The University of Osaka
Institutional Knowledge Archive

Title	大量のデータを扱う実験屋としてセンターに期待すること
Author(s)	山中, 卓
Citation	大阪大学大型計算機センターニュース. 1998, 107, p. 14-16
Version Type	VoR
URL	https://hdl.handle.net/11094/66250
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

特集

大量のデータを扱う実験屋として センターに期待すること

大阪大学大学院理学研究科 物理 長島研

山中 卓

1 高エネルギー物理学における計算機の使い方

我々の研究室は高エネルギー物理学、即ち物質を構成する素粒子に関する実験を行っている。例えば、世界最大規模の陽子加速器や電子加速器を用いて特定の粒子を作ってその崩壊を観測することにより素粒子間に働く弱い力の対象性の破れを調べている。また、地下にある5万トンの水槽を用いて太陽や上空の大気で作られたニュートリノを観測することにより、ニュートリノの性質を調べている。

さて、これらの実験はいずれも大がかりであり、その解析のためには大量のデータの処理を必要とする。例えば、米国で行っているK中間子を用いた実験では約60Tera byteのデータをテープに書いたし、来年から日本で始めるB中間子を用いた実験でも1年間に数十Tera byteのデータをとる予定である。これらの大量のデータはまず簡単な解析を行って欲しい事象の候補を選び、データ量を約一桁落としたテープを作る。このテープの詳しい解析を行った後、更に情報を圧縮したテープを作って解析を進めていく。解析の最終段階はワークステーションなどで十分まかなえるが、それ以前の段階ではより大規模の計算機が必要である。

我々の実験で用いる計算機のもう一つの主な用途は実験のモンテカルロ・シミュレーションである。これは計算機上で目的の事象を発生させ、これで生じる多くの粒子の飛跡を追いかけ、各粒子の、物質や測定器との反応をシミュレートする。こうした事象を大量に作ることにより、事象の検出効率を求めたり、実験で見えるバックグラウンドの見積もりを行ったりする。こうしたシミュレーションは大がかりなものだと約1－2ヶ月はかかる。

2 実験屋の使う計算機に要求される事

さて、我々の実験でこうした目的のために用いる計算機には次のような機能や使用環境が要求される。

1. 高速のスカラーマシンである。上で述べた様な用途では計算をベクトル化する事はほとんどできない。
2. 高速のデータ入出力が可能である。データの入出力先はDLTなどの磁気テープとディスクである。例えばK中間子の実験の場合、一つの事象を解析するのに要する時間が数msec、一つの事象が8kbyteであるから、2Mbyte/sec以上の速度でデータの読み書きができないとCPUによる計算時間よりもデータの読み書きに時間がかかってしまう。手軽にこうした入出力速度を得るためには、ネットワークを通して行っていたのでは追いつかないので、テープやディスクドライブが直接SCSIを通して計算機につながっている必要がある。
3. テープが集中的に管理され、かつ「自動的に」ドライブにマウントされる。実験で書かれた生のテープは数千本、一旦選別されたテープでも数百本あり、これらの数は実験が更に高感度になるにつれ増大する。従って、このようなテープをきちんと管理できる体制が必要である。また、ユーザーがプログラムを走らせて特定のテープを要求すると、それがユーザーから見て自動的

に所定のドライブにマウントされ、プログラムが終了するとまたもとの場所に戻される必要がある。この作業にはテープロボットを用いる、オペレータの人の手を用いる、などの方法がある。後者の場合は経験上、オペレータの質によって対応の時間が遅かったり間違ったテープをマウントされる場合がある。

4. プログラムを走らせるモードとして、interactive にその場で走らせる方法とバッチジョブを走らせる方法の2種をサポートしている。大量のデータをどんどん処理していくには、バッチジョブの形態を用いて、ある本数のテープを処理するジョブが順次走るようにするのが効率的である。しかし、そうしたジョブを開発するときや何か問題が起きたときには、デバッガを起動して interactive にプログラムを走らせる方が圧倒的に効率が高い。特に Silicon Graphics, DEC などの UNIX ではバッチで既に走らせているプロセスに対しても dbx -p *pid*, dbx -pid *pid* などでデバッガーをかけることができ、大変便利である。
5. 単純なシステムである。これは上で述べたことも含むが、ユーザーがあたかも普通のワークステーションにログインするように入り、そこに自分の全てのファイルが interactive,batch を問わず見えており、そのシステムを使うために特に複雑なことをしなくてもよいという環境が必要である。なまじっかワークステーションクラスターだの、分散処理などを持ち込むとシステムが複雑になり、使いにくい。実際、日本のある研究所に導入したクラスターのシステムはうまく機能していなかった。

