究部情報環境室 E-mail: himeno@postman.riken.go.jp

この原稿は、平成13年7月2日に情報基盤センターで行なわれたスーパーコンピューティングユーザー会の特別講演を文章に起こしたものです。

プリンルールの磁気テープなど、今では全く見られなくなったが、懐かしく感じられる方も多いかもしれない。これらは日本では 1960 代から 80 年代にかけての出来事だと思う。この時代、ユーザーは計算機センターに出向いて使っていた。

1980 年代半ばに入ると日本でもスーパーコンピュータが導入されるようになってきた。このベクトル型のスーパーコンピュータはそれまでの大型コンピュータの 10 倍以上の性能を示す、文字通り桁違いの性能を持っていた。計算機センターではカードリーダーが消え、ネットワークによってユーザーは徐々に研究室から離れずに使えるようになってきた。



The speed of digital computers has increased by more than a trillion times during the last five decades. The exponential increase in speed (measured in flops) has continued at a steady pace even though the underlying hardware technology and the supercomputer architecture has undergone major shifts.

図 1. コンピュータの性能向上の歴史[1]

ユーザーと計算機センターの蜜月はいつ終わったのだろうか。図 1 はコンピュータの黎明期から 1990 年代半ばまでの発展の歴史を示している。図の原著には示されていないが、計算機センターとユーザーの間柄を変化させたのは、1980 年代半ばに登場したワークステーションである。ワークステーションはそれまでのミニコンピュータと異なり、研究室でも無理すれば購入できる金額で、しかも、設置するのに特別な設備が必要ではなかった。このため、大型計算機センターでの課金でお金を使うよりも、このワークステーションを購入した方が経済的だというユーザーが現れ始めたのだ。もちろんワークステーションの方が計算速度は相当遅いのであるが、大型計算機センターでもジョブ投入から結

果を得るまでの待ち時間で見ると、混んでいる時はそれほど大きな差はなかった。特に計算機センターでは、年末から 3 月までの卒業研究の時期には非常に混み合い、結果を得るまでの待ち時間がひどく長くなっていた（今もだろうか）。

このときのユーザーの不満は、(1) 高価な課金、(2) 長い待ち時間、(3) ベクトル化のための書き換えの負担であった。スーパーコンピュータの速度が速いのはベクトル化による速度向上が大きく、スカラー計算であればワークステーションの速度とそれほど変わらない。このため、ベクトル化を意識したくないユーザーや、ベクトル化が効かないプログラムを持つユーザーには評判が悪かったわけである。とはいえ、ワークステーションは当初 1000 万円以上したので、簡単に購入できる額ではなく、不満はあっても計算機センターは使われ続けた。

2.2. PC の台頭と並列化

1990 年代に入るとワークステーションの値段が下がり、ユーザーにも普及し始めた。一方、21 世紀を目前にした 90 年代末期には PC が急激に速度向上を果たし、ワークステーションに迫る勢いである。ワークステーションは追う立場から、急に追われる立場になり、存在価値を問われるようになってきた。

このような情勢を情緒的ではなく、客観的に示そう。このためには PC からワークステーション、スーパーコンピュータまでを同じ尺度で測る指標が必要である。そこで、筆者が作成したベンチマークテスト (himenobMT [2]) を使う。これは熱伝導方程式や非圧縮性流体の圧力方程式など、物理から工学までの広い範囲で出てくる Poisson 方程式を 3 次元の一般座標系を使う差分法で解くもので、ベクトル化と並列化が完全に行えるように Jacobi 法を使っている。このベンチマークテストの特徴は、測定結果を相対値ではなく、絶対値として MFLOPS で与える。また、FORTRAN と C のソースコードがあり、それぞれ逐次実行版と MPI で並列化された並列実行版もある。全体で 100 行程度のものなので、MPI の練習用としても使えるものだと思う。さらに PC/AT 互換機用と MAC (Machintosh) 用に実行形式のものも用意されている。測定結果は web に一覧表として整理されており、PC や MAC からスーパーコンピュータに至るまでを比較できる。

