Ciência Téc. Vitiv. 6 (2): 45-65. 1987.

COMPUTERS IN SCIENTIFIC RESEARCH (*)

F. JAMES ROHLF
Department of Ecology and Evolution
State Udiversity of New York at Stony Brook
Stony Brook, NY 11794-5245

SUMMARY

This paper is concerned with the use of computers in scientific research and with some of the important considerations that need to be taken into account in order for an institution to plan for the provision of computing services to its researchers.

The features of large batch systems, minicomputers, microcomputers, and supercomputers are compared to give an overview of alternative approaches to supplying computing services to researchers. Examples of different types of applications of computers are also presented to illustrate how different users can have *very* different computing needs. These applications range from wordprocessing, text processing, statistical computing, laboratory automation, graphics, to large-scale simulations and modeling. Some of the latter applications may require very large amounts of processing time — even on modern supercomputers.

Finally, a number of points are discussed that should be taken into consideration when planning computing facilities — whether for an individual, a research laboratory, or an institute or university. These included the consideration of the importance of rapid response time from a computer (to maximize user efficiency), economies of scale (large centralized computer vs. distributed microcomputers), availability of good software, and availability of technical support.

INTRODUCTION

This paper is concerned with the role of computers in making researchers more productive. The level of productivity of researchers is important for many reasons. An important

---

(*) Presented at the Simpósio Internacional Investigação Científica e Desenvolvimento Económico. Oeiras, 1-3 October 1986.

reason, relevant to the present symposium, is that it will have an impact on the long-term potential of a country for a high-level of economic development.

First of all, the modes by which institutions have provided computing services in the past will be reviewed. Then some examples of different types of use of computers in scientific research will be presented to illustrate how computing needs vary greatly in different fields. No one computing facility can be optimal for all users. Thus a computing system must, to some extent, be a compromise of the needs of different groups. Finally, a number of important issues that should be considered when making institutional decisions about the amount of computer support for research and the form of the distribution of computing services will be discussed.

I believe that it is important that the reasons for wanting a particular computing system be stated and that the decisions be based on reasonable principles with a clear statement of the probable consequences. If one can't afford a system that will meet all of the computing needs, then it is important that the potential users understand this so they will not have false expectations. It is, however, becoming increasingly difficult to make optimal decisions since rapid developments make it difficult to make correct predictions about future directions in computing technology and their implications for users. Despite the advertisements one sees in the popular press, it takes effort to learn how to use a computer and it takes time to enter large databases into a computer and to develop software or to learn to use existing software and adapt it to your needs. It is unfair to get a group started on this process if the available computer does not have sufficient capacity for the expected level of use (especially since one can count on the fact that the level of usage will be higher than one expects). In such cases users will lose a lot of time and become very frustrated — and rightly so. This can create a problem for the administration since many more people now realize that they have a stake in the issues and may apply considerable pressure if the computer installation does not meet their expectations.

SOME HISTORY

It is useful to review, briefly, the stages that scientific computing has gone through. This helps one appreciate its current state and prepare for future developments in computing. In looking back, one can see that different eras can be classified into fairly clear classes with respect to the ways in which computing power was made available to the average user. The present time is more difficult to classify due to the recent rapid developments in computing technology.

— *Early computers:* When electronic digital computers became generally available in the mid to late 1950s they were seen mostly as tools for engineers. They were to be found in basements of buildings and protected and cared for by teams of technicians, users left either boxes of punched cards or punched paper tapes to be processed. In some cases the user could make an appointment to work with the computer for some specific period of time (very helpful when debugging programs). The computers of this era worked on only a single problem at a time. Potential processing time was wasted while the computer waited for the operator to read in the next job, for the next user to get setup, or simply for someone to empty its output hopper. Because of its slow processing speed and limited memory (although impressive at the time) everything was done to make the computer work with maximum efficiency. Most programmers had to be very familiar with the internal architecture of the machine and used every trick possible to squeeze as much information as possible into a small amount of memory and at the same time make the computer operate as quickly as possible. One did not speak of «user friendly» computing systems — it was up to the user to make things as easy as possible for the computer.

