COMPUTING AT THE UNIVERSITIES OF DURHAM AND NEWCASTLE UPON TYNE IN THE 1980'S

Report of an Ad Hoc Working Party

chaired by Professor H. Whitfield

20 October, 1980

TABLE OF CONTENTS

SECTION CONTENTS

COMPUTING AT THE UNIVERSITIES OF DURHAM AND NEWCASTLE UPON TYNE IN THE 1980'S

Report of an Ad Hoc Working Party

chaired by Professor H. Whitfield

20 October, 1980

1. INTRODUCTION

This report discusses some major policy issues concerning the future provision of computing at the two universities, and is addressed to the Newcastle Computing Laboratory Committee and the Durham Computer Unit Committee. It has been prepared by an informal working group whose membership has been drawn from the staff of the Laboratory and the Unit. The need for such a report arises from two factors, namely

(i) the likely constraints on the human and financial resources that will be available to provide for an ever-growing extent and variety of computer use, and

(ii) the continuing rapid pace of technological developments in computing.

In common with a number of other major university computing services, on both sides of the Atlantic, we believe that these factors could and indeed should cause the early 1980's to be as much of an era of change as was the late 1960's, when the first large scale time-sharing systems, such as MTS, were introduced. Accordingly we believe that there are a number of problems and major policy decisions, which need to be actively considered by the two Committees, and indeed by the existing and potential future computer user community at large.

In section 2 of this report we provide a brief analysis of the expanding variety of uses to which computers are being put. The section following describes relevant major developments in computer hardware and telecommunications. Section 4 is an account of the present state of computing in the two universities, as seen from within NUMAC. Against the background provided by these three sections, section 5 attempts to take into account the realities of the resource constraints we and the users will face in the coming years, and to propose some of the major components of a development plan for computing at Durham and Newcastle. Our various specific recommendations are summarised in the final section of the document.

2. COMPUTER USAGE

There is a two-dimensional classification of computer users. One is by computing requirement and the other by subject area. Computing requirements of course vary widely. Many users will have no specialised demands, and will be happy to use a computer (with adequate storage and processing) as a black box. Other users will require access to special types of hardware (for example, vector processors for solving large scientific problems), while a third category consists of what might be termed "system stretchers" who have problems which can be extended to saturate whatever computers are available, however large.

A number of (computing) subject areas are:

(i) **General (traditional) computing.** Most present-day computing is done by user-written programs, aided by editors, program writing and debugging aids (often referred to as software tools), and a variety of programming languages. Studies elsewhere suggest that the computing power required by these existing users is growing at 15-25% per annum; it would seem unlikely that Newcastle and Durham are any different. Nor will improved software tools reduce this rate of increase - people are simply likely to write more programs.

(ii) **Software package usage.** Many users, of course, do not write their own programs but make use of large program packages (of which the statistical package SPSS is perhaps the best known). Package use amounts to about 30% of current computer usage and this proportion seems certain to grow as more computationally unsophisticated users start to use a computer.

It is worth pointing out that the package/user program split is not between light and heavy machine users; indeed NUMAC's biggest system stretchers are the quantum chemists who make use of the ATMOL package. It is also worth remarking that almost all present computer use manipulates numbers; as the sections below will show, this will change as computers are used more to manipulate text.

(iii) **Word and text processing.** This subject, which is beginning to expand rapidly, needs consideration at three different levels. At its simplest, many cheap machines are able to support word processing, which is the production of fairly simple documents, such as letters and memoranda. The second level is the production of more complex material, such as lengthy reports and books, which may require tables, figures, equations, and several different fonts, together with indexing and cross-references; for this more sophisticated facilities and substantial processing are required. The third level concerns literary and linguistic research in which the text itself is regarded as data, for example in stylistic analysis. The more complex of these applications can consume fairly large quantities of computer time (for example, book production can use one second of IBM 370/168 processor time per page) as well as significant file space.