さて、ベクトル計算機とスカラー計算機の性格の違いが如実に現れるのは、たとえば科学技術計算で多用される $A(I)=B(I)*C(I)+D(I)$ などという四則演算を DO ループで繰り返したときである。図 2 は NEC の SX4 と HP (Hewlett Packard) のワークステーション HP735 とで、この演算を配列サイズの小さいところから大きなところまで測定し、両者の特性を比較したものである。ちなみに、このテストは cache test として himenobMT と同じホームページに掲載している。図は両者ともに配列サイズを横軸に Byte で表し、縦軸に MFLOPS で性能を表示している。この図は性能の直接の比較を意図したものではなく、縦軸の目盛は両

者で異なる。特性の違いはあきらかで、ベクトル計算機はサイズ（ループ長）が小さいと性能は出ず、サイズ（ループ長）が十分大きいと一定値になる。スカラー計算機はこの逆で、サイズが小さく、キャッシュに納まると性能を発揮するが、サイズが大きくなってキャッシュに納まらなくなると性能は大きくダウンする。このような特性の違いから、性能測定を行うときにはプログラムのサイズについて十分な注意が必要で、himenoBMT では最小の問題でも 20MB を超えるように作っている。キャッシュのサイズは徐々に大きくなってきており、これでも十分大きいといえない時代がくるかもしれない。

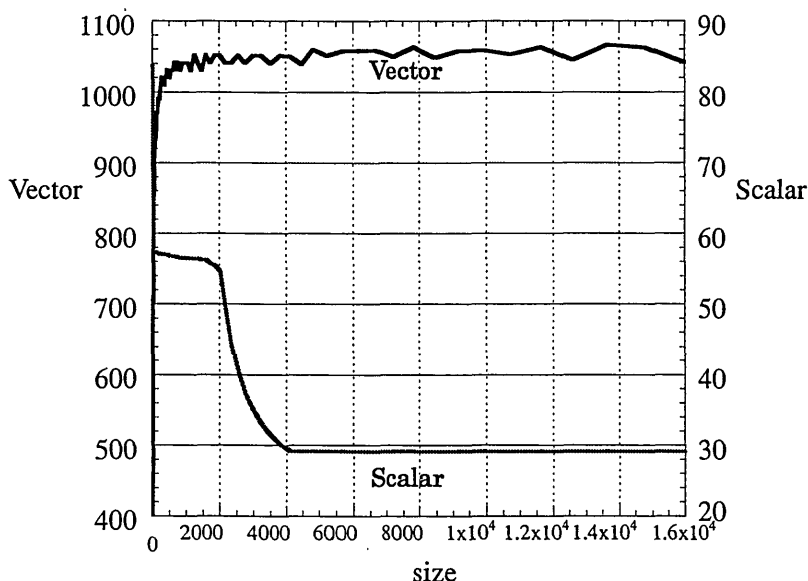


図 2.ベクトル計算機とスカラー計算機の違い

さて、前置きが長くなったが、ベクトル型スーパーコンピュータの 1CPU（または node）の himenoBMT での性能測定の結果を図 3 に示そう。同じベクトル型ではほぼ同じ理論性能を謳うマシンであるが、メモリーと演算器の間のデータ供給能力（メモリーバンド幅）の違いで、かなり大きな差が生じている。一方のワークステーションと PC ではどうだろう。図 4 は Intel の Pentium4、Compaq の Alpha Server、SGI の ORIGIN2000、AMD の Athron 1 GHz を比較したものである。このテストに関しては Pentium4 が圧倒的に速い。すでにクロック周波数ではワークステーションの CPU は PC に負けている。PC で弱いといわれてきた浮動小数