— *Large batch systems:* Memory was large enough on these systems so that users could submit jobs which were read into the computer ahead of time so that the computer could quickly start the next job as the previous one finished. Users would then come back later for results. If a job aborted due to a minor errar (*e. g.*, a comma was left out), then a user would have to try again several hours later.

— *Timesharing systems:* Users at terminals are connected to a single large computer. Since the computer rapidly gives each user a turn (a time slice), users get the impression that the computer is working on everyone's jobs at once. These computers became much more accessible since users could now have terminals in their labs or offices. The computer then started to become more of a part the normal research routine. The types of applications for computers greatly expanded during this era. Time-sharing requires a more complex operating system and more reliable hardware than were needed before. When a batch machine failed the user just experienced somewhat longer delays. But when a timesharing machine fails the users work comes to a stop and they have to do someting else until the computer is repaired. This is because the computer is now also needed for the preparation of input (both data and programs) for future computer runs.

— *Minicomputers:* Next, small minicomputers became available for use in laboratories. They were used for a variety of tasks but were found to be especially useful for the automation of experiments. They could be programmed to adjust the parameters of an experiment and then to record the results within the computer. With the data in the machine it was a natural step to analyse the data in the minicomputer rather than to somehow transfer it to the central mainframe computer. This meant the development of special software for these minicomputers to get around the problems of their small memory and slow processing speeds. But once the software is developed, an individual researcher no longer has to wait in a batch queue for results (of course one has to wait for the small slow minicomputer but that is not so bothersome since one is just waiting one's own results).

As this continued minicomputers were used for most of the functions previously given to the central mainframe. They have gone from being single-user systems to small multiuser timesharing systems (so their costs could be shared). They have now taken on many of the problems of the large multiple user systems and their operating systems have become quite complex.

— *Microcomputers:* These small personal computers have become phenomenally successful for at least two reasons. First, their low price has made it possible for individuals to have their own *real* computer (not just a toy) which gives their users a feeling of freedom from the frustrations of dealing whith large, complex shared systems. Second, there soon became available many high-quality programs that allowed users to do things that they had not been able to do before. For example, what you see is what you gest (WYSIWYG) word processing, spreadsheets, and interactive graphics). It was also important that at the time this equipment became available the programming practices changed and much more emphasis started being placed on software being helpful to the user («user friendly») and reliable than before. This was necessary since a high proportion of microcomputer users are now less technically trained in computing (most users are no longer programmers). For laboratory scientists it allowed data acquisition at even lower cost than on a minicomputer since standard analog to digital (A/D) conversion boards and supporting software were available. Since the equipment is so cheap it is possible to dedicate a microcomputer to a single experiment. With random access memory (RAM) and disk memory now so low in cost it is also possible to run standard scientific software for the analysis of experimental data — rather than having to develop special small versions of these programs. Of course, microcomputers execute programs more slowly, but that is often not a problem for routine data reduction and analysis.

— *Communication networks:* Many computer users have now found that they don't really want to be *completely* independent of the main computing system of their institutions. Some of the more important reasons are:

- — They have found that they now need more computing power and they are not ready to buy a larger microcomputer (there is little trade-in value on an old computer).
- — They would like to have access to expensive peripheral devices that they can't justify for just their own use

(*e. g.,* laser printers, photocomposition devices, 9-track tape drives, large plotters, etc.).

— Mainframe computers are now being connected to computer networks such as BITNET, EARN, JANET, etc. that allow a user to send electronic mail, manuscripts, data, and programs to users on other computers — whether across the same campus or around the world. Not only is this electronic mail much faster than sending the same information through the normal mails but it is simple and avoids problems of compatibility of magnetic tapes or floppy disks. However, unless a user has an account on a mainframe computer and logs on from time to time, they will not discover that someone has sent them something.

Many have been surprised at how the character of computer usage has changed over the last few years. No longer are the majority of users just interested in large-scale numerical computations. Now most users are interested in information processing — word processing, data bases, and electronic mail. This is because computers are now being used more by people who do not normally have problems that require large-scale calculations.