(iv) **Data bases.** A great deal of information is now available in a structured form suitable for computer analysis. Several data bases and their manipulative software are mounted on the NUMAC machine, while others are accessed remotely. (Indeed, many of the computer networks now becoming available are marketed as data-base access services, as of course is Prestel.) The applicability of the data-base approach to information management seems limited principally by the availability of storage space. (In the commercial world, a data-base holding 100 million characters of information would be considered small - this is already nearly half the size of the entire text of the Encyclopaedia Britannica.)

(v) **Teaching.** In the future many more University departments might well wish their students to have an awareness of the potentialities of the computer in their own fields, and to have some training in computer use. It also seems likely that the computer will become a classroom tool in many disciplines, just like the oscilloscope. The constraining factors will be the numbers of terminals and small computers which can be provided, the availability of suitable software, and the rapidity of acceptance of the new possibilities by teaching staff.

(vi) **Graphics.** For many applications pictorial (and preferably dynamic) output is more satisfactory than lists of numbers. Much engineering design in the commercial world is now done using graphical displays and this trend is certain to advance further. It should be noted that interactive graphics is remarkably expensive in computing terms (both of processing and data transmission).

(vii) **Electronic mail.** An incidental benefit of a widespread penetration of computer terminals into University offices would be the possibility of evaluating the usefulness of an electronic mail system. Given a suitable set of facilities it is possible to use electronic mail to eliminate much conventional internal and even some external mail. The effectiveness of such systems stems not just from their efficiency at transmitting information to specific individuals or groups, but also from the fact that previous correspondence can easily be retained in a computerised data base, and be much more readily accessed than if it were in filing cabinets.

(viii) **Administration.** Although there are many advantages in keeping administrative computing separate from academic computing, the question of whether the separation should be total is worth addressing. Two examples of situations in which links would be useful are in interactions with the University Libraries and with UCCA for student admissions. Additionally, computers are likely to be used increasingly as an administrative tool - at personal, departmental and university levels.

(ix) **Data collection.** In the Chemistry Department of the University of Newcastle upon Tyne a cheap microcomputer is used to oversee an expensive experiment continuously, collecting data and passing batches of data to be analysed to MTS whenever MTS is prepared to listen. Thus some expensive equipment is used far more intensively than could be the case with manual control. Such automatic collection of data will certainly spread.

(x) **Support for microprocessors.** The storage, input/output devices and software facilities of many microprocessors are so limited that it is often necessary to use a larger computer to assist the task of program preparation, with tools such as cross-compilers, cross-assemblers, simulators, etc.

It is generally accepted that user demands in universities have for some time caused the underlying computing growth rate to be about 25% per annum. There is every reason to believe this will continue, and be accompanied by at least equal continuing growth in storage requirements, though with the increasing proliferation of computing equipment throughout the universities it is becoming more difficult to estimate the total amount of computing resources that are being used. Some of the more recent applications, such as word processing, data bases, and graphics are likely to provide particularly heavy demands on future computing resources. Such growth is of course facilitated by, and indeed partially caused by, the decreasing cost of computing, which is allowing an ever increasing range of uses to be judged as being fully cost-effective.

One final and important point that has to be made is that an ever-growing proportion of users will not want to be programmers, let alone system programmers. They, very reasonably, will want the computer to be an easy-to-use and stable tool to help them with their real work, rather than provide them with irrelevant challenges and tasks. Similarly, existing users usually have a large investment embodied in the programs and data files that they are using, and thus place great reliance on the stability of their computing environment and tools.

3. TECHNOLGICAL DEVELOPMENTS

The three principal areas affected by the current rapid improvements in hardware technology are processing power, long-term storage facilities, and data communications. Performance and cost improvements in processors and main storage derive from developments in semiconductor technology. The basic and amazing fact is that, for the last twenty years, the number of logical elements that can be placed on a single chip (i.e. its "integration level") has increased by a factor of 10 every five years. Moreover it seems certain that this increase will continue for a number of years to come.