点演算やメモリーバンド幅も、今やワークステーションを完全に凌駕している。PCはワークステーションに比べ生産台数が一桁以上大きく、コストの点で絶対的に有利である。現在、PCとワークステーションはほぼ同じ使われ方をしているため、コストで不利なワークステーションが生き残れるシナリオが見つからない。PCのグラフィック性能も向上してきた今、ワークステーションは、もう使命を終えたようだ。

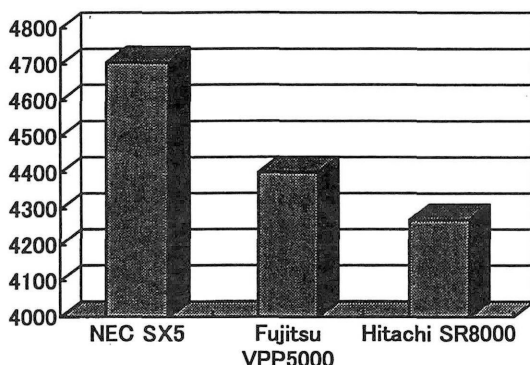


図3.ベクトル機の1CPU(node)の性能(himenoBMT、MFLOPS)

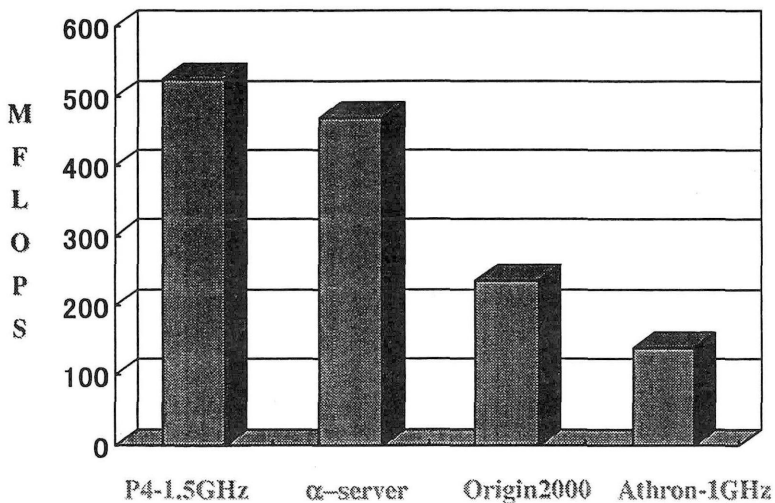


図4.PCとワークステーションの比較 (himenoBMT 結果)

いよいよ、ベクトル型スーパーコンピュータと Intel 製 PC との性能差が近年どう推移してきたかを見てみよう。ベクトル型スーパーコンピュータはシステム全体の性能ではなく、要素プロセッサとしての 1CPU (node) と性能を比較す

る。図5が1994年から2000年までの7年間の変化である。1994年当時、ベクトル機とPCでは約100倍の性能差があったものが、2000年には10倍にまで縮まった。その性能向上のトレンドがこのまま続けば、今後さらにPCの性能は上がり、差が縮まってゆくことだろう。このPCの性能向上にはマルチメディア用のストリーミング命令が寄与している。このストリーミング命令（SSE, SSE-2）とはベクトル命令に他ならない。AppleのPowerPCにもVelocityEngineと呼ばれるベクトル機構が組み込まれ、性能向上を図られている。しかし、残念ながらコンパイラが今のところ対応していない。

一方のベクトル機の方は、これまでプロセッサが1つのLSIに納まらないマルチチップ構成なので、クロック周波数を上げることができなかった。このため、これまでの速度向上は演算器を増やして対応してきた。この方法では限界があり、今後の速度向上には演算器を減らしてでも1チップ化してクロック周波数を上げるしかないだろう。