こうした要求を満たすシステムとして我々が用いているのは多数の CPU が、テープ、ディスクドライブの直結した一つの箱に入り、メモリーを共有しているものである。この箱では単に普通の UNIX が走っているので、自分の机の上のワークステーションと変わらない環境で使うことができる。そして、ドライブが箱に直結していることによってデータの入出力の速度を高めている。また、多数の CPU がドライブとメモリを共有する事により、資源を有効に利用することができる。反対に例えば、単にそれぞれテープとディスクのドライブのついたワークステーションを並べるだけだと、ワークステーションによっては使われないテープドライブがあったり、メモリーが余ったり足りなかったりするし、速度を上げるために同じデータのコピーをそれぞれのワークステーションのローカルディスクに置いたりする無駄が生じる。また、CPU の負荷も不均一になる可能性が高い。但し、こうした問題は将来より高速のネットワークスィッチなどを使えば解決される問題かもしれない。

具体的に米国の実験で用いているシステムは、433MHz の CPU が 10 個、メモリーが 1Gbyte 入った Alpha 8400 のサーバー2台に、それぞれに DLT のテープドライブを約十台、ハードディスクを 340~500Gbyte つけたものである。これでも十分とは言い難いが、それでも８０人の実験グループのデータ処理、解析、シミュレーションのほとんどを担っている。テープのマウントはオペレータの手によっている。日本のＢ中間子の実験のためには筑波の研究所で 28 個 CPU の入った Sun UltraEnterprise 6000 を 7 台つなげたシステムを用意し、テープドライブを 26 台これはネットワークを介してつなげている。

3　大型計算機センターに期待すること

さて、現在我々は阪大の大型計算機センターは全く使っていないが、上で述べたような要求を満たすシステムが導入されればそれを使う用意はある。こうしたシステムを自前（研究室、学科など）ではなく、センターで持つことの利点には次のようなことが考えられる。

1. システムが購入しやすい。これらのシステムは億円の単位の買い物となり、まずこうした財源は研究室、学科の単位ではなかなか得られない。運良く得られることがあるにしても、割のよい宝くじを待つようなもので、研究の計画をそうした不確かなものに依存させるのは難しい。従って、定常的もしくは当たりやすい予算を持つセンターで購入する方が確かである。また、１億４

千万円を越えると政府調達の対象となり、使用作成、官報公告、説明、入札など、その手続きは複雑である。こうした作業は我々研究者よりも、そうした作業に慣れておられるセンターの方などにやっていただいた方が効率的である。

2. 研究者のシステム管理の負荷が下がり、且つ管理が責任を持ってなされる。我々の研究室では、多くのワークステーションを買うよりも、多くのCPUの入ったマシンを購入することにより、管理するべきシステムの数を減らしている。こうしたシステム管理は大学院生や助手の手による事が多く、その労力を減らして本来の仕事である研究に専念させることは重要である。それでも依然、やれネットワークがおかしいだの、ディスクが壊れただの、バックアップを取るだの、と仕事はある。こうしたシステム管理の経験は、学生にとってはよりシステムを理解する機会となって、就職に役立つ利点は確かにある。しかし学生は常に研究室を巣立っていくので毎年super userの資格を持つ新人ができ、毎年何らかの事件が起きる。こうした点からも、専門のシステム管理者がいると大変ありがたい。

さて、こうしたシステム一つでは阪大全体の要求を満たすのは難しいであろうから、学部、もしくは学科単位でシステムを構成するのが適当であろう。これらのシステムは基本的にはセンターで購入するが、ユーザーの方で宝くじに当たった場合も本体をセンターに置き、管理をセンターに依頼できればなお良い。他のグループより特に大きな容量のディスクが必要な場合は自分達専用の物を各研究室で購入してシステムにつけられるようにするとよい。一つのシステム内で研究室間の資源の取り合いを制御するにはバッチジョブの枠を各研究室に割り当てるなどの方策が要るであろう。使用料はもちろん只に越したことはないが、校費でまかなえるくらいは仕方ないだろう。

研究室で到底買ったり管理したりする事のできないベクトル計算機などは、当然今後も大型計算機センターで面倒を見ていくべきである。ワークステーションの性能の向上という背景を受けて、ワークステーションと性能の変わらなくなってしまった汎用の大型計算機を捨て、ベクトル計算機に絞ったセンターの判断は間違っていないと思う。しかし、大量のデータ処理のできる大型のスカラー型計算機をサポートしてもらえれば、より多くの研究者のために役立ち、センターのユーザーも増やすことができると考える。