The operation of large computer systems

By Florence A. Rigg†

This paper describes the changes in methods of operation of large computer systems. The experiences cover a period of seven years during which four major computers were installed in turn.

The purpose of this paper is to show how our ideas at AWRE, Aldermaston, on the organization of large computer systems, have developed with our experience of different computers.

Early computers

The first AWRE computer was a Ferranti Mk I*, installed in 1955. This was followed in 1956 by an English Electric Deuce. Both were programmer operated, the programmer himself sitting at the computer and monitoring his own work through sometimes lengthy testing and production sessions.

The IBM 704 followed by the 709

In February 1957 an IBM 704 arrived. This machine had an 8K core store, 16K magnetic drums and 10 magnetic-tape units, with card reader, card punch and printer on-line. After about six months, the 8K core store was increased to 32K and the magnetic drums were withdrawn. Only one input–output unit could be used at a time; contents of storage registers and arithmetic units could be displayed in lights on the operator's console, and corrections could be inserted manually using switches. The first few months were spent in program testing and development, and there was sufficient time available for each individual programmer to take his work to the computer and test it by the old method of pressing buttons and reading lights. However, as the programs reached production stage, it soon became necessary to run the computer for every available hour. It would clearly have been inefficient to continue to use programmers as operators, and so professional operators were introduced.

Initially it was difficult to assess the calibre of staff required for computer operation, and so experiments were made in using members of various grades already existing in the Atomic Energy Authority. From this it became apparent that suitable individuals may be found in several different grades, but that the requirements of the grades do not necessarily ensure that the individual will suit the job.

The introduction of operating staff was not immediately appreciated by the programmers, even though it relieved them of the necessity of working during the evenings and weekends. They complained that they had to spend time in writing instructions to the operators, and that it took much longer to get their work through the computer. As a compromise, programmers were allowed access to the computer during normal working hours, the remainder of the time being under operator control. The day time was scheduled according to the current requirements of the research program for about a week in advance, so that each group of programmers knew precisely at what time and for how long the computer would be available to them. A small amount of time, of the order of three half-hour periods per day, was left free for any programmer to queue to run his job provided it was of not more than 2 minutes duration.

With a rented computer like the 704, it was uneconomic to use the main computer just for loading cards and for printing results. So peripheral equipment was obtained to perform card-to-magnetic-tape and magnetic-tape-to-printer operations. Then in July 1959, the 704 was replaced by a 709 which, for the first time, gave us the possibility of simultaneous input–output operations. The 709, in fact, had four input–output channels, channel A which carried the card reader, card punch, printer and 4 magnetic-tape units, channel B which had 4 magnetic-tape units, and channels C and D which had 2 magnetic-tape units each.

Replacement by the IBM 7090

Fifteen months later, a 7090 was installed to take the place of the 709. This increased the speed of computation by a factor of five and led to a much faster flow of work through the installation. A job which had previously taken 10 minutes to execute now only took 2 minutes, barely enough time for the operators to read the instructions for the next job and load the tapes in readiness. With the two earlier computers, tape handling had presented no particular problem, but during the peak periods on the 7090, tapes had to be loaded and unloaded at the rate of something like one a minute, and the operators had to work at their top speed in order not to allow the computer to be idle. Also, the peripheral equipment was no longer able to keep up with the demands of the computer and so, early the next year, it was superseded by a 1401. This gave us a greater speed of loading cards onto tape, and more important, raised the speed of printing of results

† AWRE, Aldermaston, Berks.

from 150 lines per minute to 600 lines per minute. Before the installation of the '90, it was decided to recruit and train operations supervisors, either mathematical graduates or people with comprehensive punched-card and computing experience, who would work on a shift basis and, between them, be responsible for the smooth and efficient running of the installation. By the end of 1960 these people were in post, and operation of the computer was almost entirely in the hands of the professional operating staff, though programmers were still allowed to watch their jobs being executed. About the same time, the first version of the FORTRAN monitor became available. This system control program automatically called into the computer the successive jobs of varying type which had previously been written on magnetic tape. It proved convenient for both programmers and operators, and opened the way towards the use of even more comprehensive system control programs.

Use of computers at a distance

During all this time, other stations of the UKAEA had been making use of the computers at AWRE. Much thought was given to the problems of using computers at a distance, and to the best methods of transmitting data. In an effort to counter the difficulties of the off-site programmers, fixed periods of computer usage which would produce the most effective turn-round were agreed and guaranteed. In practice, the most acceptable periods proved to be one at about midday and another during the night for each station.