In 1985 it was estimated that there were about 600,000 local area networks (LANs) installed and that over the next 3 years that number could grow to 5 million (Cowart, 1985). Most local networks have been installed in offices but their use is spreading to laboratories and universities. These local networks are now being connected to national and international networks through gateway computers. A variety of physical media are being used for networks but baseband (*e. g.,* Ethernet, IEEE 802.3 standard) has been used most commonly. Fiber-optic networks are now also being installed quite commonly.

— *Supercomputers:* Despite all the recent emphasis on microcomputers, one must not ignore the fact that in some fields research computations are becoming even more demanding upon the calculational power of a computer. As a result, very large and fast supercomputers have been developed in the last few years. These are very important for many applications in physics, chemistry, molecular biology, and engineering (see below for additional discussion). Researchers and institutions must take into consideration the availability of these very large and very powerful new computing systems. Supercomputers are expected to reach speeds of at least 1 billion instructions per second (BIPS) and several hundred milion floating point instructions per second (MFLOPS) through the use of multiple processors. They will also have much larger storage than present-day computers. These developments will have important implications for many fields of scientific research since it will greatly change the nature of what one considers to be a reasonable-sized computational problem.

Interesting current development are the mini-supercomputers and personal supercomputers (the latter are expected to cost around $125,000). The advantage of the personal supercomputer is that they would permit interactive applications to be run that require supercomputer performance (such as very complex graphics). Researchers now usually wait days for access to blocks of time on a supercomputer. It is not clear whether the smaller supercomputers will be economically successful. The problem is that no matter how much CPU time one makes available to some researchers it is never enough. For some applications access to the largest available supercomputers is always needed. The difference is important since the speeds of different systems may differ by factors of 50 to 100.

The future?

One cannot help but be impressed with the constant increase in technological developments in recent years. The expression «technology push» is sometimes used to describe the fact that, whether we like it or not, technological changes keep coming. It is hard (impossible?) to try to hold it back without it having a serious impact upon the character of an institution.

A problem for the planner is *options shock* — there are just too many choices presented by the different vendors. A careful evaluation can be more than a full-time job. This can be a real problem. One feels like giving up and deciding on just one vendor and then buying whatever they sell. This feeling seems to be encouraged by some vendors. It does save the buyer from having to keep up with all the new products and studying ther advantages and disadvantages. Even salesmen for a single large company have to work full time to keep up with the detailed specifications and interactions between their own products. This is an important task since the different options must work together in a compatible and coordinated way in order to be efficient.

APPLICATIONS OF COMPUTERS IN SCIENTIFIC RESEARCH

Several areas of applications of computers are described below. In addition to research calculations *per se*, applications in support of scientific research are also considered.

Statistical data analysis

It goes almost without saying that routine analysis of experimental data is more efficiently performed with the use of high-quality statistical software. There are now many such program packages. Standard software now makes it almost routine to perform efficient analyses of complex unbalanced designs which previously either could not be analyzed or else coul only be approximately analyzed since the correct methods involved so much computation. Multivariate analysis is a branch of statistics in which applications are not really feasible without access to a computer to help setup the data sets and to perform the computations. But these methods do not require a very large computer by today's standards — unless very large data bases are required (as in many sociological studies where the database may fill several reels of magnetic tape).

Access to good graphic facilities is also very important. Modern data analytic practices require the routine use of graphics to supplement the routine numerical output produced by traditional statistical methods. Fortunately this is now quite easy. It has the further benefit that publication-quality ilustrations can now be routinely made using standard software and a good quality plotter or laser printer.

Ecological and genetic modeling

Some of the earliest applications of computers to biology were in the area of the mathematical analysis or simulation of ecological and genetic models. Many models can only be solved numerically. The amount of computer requirements depends upon the types of models investigated. Usually memory requirements are fairly modest so that it is quite possible to run these programs on microcomputers (although with running times sometimes measured in weeks or months). These programs are often ideal applications to run as low priority background (batch) tasks on large computers.