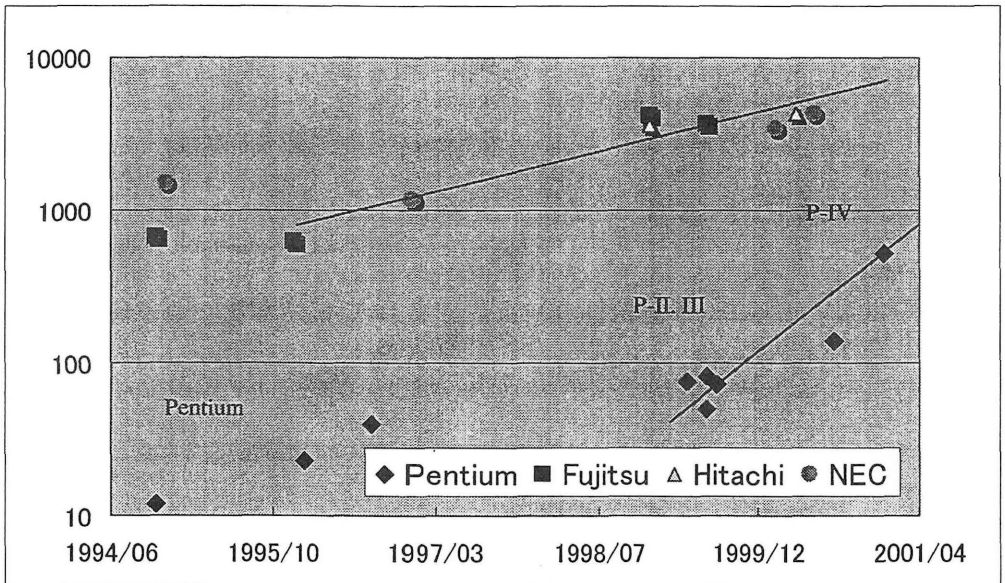


図5.ベクトル機とIntel製CPUの速度の推移(himenoBMT、縦軸はMFLOPS)

2.3. PC クラスタ

これまでにPCはスーパーコンピュータの1/10の実効性能があることを示した。そのPCをネットワークで結合させて並列計算機として使うのがPCクラスターである。単純な100Base-Tのスイッチングハブを利用した安価なシステムでも8台のPCで6倍以上の性能が出る。筆者のところで、64台のPentium 4、1.7GHzをMyrinetで結合したPCクラスターを導入予定で、このシステムでのテ

ストでは himenoBMT で 30GFLOPS を超える実効性能を出した。これらは本当にスーパーコンピュータと呼べる性能である。現在、PC クラスタでは 512 台構成 1024CPU の規模のものが稼動している。

この PC クラスタは計算機センターにとって大きな脅威である。ユーザーは自分の計算規模に合わせて、小さいものから大きなものまで必要な規模のクラスタを構成できる。要素プロセッサは PC であるから、価格は従来のコンピュータに比べ、桁違いに安い。OS やコンパイラはフリーで手に入る。特に小規模なクラスタは Myrinet などの特殊なネットワーク機器を使わなくてもよいので、コストパフォーマンスに優れたシステムを作り上げることができる。たとえば 8 台から 16 台程度のクラスタであれば、PC やスイッチングハブを集めてきて、図 6 のように簡単に組み上げられる。

この先もユーザーは計算機センターを使うのであろうか。それとも、各研究室に必要な計算機資源を確保するようになるのだろうか。

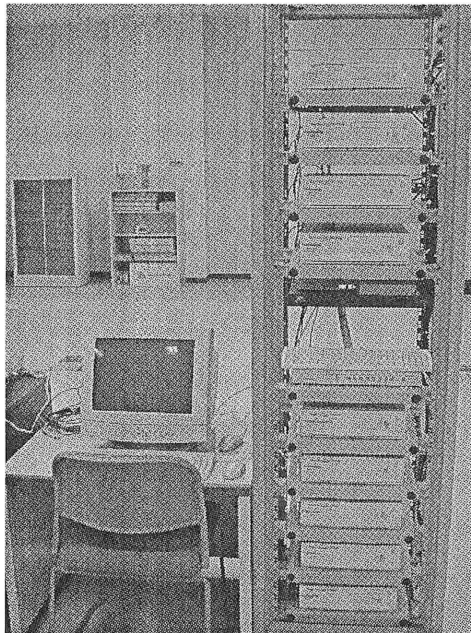


図 6. 手作り PC クラスタの例 (8 台の PC と制御用 PC の 9 台構成)