Such advances in integration level are permitting two somewhat distinct paths to be followed by manufacturers - the production of more powerful expensive computers and of cheaper small computers. Machines with fifty times the number crunching capacity of the 370/168 are already under development. It is confidently predicted that machines of the power and logical complexity of a 370/168 could be available on a single chip by the mid-1980's. In fact experimental machines of this power intended for use as personal computers by individuals in their offices, and occupying little more space than a standard three-drawer filing cabinet, are already operational at the Xerox Palo Alto Research Center.

At the other end of the scale, microcomputers capable of performing many of the simpler tasks required are already cheap enough to be within range of a department's equipment grant. The earliest microcomputers had quite archaic and inefficient architectures, but their manufacturers (and thus their users) are reliving the history of the development of computing at an accelerated rate, and more reasonable machines permitting the use of more up-to-date programming practices are now becoming available.

Developments in backing storage are less easy to predict. Cost/performance improvements in magnetic disk storage are continuing, but there is a growing gap between the performance of such storage and that of semiconductor storage. Some new technologies such as bubble memories may plug this gap and provide very fast backing stores of modest size, suitable for use in place of a paging drum, for example. However the improvements that are being made in disk storage densities will make large (and therefore usually shared) file servers, i.e. backing and archival stores, economic for some time. (The economies of shared file storage resources become even more obvious when one takes into account the needs users have to share actual data, and the costs of the largely manual activities involved in taking precautions to prevent occasional loss or corruption of stored information.) Small scale backing stores suitable for microcomputers are in a particular state of flux - tape cassettes are already being superseded by floppy disks, which may in turn be ousted by bubble memories and/or small versions of conventional, magnetic disks based on so-called "Winchester" technology. One other development, whose impact is very difficult to predict, is the advent of the video disk, which will permit extremely cheap storage and fast access to huge amounts of data, but does not allow data, once written, to be subsequently modified.

The other important area of hardware development concerns communications. Existing communications links, such as those often used for time-sharing terminals enable communications at a rate of at most a few thousand characters per second. However the costs of high speed communication (i.e. at a rate of millions of characters per second) such as would be appropriate for connecting computers to large backing stores and directly to other computers are falling fast, aided by such developments as optical fibre communications. The Post Office's monopoly marketing position may cause the price at which they provide high speed communications to remain high, but this need not apply within a single campus. Such high speed networks already exist in a number of places such as the Xerox Palo Alto Research Center and (on a smaller scale) within the Cambridge Computer Laboratory. High speed networks covering an entire campus are in fact being actively planned by many other universities, including Cambridge, Edinburgh, Kent, MIT, Stanford and Waterloo.

It is these communications developments that are central to the view that university computing (and for that matter much large scale industrial and commercial computing) will change as dramatically in the next few years as it did when time-sharing was introduced. Thus in some cases such networking plans are combined with plans to install large numbers of quite powerful personal computers. Xerox has had internal networks of some hundreds of powerful 'Alto' personal computers, each fitted with a high quality graphical display, operational for quite a few years, but similar machines are only now about to come on the market, priced at about $20,000. It will therefore be some time before such computers generally replace ordinary terminals connected to shared computing facilities, whether these be large central main-frames, or so-called 'multi-user minis'. However, the trend is there to see, and the quality of service and of user interface that can be provided by networks of powerful display-oriented personal computers like the Alto and its successors, backed up with appropriate centralised services for mass storage, high quality printing of text and graphics, and links to other user communities, has to be experienced to be believed.

There are of course many other hardware developments which are likely to effect future computing provisions. Briefly, some of these are:

(i) High-quality printing using photo-typesetting, as well as high-performance line printing, seems likely to be overtaken by the use of laser-based xerographic printers, particularly if small cheap printers of this type, suitable for both text and graphics, can be developed.

(ii) Raster-scan (i.e. television-like, as opposed to vector-drawing) colour graphics, suitable for high resolution drawings as well as text, are likely to become much more prevalent.

(iii) Voice output is already available, but voice input is still a subject of (mainly software, indeed "artificial intelligence") research.

(iv) There are definite moves towards integrating voice, data and facsimile transmission in very high speed transmission and switching systems.