Molecular modeling

The computational needs in this field are rather different from those applications described above. Present-day supercomputers are now powerful enough for researchers to investigate not only the static structure of a protein (as in the pictures produced by x-ray diffraction studies), but also, though the use of simulations, some of the dynamics of the internal motions of a molecule. This is important because, for example, it can help identify the optimal shape for an enzyme's binding site as a first step toward the design of enzyme inhibitors. For any reasonably sized protein the required calculations are enormous. For example, the protein bovine pancreatic trypsin inhibitor (BPTI) is a small molecule, composed of only 58 amino acids. To reach resolution of the dynamics in the nanosecond range (10^{-9} seconds) it has been estimated to take someting like 1,000 hours of computing time on a Vax 11/780. On a Cray-XMP the time would be reduced to only 10 hours. However to watch this small molecule go through its entire dynamic range would require about 10^{10} hours of computer time on the Cray (Patrusky, 1985). Larger molecules require much larger amounts of computation.

Clearly, there is a need for both much more poweful computing facilities and other mathematical approaches than straight-forward simulations. One answer is to develop special purpose computers for just these types of computations. An example of this is the FASTRUN computer that is hardwired (no software required) to perform those calculations required to determine the forces between each of the paired atoms in the molecule. An array processor will do the rest of the calculations. This computer will be able to perform these calculations 25 to 50 times as fast as the Cray I. While impressive, it only allows slightly larger problems to be studied so further developments are urgently needed.

Very high quality graphics devices and software are also often needed in order to display the results of these computations. Ransom and Matela (1986) contains a brief introduction to this field.

Office automation

Office automation needs are not limited to businesses. In a scientific laboratory one needs to keep track of grant budgets, purchases of supplies and equipment, prepare technical reports, grant proposals, scientific manuscripts for publication, organizing symposia, and keep up with correspondence with one's colleagues. Time spent on these activities can be made more productive through some degree of automation. Wordprocessing in particular has been found to be very helpful and most users now find it difficult to imagine doing without it. This is the most common application of personal computers (amounting to perhaps 80 % of their use). It is a particularly successful application for a microcomputer since this activity requires relatively little real computation but long hours of user interaction (editing and re-editing documents). Larger processing speeds are only needed when it is desired to produce output that is of typeset quality (desktop publishing using laser printers) rather than being just a simulation of typewriter quality output.

Desktop publishing, which gives a computer user the ability to produce high quality (like printed text) output, is an important new development. Some of the advantages are as follows:

— Manuscripts with mathematical formulas can now be prepared using normal mathematical symbols. This reduces the chances for errors in manuscripts since one no longer has to remember to write them in by hand on the final typescript or to use non conventional typewriter-style mathematical notation that has to be changed by the typesetter.

— Laser printers allow one to create good quality black & white graphics for publication. The normal 300 dots per inch resolution is sufficient for many plotting needs. Higher resolution devices are also available.

— It is now possible to prepare «printed» technical reports. Camera-ready output can be produced directly with little extra effort. This can be a great time-saver for the preparation of reports, proceedings of scientific meetings, etc.

Individual researchers and research groups act like small independent businesses and thus maintain their own data except for personnel information. This is in contrast to large corporations where many users must tie into a centralized corporate database. Thus one expects researchers to be able to take advantage of much of the technology being developed for small businesses.

At present, most businesses have not fully automated their offices so it is not surprising that most research labs have also not achieved the level of automation that is possible and that one reads about in the popular press. While most use wordprocessing, offices have only recently begun to network their PCs together and to tie them in to their central mainframe computer. It is interesting that, while in theory a combination of PCs and mainframe computers should provide all the power and functionality that is needed, small departmental minicomputers are also common. Groups like to maintain and manipulate their own data on their own computer where they can set their own priorities for when and how different jobs will be performed. Thus, the psychology of the user must also be taken into account in planning an effective computer installation. Networking is important but software is only now becoming available to fully integrate PCs and mainframe computers and their rather different software. Many users have taken advantage of the fact that by connecting PCs to a mainframe (as intelligent terminals) one has, in effect, a simple network in which users can send files to one another and to share a common printer. Dedicated wordprocessors seem to be on the way out. They are very effective machines but are more costly and less flexible than microcomputers programmed to act as wordprocessors.