3. 理化学研究所のスーパーコンピュータの利用実態

われわれのセンターでの使われ方を紹介しよう。理化学研究所のスーパーコンピュータは VPP700E の 160PE 構成である。ユーザー数は約 100 人、うち半数

は理研と共同研究を行っている外部のユーザーで、海外からのユーザーも全体の1/3を占めている。現在、CPUへの課金は行っておらず、ディスクとテープの使用にのみ課金している。

図7は2000年の統計数値で、会話型の使用比率を示している。やはり会話型の利用はデバッグやテストには便利なようである。この会話型では並列ジョブも利用でき、かなり多くのユーザーが実際に並列ジョブのデバッグ・テストに利用している(図8)。

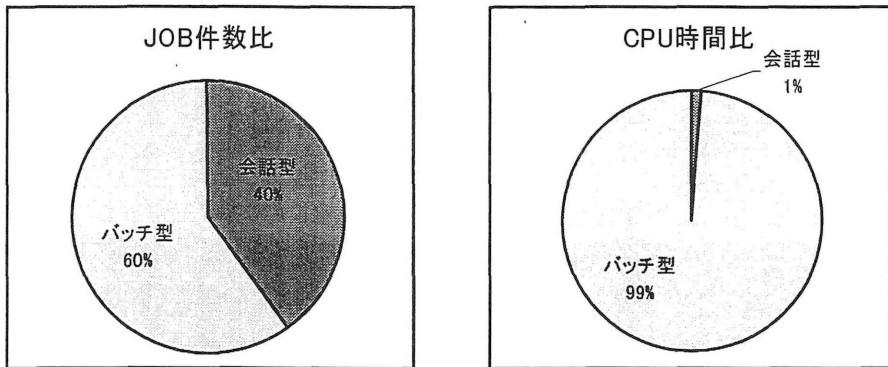


図7. 会話型とバッチ型のジョブ件数比とCPU時間比

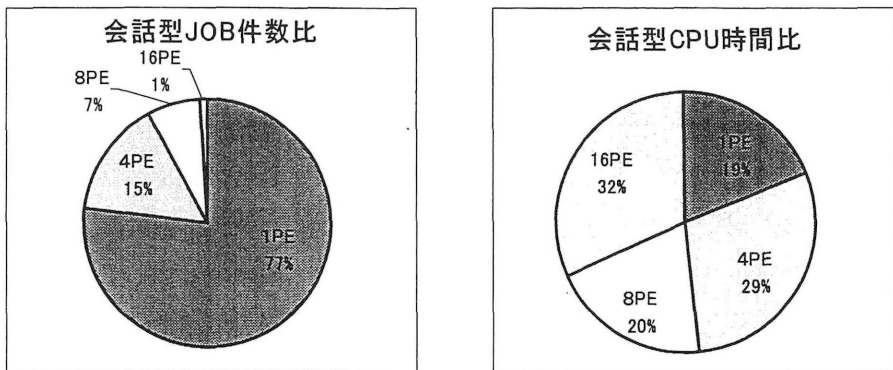


図8. 会話型ジョブのPE数ごとの件数比とCPU時間比

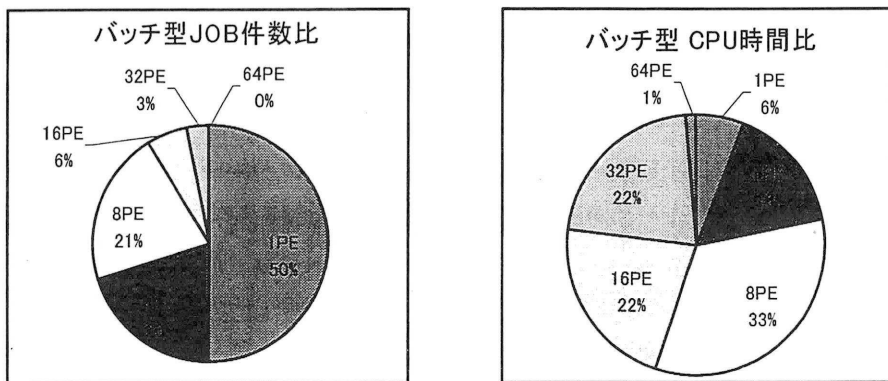


図9.バッチ型のPE数ごとのジョブ件数比とCPU時間比

図9はバッチ型ジョブの使われ方を示しており、CPU時間で見ると並列計算が主流になっていることが分かる。しかしながら、ジョブ件数で見ると半分のジョブが並列化されていない。理研のセンターでは商用パッケージの利用はGAUSSIANが主で、他の利用はほとんどない。このGAUSSIANは並列化されているので、ユーザー自身のプログラム（フリーウェアなども含む）が並列化されていないものも、まだ、かなりあると見ている。