(v) There is much research on novel forms of computer organisation, in some cases aimed at special number crunching applications (e.g. wind tunnel simulators), in others intended for new approaches to general purpose computing.

There have also been extensive software developments in recent years, and a huge growth of application packages. These developments concern the provision of tools for developing new programs and applications. Developments range from improved high-level languages (of which Pascal is the most widely used) to very sophisticated suites of display-oriented interactive tools intended for assisting all phases of program development. However, such new tools are difficult to fit into existing computing services, and to combine with the desires of existing users to preserve the value of their past software investment (let alone their wish to go on using the particular language they first learnt perhaps many years ago). Thus the most successful examples of software development tools are often in new environments, using new hardware, and little known to users elsewhere.

One successful compromise that has been achieved at Bell Telephone Laboratories is a development system based on the Unix operating system. (Unix itself is in widespread use in universities on both sides of the Atlantic, including Newcastle and Durham, and is becoming available not just on PDP-11/VAX computers, but also on a variety of processors, ranging from main-frame computers to microcomputers.) The development system is entitled the Programmers Work Bench, and runs at Bell Labs on numerous PDP-11 and VAX computers, providing users with a high quality interactive environment for developing and documenting programs which are eventually intended to be run on a variety of large-scale computers, including the IBM/370. The submission of jobs to such computers is handled automatically by the Programmers Work Bench, which also provides extensive facilities for the examination and analysis of the resulting output.

Despite such developments, software costs are continuing to rise, particularly relative to hardware costs. This is hardly surprising, since the main function of software is to fit general purpose mass-manufactured hardware to some more specialised (and often unique) environment. Thus in contrast to hardware, software development costs cannot usually be amortised over the subsequent production runs needed to produce large numbers of identical copies. In a university environment, even though, except for purchased software, software costs are not usually assessed in monetary terms and therefore not so visible, they are nevertheless equally real. They are likely to come to dominate many decisions concerning the type, scale and use made of computer facilities in universities.

4. THE CURRENT STATE OF COMPUTING AT NEWCASTLE AND DURHAM

The pioneering decision by the two universities in the late 1960's to combine their resources and to install one of the first large-scale time-sharing systems in Europe has proved to be very successful. The user population has grown to about 3000, to whom a very large variety of compilers, packages and facilities of all kinds have been made available. In recent years much effort has also gone into the development of the NUNET network, which currently handles 90% of the terminal traffic to MTS. Thus twelve years after we settled on MTS as our primary system, it is still serving us well, though it is beginning to show its age and still has some deficiencies. However as the system has grown in complexity, so the maintenance task has become larger, as has the task of providing an adequate advisory service and of coping with the materials handling and operational problems associated with a growing number and diversity of users. Indeed at present these aspects of the NUMAC service are under particular strain, especially at Newcastle - a situation which will have to be remedied before we can make much progress on more long term plans.

Many other computer-based facilities have grown up alongside NUMAC, ranging from those used for library and general university administration, to those provided in particular departments, or for use by particular groups of users (e.g. those funded by the S.R.C.). A growing number of small machines are used for word processing and of course various external services are used, via several different communications paths.

Most of these developments are well-justified, and were only carried out after appropriate consultations over the choice of hardware and software to be used. Nevertheless they have already given rise to a number of compatibility problems that are preventing fully effective use of the various facilities and are causing significant amounts of wasted effort.

Finally, we turn briefly to the resources available now and in the future to NUMAC. Our staff are now fully extended running a large and complex service, yet there seems little likelihood of increases in staffing to cope with predicted growth in the demands made on us. Under normal Computer Board policy we can hope for only modest funding for minor upgrades until a major upgrade of our hardware resources in about 1985, though future Board policy, and indeed the future of the Board itself, is regretably in doubt.