Applications to scientific education

Early attempts at computer assited instruction (CAI) were not too successful due to its high cost and by the fact most teachers were not really ready to use computers. Another problem was the fact that the software (courseware) took a lot of work to develop if it is to represent more than just the automation of the reading of a textbook. There is clearly the potential for much more. Computers are now routinely being used to perform computations needed for a class. They have even greater potential as a laboratory tool to allow a student to manipulate a simulation of a complex process. For example, Peterson (1984) described a simulation of cardiac mechanics for use in a medical laboratory. Beginning students can get a «hands on» feel for the consequences of changing various parameters without having to sacrifice large numbers of laboratory animals (which is an important ethical consideration). Such programs must use animation style graphics with quick responses to student input in order to keep the student's attention. They must be as lively and captivating with its use of visual action and sounds as a video game. Curent CAI application are more likely to be successful than the early attempts. With the widespread use of microcomputers, students (and even many teachers) are more receptive to the idea than before.

Note that these activities are not directly related to the topic of teaching students to program. Even though more and more people are beginning to use computers one does not expect a large percentage of them to become programmers. With new software most of the nontechnical users will be able to use programs that allow them to get their work done without them having to learn to think like a programmer — which is a particular and disciplined skill. It has been thought by some that people should be taught to program simply because it would be good for them — it would help them learn to think. The study by Mayer *et al.* (1986) does not support that idea. Learning to program only seems to help closely related skills such as the ability to follow procedures, not general intellectual ability.

Issues in the delivery of computing services to researchers

It seems strange to have to justify the use of computers in 1986 but some research organizations still have very primitive computing facilities. When such institutions now begin to make plans for the purchase of their first computing system it is important that it is understood that one must do more than just buy a box labelled «computer». It makes a difference which vendor one buys from and the particular model one selects. This is not because one is «better» than another in some absolute sense. That is often the wrong question to ask. What one wants is the best computer for one's needs and for the local situation (availability of support in terms of both hardware and software). Conversion of software can be very expensive. So one must also take into account the fact that certain families of computers have much more software available for them (and one needs to take into account the fact that one needs not only scientific software but also various software packages that help one be more productive, *e. g.*, wordprocessors, compilers, statistics programs, graphics routines, application generators, etc.). Funds for software and for updates of software must be budgeted so that the systems can be kept current. Even if the software still works fine one is generally forced to obtain the new versions in order to stay current and compatible with similar systems at other institutions.

There are many definite advantages to cooperating individuals and institutions having compatible computers. Costs of software may be greatly reduced if locally developed software can be traded with researchers working on similar problems. In addition, psychological studies (Rushinek and Rushinek, 1986) have shown that users are happier with the computer they are using if many of their colleagues also use the same type of computer. However, this may be less important in the future as many are now coming to appreciate the advantages of portable software. But, for some applications, it will always be necessary to get the maximum performance from the computer to make the application feasible. In such cases programmers usually need to use system-dependent tricks and compatibility will remain an important consideration.

Institutional control of computing

This can be a rather controversial issue. *Should* computing be controlled? Microcomputers are very hard to control due to their low unit cost. But thousands of them are now found in many institutions and their total cost is considerable. An institution must be careful since if it buys the wrong one or buys a mixture of incompatible systems then a lot of money may be wasted. Microcomputers probably do not represent a cheaper means of providing computing power to large numbers of users. With proper planning, however, it can provide very good quality of computing services for many types of applications.

Training is very important. One can't just buy someone a computer and then expect them to make good use of it. This is especially true if the users don't think they need it (*i. e.*, they are not yet ready for it). New users must make an extra effort to acquire the necessary skills. Some clerical and technical support staffmay resist changes. There needs to a generally higher level of what has been called *computer literacy* among the users in order for them to be able to cope with the problems that normally arise in using computers. While the quality and friendliness of programs have increased dramatically in the past years, it still will be some time before an untrained user will be able to cope with all of the unforeseen problems that normally arise when one is dealing with computers. It takes some understanding in order to known how to cope with various types of problems.