図10は計算資源をどれくらいユーザーが使っているのかを見たものである。これから、上位のたった3人で全資源の1/3を使っていることが分かる。上位10人で広げると2/3が、上位20人では90%にもなり、スーパーコンピュータは少数の研究者で使われていることが分かる。この傾向は特殊なものではなく、普遍的なものではないかと考えている。

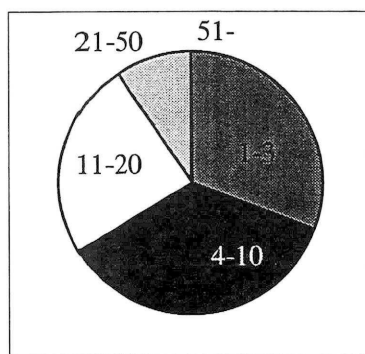


図10.上位何番目のユーザーがどれくらい計算機資源を使っているか

4. 今後の行方

4.1. ユーザーの行方

全コンピュータのユーザーを考えたとき、図 11 に示したように、現在ワークステーション（図中では WS）、あるいはグラフィック・ワークステーション（図中では GWS）を使っているユーザーは、すでに述べた理由で性能の向上する PC か、あるいはクラスターの方へ移ってゆくだろう。現在ワークステーションのサーバーを使っているユーザーは図の中では SMP として表示しているセグメントであるが、これも徐々に性能の向上する PC のクラスターに移るだろう。

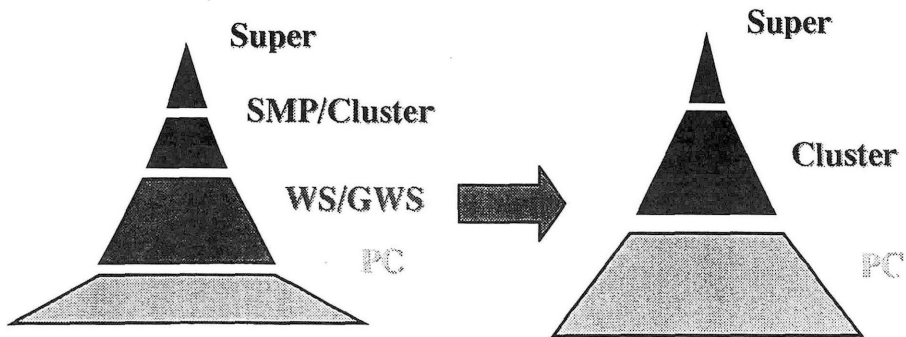


図 11.今後のユーザーの行方

スーパーコンピュータにはどのようなユーザーが残るのであろうか。その一つは使いたい市販パッケージが簡単に使えるとか、煩わしいコンピュータの管理がないという理由から残るユーザーのグループであろう。スーパーコンピュータでなければ計算できないというユーザーも、もちろん少数ではあるが、残るに違いない。計算機センターとしては、この少数の、他に代えられないユーザーを積極的に支援してゆくことが肝心と考えている。理研のセンターとしては、彼らのニーズを積極的に反映し、今後のセンターをより良いものとしてゆく。