5. FUTURE POLICY

We believe that there will for some years to come be a continued demand for MTS on centrally provided hardware facilities and for the large range of software packages and languages it provides. However as demands grow, it will be neither desirable, nor perhaps even feasible, to continue to base the provision of computing power on just a single computer. Rather, much new computer usage will and should be based on the use of other computers, some located centrally but with many in various locations around the two universities, and often being used solely or primarily by particular departments (e.g. for laboratory automation or computer-aided teaching) or even individuals (e.g. for text processing or interactive graphics). Thus we expect mini and microcomputers to proliferate, although we believe that we will continue to have a lot of users below the sophistication threshold of wanting or even being able to manage their own computer system.

We believe that virtually all of the two universities' computer systems, together with their multifarious terminals and peripheral devices, should in general be linked together via a high speed communications network, and that the development of such a network and its associated file servers and gateways to external services must be pursued vigorously. The network will need to have alternative routes providing the means of masking most hardware faults, and ideally should be operational twenty-four hours a day. The file servers should provide reliable and secure storage for large volumes of data, and if only because of communications costs, will probably be needed in both Durham and Newcastle. They should be general in nature, rather than associated with particular processing resources.

Like a number of other organisations where similar plans are being formulated, we believe that the availability of a network and file servers with well-chosen and well-established interfaces and protocols will

(i) provide a basis for much less disruptive evolution of the computing facilities (both central and departmental) as processing and storage resources are added to cope with increased demand and as obsolete facilities are withdrawn,

(ii) enable the avoidance of much wasteful duplication of hardware and of software and administrative effort (by both computing service personnel and users),

(iii) facilitate collaboration between users,

(iv) provide the basis for the the provision of a coherent plan for word processing, and

(v) aid the cost-effective use of personal computers, which we expect to become prevalent in due course.

The provision of high speed communication facilities and file servers should automatically encourage the acquisition and use of those computers, terminals, and software systems that already have provisions enabling them to adhere to the communication interfaces and protocols that are needed for use on the network. Nevertheless, much effort to support and promote various standards will be necessary. Failure to standardise on a small number of interfaces between computer and communication systems will ultimately result in high monetary and man-power costs to both universities. Indeed unnecessary incompatibilities between the various hardware and software systems used by different users will often prove to be a luxury that the universities can ill afford.

Ultimately, in individual cases it may be appropriate to make an explicit decision that a particular standard need not be adhered to - in some cases such decisions may best be the responsibility of individual users or departments, but in other cases more extensive prior consultation should be required as should certainly be the case prior to introducing a new or changed standard. (It would be naive to assume that standards never have to be updated - in any case an occasionally changed standard is much better than a complete absence of standards.)

Standards will be needed for terminal communication, computer-to-computer data transfer, the representation of text to be formatted by word processing systems, and for using various centrally provided compilers, packages, data base systems, etc. In general it is easier to achieve such standards by minimising the number of different processors and software systems involved. (Edinburgh, for example, have recently decided that they can afford to provide microprocessor support facilities for only three specific makes of microprocessor, and that they should strongly encourage the use of just one software system on these micros, namely UCSD Pascal). Similarly, by far the easiest way to achieve the savings and advantages that word processing can offer the universities will be to minimise the number of different makes of word processing systems that are installed. For different but equally valid reasons, namely user convenience, and the costs of training and documentation, it will also be very advantageous to settle on certain standard user interfaces (e.g. for file editing, job control, and document preparation), even when these are provided on several different processors.

These and many other standardisation issues will require prompt but careful consideration and consultation. Wherever possible we should adopt existing international or national standards, such as those for basic network communications being promoted by the Joint Network Team and the British Post Office. Where agreed standards do not exist we should attempt to cooperate with other universities in developing the required standards, and, where appropriate, in sharing the software development efforts involved.

The benefits of continued close cooperation between Durham and Newcastle in the form of NUMAC are therefore obvious. (A single network is feasible because relatively high speed Post Office-provided communication links between the parts of the network at each campus would be economically viable, whereas multiple direct connections between individual systems at the two universities would not.) However, it is clear that the role of NUMAC must change somewhat to include much more emphasis on the development of high speed communication facilities and of those computing resources that are best provided centrally, such as file servers and gateways to other computing services, as well as much work on standards definitions.