Computer anxiety has often been discussed as a concern in automating procedures performed by nontechnical personnel in offices. Recent results (Howard and Smith, 1986) indicates that this is not as important as once thought. It was found to be a problem in perhaps only 3 % of people in managerial positions and it was correlated with mathematical anxiety but unrelated to age or sex. Institutions must take this into account. They cannot just give a computer to department and expect them to make efficient use of it without appropriate help.

Should one develop one's own software for special applications? One needs to be very careful in judging the costs and benefits of in-house development of software to perform such general tasks as accounting, wordprocessing, data bases, etc. A lot of good quality general software already exists for the more popular computing systems. Perhaps you have a person with special talent who could develop such software, but will you be able to maintain it after they leave the institution? Cost of software development are always much greater than what one expects. It is often a good strategy to let a well-established company develop and maintain software. The quality of commercial software has increased considerably in recent years. It is now much harder for amateurs to compete.

In many institutions researchers have been frustrated because they were unable to purchase the computing equipment they needed in their work even when funds were available. Computers were treated in a special class. At one time I was able to purchase a $20,000 programable desktop «calculator» with state funds but was prevented from purchasing a $1,000 terminal that could have been used with the campus computer. With the advent of microcomputers most of these bureaucratic controls have been greatly reduced. For many researchers the computer is now just another piece of equipment. Some of the problem may have been that institutions justified their large expensive central computing systems on the premise that it would handle all of their computing needs. If individuals and departments were allowed to purchase their own computers then what would happen to the expensive central facility? They could not imagine the insatiable appetite users have for computing power. More computing power allows researchers to progress to the stage where they need even more computing power to make even further advances.

Thus the question is not «how much computing power is enough» (with the implication that one could make a long term budget that researchers should work within), but how much scientific progress is enough. In some fields these two are completely linked. It is difficult to imagine how certain questions can be answered without the use of computers.

System performance

Does it matter that one's institution has an old overloaded mainframe computer? A number of studies have shown that response time has an important impact on productivity of computer users when the response time of the system rises above a threshold. When response times are near 1 second or larger various measures of productivity rapidly decrease. This has been found to be especially true for experienced computer users (Thadhani, 1984). Doherty and Kelisky (1979) attribute the phenomenon to the individual's attention span. A user does not have to think after each computer response but tends to use short sequences of operations. When response time becomes too long the user may have to rethink the sequence of actions to be performed. In a study of engineers it was found that it took twice as long for them to complete a standard task when the response time was 2.0 seconds than when it was 0.25 seconds (Anon). This is a clear advantage of PCs. For many simple tasks they give a very good response time since they are working on a single task.

These observations have been supported by psychological studies (Rushinek and Rushinek, 1986) of what makes users happy. The most important item was good response time. The computer is used by most as a means for getting various tasks done. Users sense that when response time gets very slow they are not being as productive as they could be. In the extreme it might even be faster to do various tasks by hand. They also found that users were unhappy with using old computing systems. This is most likely due to the fact that the new powerful software packages are written for the new computing systems that are selling well. Only later, if at all, will software be converted to still existing older and less powerful systems. Thus one is forced to keep up to date, just because of the way the market works. One may be forced to abandon perfectly well running hardware in order to have the software one needs. This is necessary because without the software one needs, hardware is just something that uses electricity to generate heat (and there are cheaper ways of doing that).

An institution must also be prepared for the fact that, while computing knowledge and software is a long-time investment, computer hardware is not. It may cost more in power consumption and hardware maintenance per year to run an old system that it would cost to buy a new computer. This is also reflected in the fact that old computers have little re-sale value. Some still working computers can't even be given away! Very frustrating — but that is the way it is.

Are there still economies of scale?

There has been a continuing debate over the years concerning the relative merits of having centralized large computer versus many smaller decentralized ones. The correct strategy is difficult to determine since there are many components to be considered.