4.2. 計算機センターの行方

今年度中に世界最速のスーパーコンピュータ：地球シミュレータが稼動を始める。計算機は規模のメリットがあるので、日本にいくつもある計算機センターを統合して大きなセンターを作った方がいいような感じがするだろう。計算機の速度・容量・専門スタッフなど、どれも集約して大規模化するメリットを

持つ。しかし、果たしてそうなのだろうか。よく考えると、これは二つの意味で間違いである。ひとつはコストパフォーマンスの問題、もうひとつはユーザーの使い勝手の問題である。

ハイエンドコンピュータはまず開発コストが非常に大きく、しかも台数が出ないので、開発費を回収する母体が小さく、その結果、非常に高くつく。また、たとえばクロスバススイッチや、密結合の場合のメモリからプロセッサまでの配線など、半分の規模で作ればコストが 1/4 になるものがあり、コストパフォーマンスが必然的に悪い。また、部品数が非常に多くなる結果、製品の信頼性が低くなり、初期故障だけでなく、運転時の故障が多くなる。結果的に稼働率が下がり、コストを上げる要因にもなる。

大型のスーパーコンピュータの価値は、そこでしかできない計算を行うことにある。100人のユーザーが同時に使えば 1/100 の性能のコンピュータと同程度のことしかできない。これではまとめて大きくする意味がない。数人のユーザーがあたかも占有しているような環境を提供し、かつ、そのための小回りの効く運用を提供して、初めて大規模化する価値が出てくる。しかし、日本に数箇所のセンターに集約した場合、数人のユーザーに占有させたり、大容量のディスクを提供したり、と言うわがままが許されるだろうか。あるいは急いで結果を出したいという 1 ユーザーの願いを聞いて、他のジョブを止め、優先順位を上げて優遇するようなことができるだろうか。残念ながら、大型化することは自由度を制限することにつながる。ほかに方法はないのだろうか。

百貨店方式の従来型計算機センター

これまでの計算機センターは、いろいろなジャンルの商品をすべて用意し、しかも、それぞれの商品は出店者に売ってもらい、自らは場所を提供して店賃を貰う百貨店とよく似ている。この商売は出店者（アプリケーションやパッケージソフト）に依存し、付加価値が低く、この不況下では競争に敗れていく例が多い。現在のソフトウェアはそれぞれプラットフォームとなる計算機によって速度が異なる。流体計算などはベクトル計算機だと非常に速いが、ゲノムサーチやモンテカルロ法を使う粒子計算などは遅く、スカラー計算機や PC クラスターなどが速い。もし、現在のまま、どれもアプリケーションとして取り入れようとする、それぞれに向けた計算機を複数用意するか、あるいは、どれもそこそこのレベルだが、どれにも中途半端な計算機を用意するかを選択を迫られる。

ショッピングモール方式の計算機センター・ネットワーク

今消費者に受け入れられている小売のスタイルは、ショッピングモールとコ

ンビニである。大規模な駐車場を備え、大型の専門店が集まったショッピングモールは週末ともなると大勢の消費者を集める。一箇所ではいろいろな商品を買うことができるだけでなく、それぞれの商品の種類が豊富で、しかも家族が一日十分に楽しめるエンターテイメントも用意されている。この商品種類の多さと駐車場の広さは、同じように専門店が集まった駅前商店街とは異なる。このような計算機センターはできないだろうか。