It has to be assumed that the growth of diversity of users and usage cannot possibly be matched by a similar growth in the provisions by NUMAC of advice, services, assistance with programming, documentation, etc. Thus NUMAC's system programming activities must concentrate on the provision of easy to use software tools of wide utility. For some years, future developments will have to involve the use of, and certainly systems programming for, as few different computing systems as possible. Indeed it has been observed, not just at NUMAC, that it is just not practicable for a single computing service to provide fully adequate support for more than two separate major operating systems. On these grounds, and because of existing user populations, it seems clear that for a number of years to come NUMAC should base its service on 370-compatible processors (running MTS) and PDP-11/VAX compatible processors (running Unix, where appropriate). Moreover in many cases particular software packages should be mounted on just a single system, and accessed via the network since it will not be practicable to provide the necessary conversion and installation effort needed to provide multiple versions of a given package. Indeed NUMAC support for each individual package will from time to time have to be assessed carefully to determine whether it is still a defensible use of scarce staff resources.

It will not in general be possible to provide much specific assistance or even advice to individuals or very small groups of users wanting very specialised facilities. Rather we must make it easier for such users to become self-sufficient, whether using private or centrally provided computing resources, and whether involved in programming or just in the use of pre-programmed packages. Equally we will have to avoid any significant increase in the amount of operator intervention and materials handling (cards, tapes, disks, line printer paper, etc.) that has to be done by NUMAC personnel - such tasks as continue to be necessary will largely have to be carried out by the users, who will need to be provided with convenient local means of input, and of obtaining printed and graphical output. With comparatively few users needing to visit a central site, easy to use 'help' facilities and 'message' systems will also be needed.

If the two universities can agree policies along the above lines and if NUMAC can master, and have the resources and backing necessary for, the resulting changes to its role, it should be possible for Durham and Newcastle to achieve the advantages offered by new technological developments, and for users to see the period as one of evolution and improvement rather than of traumatic change and growing chaos.

6. SUMMARY RECOMMENDATIONS

The following recommendations are put forward, not as firm proposals, but as a means of initiating discussion on long-term planning for all forms of computing at the Universities of Durham and Newcastle upon Tyne over the next five to ten years. The list does not include recommendations as to specific levels of funding as the proposed developments could be pursued at various levels dependent upon price developments and availability of resources. They presume prior resolution of short-term issues concerning the NUMAC service, such as operational problems, the provision of screen editing, improved graphics facilities, and extensions of NUNET, for example to allow computer-to-computer data transfer.

(i) The two universities should continue to collaborate through NUMAC, to minimise waste of development effort and to maximise the usefulness of our scarcest resource - skilled staff.

(ii) Even when major hardware upgrades are made the NUMAC service as seen by the users should evolve rather than change abruptly, in particular avoiding the withdrawal of, or unnecessary changes to, MTS-based facilities still in widespread use.

(iii) There should be a high speed dedicated data network serving buildings on both campuses, as well as provision for high speed connections from any terminal to the computing facilities, via the network where appropriate.

(iv) There should be high-capacity central file storage facilities in Durham and in Newcastle, based on computer systems independent of the main processing facilities, in particular MTS.

(v) Departmental devolution of much of the processing power, for both text and for numbers, should then be encouraged subject to conformity with agreed standards.

(vi) NUMAC's role should be broadened to include the task of developing and controlling the high speed network, and coordinating the standardisation of interfaces, languages, data structures and communications protocols, in collaboration with other institutions where appropriate.

(vii) Because of the impracticability of expanding the user advisory service to keep pace with the growing amount and variety of usage, more effort should be put into simplifying user interfaces, and providing coherent and comprehensive on-line 'help' and 'message' facilities.

(viii) Access to external computing or on-line information services should in general be via gateways provided on the high speed network with NUMAC being responsible for the connection facilities, and for advice on the external facilities and their economical utilisation. It should also monitor and control any resulting external expenditure.

(ix) Because of uncertainties as to the pace of some of these developments attention should be paid to the problems of keeping as many options as possible open.