- «Grosch's law» (Grosch, 1953) states that the costs of computer systems increase at a rate equivalent to the square root of their computing power. The implication of this is that, for a given amount of money, one will receive more computing power by buying one large computer rather than by several smaller ones. The original intuitive assertion has been verified by several empirical studies. Two recent studies (Ein-Dor, 1985, and Kang *et al.*, 1986) are of special interest since they include data on computers ranging from microcomputers to supercomputers. They found that the relationship still held approximately within each class of computer but not among classes. The effect was weakest for minicomputers (so that within that class there was relatively little advantage in getting the largest of the minicomputers). Among classes, the relationship was different with microcomputers, which resulted in much lower costs per MIPS than for a supercomputer. There seem to be some obvious errors in the data they used. One also needs to take into account of the fact that MIPS are very crude measures of computer performance (and biassed in favor of microcomputers). Their results must be continually updated as new computers make use of new technologies which could change the relationships between size and cost.
- The size of the smallest computer that one can use is determined, in part, by the size of the largest jobs that one must run. One would be unable to take advantage of the fact that several small minicomputers might be just as cost effective as one large minicomputer if one's largest programs would not run on the smaller machines. Not all problems can be split so that they can be run on several small machines at once. Some research computations are in fact only feasible on the largest of present-day computers.
- Most discussions about computers tend to focus on hardware but that is not always the limiting factor. One

must also take into consideration the problems of managing a computer facility (in terms of cost accounting, operations staff, system programmers, and user consultants). These problems become more complex as one distributes computing power to more than one computer. This is especially true if the computers are also decentralized (one may have to duplicate staff in remote locations or staff must travel).

— The cost of software must also be considered. Most vendors charge a licensing fee for each CPU used (with a small discount for multiple CPUs). Thus having 10 CPUs could increase annual software costs by a factor of 8 or 9 unless applications were segregated so that different types of applications were always performed on different machines. This can be a significant consideration.

— Users on separate computers will need to be able to communicate with each other. Accounting information will most likely have to be processed centrally. In addition it is much more efficient to have the systems programmers be able to distribute new versions of software, fix bugs, etc. from a central location. With a decentralized system it is usually essential that there be some sort of computer network. This will add to the complexity of the installation and the sophistication of the staff needed to manage the facility.

What is the best computer configuration for a given research institution? There is no easy answer. It depends upon what the computing needs are, the level of experience of the staff, projections for growth in computing needs, and how much money is available for this purpose.

CONCLUSIONS

While one would like to select the «best» hardware and software configuration for one's institution, that is probably impossible because the exact nature of research applications change and because new technologies rapidly make old solutions out of date. It helps to talk to people at similar institutions to see how they are trying to solve their computing problems. However, one should not plan on just copying their solutions since they may already be partly out of date. Thus, one can only try to select a reasonable configuration based on present technology (and budgets) and one's guesses about the future. Probably the most important consideration is the allowance in the configuration for significant future growth.

But the trends are clear. Over the next few years the rapid increase in the use of computers in scientific reesarch will continue. The use of computers will also continue to increase rapidly in many administrative applications that support scientific research. While the exact configurations of future computers are difficult to predict, the trend will be towards more powerful microcomputers which will be used both as stand-alone computers and as terminals into large centralized computers and networks that are linked with a variety of local, national, and international networks. The networks will be used both for electronic mail and for sending jobs to be processed on more powerful or more specialized computing systems. The result will be to make the computer and even more important tool for the support of scientific research.

ACKNOWLEDGMENTS

I would especially like to thank the Luso-American Foundation for their support for travel that made my participation at the conference possible. This is contribution number 614 from the Graduate Studies in Ecology and Evolution, State University of New York at Stony Brook.

RESUMO

Computadores em investigação científica

Neste trabalho é analisada a utilização de computadores em investigação científica, e são referidas importantes considerações que devem ser tidas em atenção quando uma instituição planeia a instalação de serviços de computadores para uso dos seus investigadores.

As características de sistemas batch, minicomputadores, microcomputadores e supercomputadores são comparadas para permitir uma visão geral das alternativas possíveis na escolha dos serviços de computação para os investigadores. São também apresentados exemplos de diferentes

tipos de aplicação de computadores para ilustrar como diferentes utilizadores podem ter as mais variadas necessidades de computação. Essas aplicações podem ir de um processador de texto, de cálculo estatístico, da automatização de equipamento laboratorial, de gráficos, até simulações e modelização em larga escala. Algumas destas últimas aplicações podem requerer uma grande quantidade de tempo de processamento, mesmo em modernos supercomputadores.