Harwell and Winfrith each installed a punched-card transceiver system, together with a Mufax facsimile transceiver and a separate speech line for the transmission of data and instructions and for receiving a limited amount of output. The bulk of the printed output was sent by courier cars which ran between the Establishments each evening. Under this system the majority of the Harwell and Winfrith programs were held, in card form, at Aldermaston, and it was found necessary to employ a man full-time to progress work from the data links through the computer and back by the appropriate link.

It was not long, however, before Harwell decided that, over the comparatively short distance of 25 miles, a courier car gave just as quick a turn-round and was considerably easier and cheaper to use than the data link. They therefore organized a midday courier car in addition to the evening one, and dispensed with their transceiver equipment. Their programmers could then work very much as on-site programmers, preparing their own runs and depositing them by specified times at a central collecting place. They could then expect to receive the results of short runs done during the day within three to four hours, and the results of a night run when they arrived at work on the following morning.

Winfrith, at a distance of 75 miles, were unable to get this same day-time turn-round using a courier car. They therefore continued to use their transceiver systems and the man at Aldermaston to obtain a comparable service.

Off-site programmers do tend to feel that the same facilities are not available to them as to those programmers who work in the building next to the computer. It is, we believe, psychologically inevitable that this feeling should exist. Experience, however, seems to prove that they do in fact achieve almost as fast a turn-round of work, and with careful organization at each end of the link they need suffer no inconvenience due to distance from the computer.

The "Stretch" installation

The installation of Stretch in May 1962 brought major changes in organization and operation. Since 1955, the programmers had used a series of different machine languages, so it was decided that the recoding for Stretch should be the last recoding for some long time. A FORTRAN-type language standard was therefore adopted, mainly because it was already in widest use on large computers. In order to ensure that a FORTRAN-type compiler would be available when Stretch arrived, a joint AWRE and IBM team, led by Mr. A. E. Glennie, set to work to write what became known as S1. A second innovation with Stretch was the decision to operate entirely under the control of a supervisory program. This Master Control Program (MCP) was written by IBM in America.

The problems of the weeks following hand-over may be imagined: a large and intricate hardware system; MCP on "field test" from America; S1 subjected for the first time to the rigours of compiling major programs; and the programs themselves just re-written in FORTRAN. Expert advice was available in all fields, but no one person could be an expert in everything. As each problem came up for investigation, it was first necessary to decide where the most likely source of error lay, and in which area to make the initial attack. The investigation then proceeded through a series of eliminating rounds until eventually the true cause of the trouble was found and corrected. This called for the co-operation and forbearance of IBM systems staff and engineers, the AWRE systems staff, and the AWRE problem programmers. This co-operation was given ungrudgingly by all concerned.

Such a complex system of hardware had to be controlled entirely by professional operating staff. It had been strongly emphasized to us that whenever Stretch was in operation we should need at least one person on duty who was conversant with MCP and knew some of the idiosyncrasies of both the hardware and the compiler. To try to meet this need, each Supervisor was given a course in Stretch programming language and in MCP. He was already familiar with FORTRAN and had merely to learn the minor differences between it and S1.

Our Stretch system includes 96K words of core store, 2 million words of disk storage, 16 magnetic-tape units, one card reader, one card-punch, one line printer and a console typewriter. This is backed by an 8K, 4 tape-unit 1401 off-line. To complete the services offered, a General Dynamics S.C. 4020 Microfilm Recorder has been added to the installation, giving the possibility of both graphical and printed output from magnetic tape onto 35 mm microfilm. The S.C. 4020 records data at speeds in excess of 17,000 characters per second. When used exclusively as a printer it will print 7,000 lines per minute. An average annotated graph can be recorded in fractions of a second. A further option allows the simultaneous recording of data onto photo-sensitive paper, though this slows the recording speed from about $\frac{1}{2}$ sec per frame of microfilm to 3 secs per frame of hard copy.

We operate 2 shifts per day 7 days a week. To do this, three teams of shift staff are needed, each with its Supervisor, Deputy Supervisor and a total of 6 operators of varying degrees of skill and experience. Day release for further training, holidays and sickness cut the effective number. In addition 9 machine-grade staff are employed, during the day-time only, on reception work, distribution of results, magnetic-tape library, punching and processing of cards, etc.

The present organization of the installation is largely governed by the facilities offered by MCP, and also by its limitations. MCP automatically loads programs sequentially, requiring from each job some identification and instructions for use in scheduling the queue of work. It scans the incoming queue for magnetic-tape requirements, and issues instructions to the operators to set up units in advance of the execution of the job which requires them. Simultaneously it is executing jobs which appeared earlier in the queue, ending each unsuccessful job with diagnostic statements, and each successful job with instructions about the unloading and saving of magnetic tape. As a third parallel operation, it may also be printing and punching results from a previously-written output tape.