非常に高速なネットワークで結合された、それぞれに特徴のある計算機センター群を考えよう。今はもうユーザーは計算機センターに向向くことなく、ほとんどネットワーク経由で利用している。高速なネットワークは大規模な駐車場に匹敵する価値があろう。それぞれの分野に特化した特徴ある計算機センターは大型の専門店にたとえられる。どれにも中途半端な性能を出す計算機の代わりに、ある分野では最適の計算機を選ぶ。それ以外のジョブは、それぞれのジョブが得意なセンターと提携し、ネットワーク経由で計算の実行を依頼、結果を返してもらおう。ハードウェアに限らず、アプリケーション・ソフトや性能向上のための相談・チューニングなどの専門要員もより深く対応が可能となる。規模が大きくなっても、専門性が深まるため、ユーザーにとってもメリットが大きくなる。このような特色ある計算機センターの連携した未来がひとつの未来ではないだろうか。

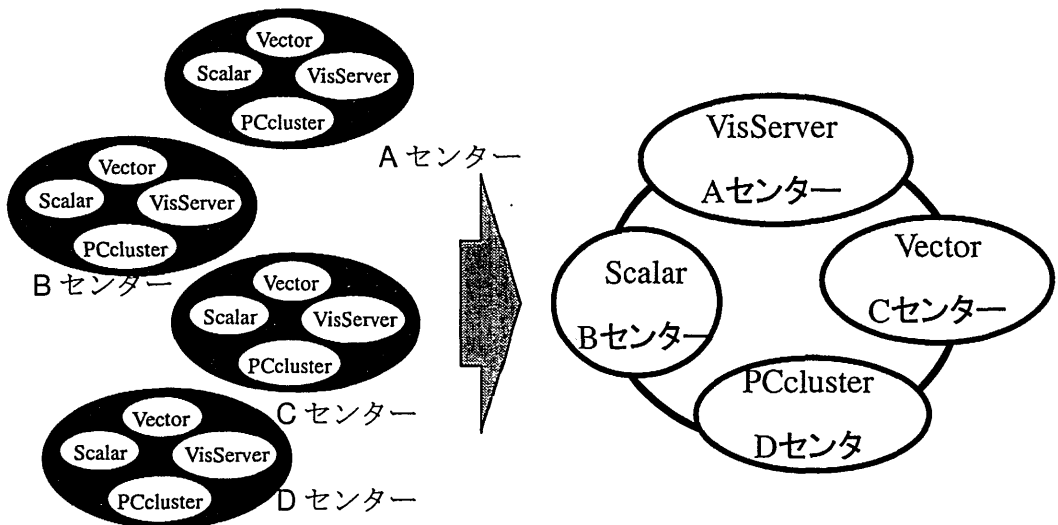


図 12.連携しない計算機センターから連携した計算センターへ

理研ではモレキュラー・ダイナミクス専用計算機(MDGrape)があり、また、生体力学シミュレーション研究で築いた流体構造連成シミュレーションがある。産業技術総合研

研究所の計算機センターとの間では、理研の計算化学の分野の GAUSSIAN ジョブを引き受けてもらう代わりに、MD 計算や流体、構造計算を引き受けるというような連携について話し合いをはじめた。都合のよいことにスーパーSINET が近々稼動する上、ITBL(Information-Technology-Based Laboratories)プロジェクトもこの4月から始まった。ITBL は旧科学技術庁傘下の6研究機関(原研、理研、JST、防災研、旧金属材料研、航技研)が集まり始めたもので、このショッピングモール形式の計算機センター・ネットワークを作るプロジェクトの側面も持つ。

5. 終わりに

これまでのコンピュータの性能の発展と計算機センターの役割の変化を調べてみた。その結果、計算機センターは従来の百貨店方式から、特徴のある計算機センターが高速ネットワークで連携したショッピングモール方式に移ったほうが明るい未来が来ることを示した。ぜひ、九州大学を始め、各計算機センターにも協力をいただき、推進してゆきたい。

参考文献

- [1] William J. Kaufmann III, Larry L. Smarr:
Supercomputing and the transformation of science, Scientific American Library, 1993.
- [2] <http://w3cic.riken.go.jp/>