Finalmente, são discutidos vários pontos que devem ser tomados em consideração quando se planeiam serviços de computação — quer para um só utilizador ou em laboratório de investigação, quer para um instituto ou uma universidade. São feitas considerações sobre a importância da rapidez de tempo de resposta do computador (para maximizar a eficiência do utilizador), economias de escala (grande computador centralizado Vs microcomputadores distribuídos por vários locais), disponibilidade de bons programas (software) e disponibilidade de apoio técnico.

RÉSUMÉ

Les ordinateurs dans la recherche scientifique

Ce travail concerne l'emploi des ordinateurs dans la recherche scientifique et réfère d'importantes considérations à tenir compte quand une institution projette l'instalation de services d'ordinateurs pour l'usage de ses chercheurs.

Les caractéristiques des sistèmes batch, des micro-ordinateurs et des super-ordinateurs y sont comparées afin de permettre une vision générale des alternatives possibles concernant le choix des services de calculs avec ordinateur pour les chercheurs. Des exemples d'applications diverses d'ordinateurs sont égalament présentées dans ce travail pour illustrer comment des utilisateurs différents peuvent prétendre des réponses les plus diversifiées possible. Ces applications vont dès le traitement de texte, le calcul statistique, l'automatisation de l'equipement de laboratoire, tracer des graphiques jusqu'à des simulations et à la création de modèles en grande échelle. Quelques unes de ces dernières applications peuvent solliciter beaucoup de temps, même en utilisant des super-ordinateurs modernes.

Finalement divers points sont discutés au long de cet exposé. Ils doivent être regardés attentivement quand des services d'utilisation d'ordinateurs sont projetés, soit pour l'usage d'un seul utilisateur, soit dans un laboratoire de recherche, soit pour un institut ou une université. A noter les considérations sur l'importance du temps de réponse de l'ordinateur pour donner le maximum d'efficience à l'utilisateur, les économies d'échelle (grand ordinateur centralisé versus micro-ordinateurs distribués dans différents endroits), sur la disponibilité de bons programmes (software) et d'assistance technique.

REFERENCES

Anon.

The economic value of rapid response time. GE20-07520-0:IBM Corp.

Cowart, R.

1985 As the market grows, standards and long-term solutions come into view. PC Week, July 1985: 55-62.

Ein-Dor, P.

1985 Grosch's law re-revisited: CPU power and the cost of computation. Comm. ACM, 28: 142-151.

Grosch, H. A.

1953 High speed arithmetic: the digital computer as a research tool. J. Opt. Soc. Am, 43: 306-310.

Howard, G. S. and R. D. Smith

1986 Computer anxiety in management: myth or reality? Comm. ACM, 29: 611-615.

Kang, Y. M.; R. B. Miller and R. A. Pick

1986 Comments on «Grosch's law re-revisited: CPU power and the cost of computation». Comm. ACM, 29: 779-781.

Mayer, R. E.; J. L. Dyck and W. Vilberg

1986 Learning to program and learning to think: what is the connection? Comm. ACM, 29: 605-610.

Mayfield, A.

1986 Truths, half-truths, and statistics. Datamation, 32 (16): 85-86.

Moad, J.

1986 Too much ain't enough. Datamation, 32 (16): 26-28.

Patrusky, B.

1985 Biology's computational future. Mosaic, 16 (4): 2-13.

Peterson, N.

1985 Designing a simulated laboratory. Byte, June 1984: 287-296.

Ransom, R. and R. J. Matela

1986 Computer graphics in biology. Croom Helm: London. 211 pp.

Rushinek, A. and S. F. Rushinek

1986 What makes users happy? Comm. ACM, 29: 594-598.

Thadhani, A. J.

1984 Factors affecting programmer productivity during application development. IBM Systems J., 23: 19-35.