A further innovation with Stretch was the necessity for a fully documented magnetic-tape library. Each tape carries a serial number visually on the reel and magnetically on the tape. These numbers are used by the librarian for cataloguing purposes, by the programmer to identify the tapes needed for his job, and by MCP to check that the correct reels are mounted on the correct units at the required time.

Jobs arrive at the Reception desk in the form of a pack of punched cards comprised of control cards followed by program and/or data cards. Each pack is accompanied by a written instruction slip giving the programmer's name and job number, the maximum time for which the job should be run, and the serial numbers of customer magnetic tapes required by the job. The Reception staff record these details before placing the cards in one of the queues built up according to any priority rules which may be in force at the time. Due to the overlap facility in MCP, it wastes little or no time to read cards directly into Stretch, so these queues form the standard input to the computer. As the contents of the queues are moved periodically to the computer room, the related magnetic-tape reels are extracted from the library so as to be at hand when called for by the computer.

The output from each job, preceded by its identifying job number, is first written onto magnetic tape. It is then printed, punched or filmed either by further processing on Stretch itself or on the 1401 or S.C. 4020 microfilm recorder. The output from the various sources is returned to the Reception staff who sort and collate it ready for return to the originator. Clearly, under this system, no one can know precisely when any particular job will be executed, but average turn-round times can be forecast and can also be regulated by manipulation of priorities on the incoming queues.

MCP also contains certain logging features which supply statistics for costing purposes and for the guidance of management. At the end of every job, MCP records, on the common output tape, the duration of the job and the serial numbers of all tapes used during the job, together with the number of times there has been a detected error connected with tape reading or writing. When the output tape is processed, this information is punched into logger cards which can be sorted, added or listed as required.

The information about tapes enables a life history of each tape in the library to be compiled and a figure to be calculated which gives a measure of its performance. This figure is used to identify faulty tapes which require attention. Experience has shown that tapes become most worn near the beginning, and that by cutting off a length of tape from the load point there is a high chance of eliminating the faulty piece. Only in cases of severe physical damage has it been found necessary to discard the whole reel.

Other Establishments in the Authority have continued to make use of the AWRE computer installation. Culham, Harwell and Winfrith each now have their own 1401 which they use to prepare input tapes for Stretch and to process output tapes from Stretch. Culham sends a courier car three times daily, and Harwell twice daily, each car bringing input tapes for Stretch and waiting to carry back the corresponding output tape. Winfrith has installed a 7702 magnetic-tape link which is used throughout the day for transmission of the more urgent work to and from Stretch. By means of the link a programmer may put a job in for running in the morning and expect to receive his results in the early afternoon. This link would be quite unable to carry the large bulk of production work which is executed during the evening, and in any case there is not the same urgency about receipt of results. So the majority of the work is still transmitted in the form of magnetic tape carried by courier car each evening, thus ensuring that work left by programmers before they go home is completed by the time they arrive at work the next morning. By these means, each Establishment has complete control over its own work load and can arrange its own internal priorities.

The Stretch system has now settled down at a high level of efficiency, an average of 96% measured by our standard of

$$\text{percentage efficiency} = \frac{\text{good computing time}}{\text{good computing time} + \text{unscheduled maintenance}}.$$

So the time has arrived for the energies of the computer staff to be diverted towards planning for the next innovation, an Atlas 2 which is expected at the end of this year.

Acknowledgements

I should like to thank Mr. A. H. Armstrong, my Senior Superintendent, for his help in the preparation of this paper. It is published by permission of the Director of the Atomic Weapons Research Establishment.

Reference

CHAPMAN, F. G. (1963). "Use of large computers at a distance," *The Computer Journal*, Vol. 6, p. 214.

Book Review

Natural Language and the Computer. Edited by PAUL L. GARVIN, 1963; 398 pages. (Maidenhead: *McGraw-Hill Publishing Company Ltd.*, 99s.)

This book is a collection of papers based on lectures given at the University of California in 1960–61. Its purpose is to "survey the emerging field of language data-processing, covering all aspects of the computer processing of natural-language data and providing background material in linguistics, mathematics and computation. . . . The papers offer a cross-section of opinions and approaches representing the major trends in this new area of study." The papers are arranged thematically: the first section of the book deals with various models which may be used in the study of language—the "definitional" model of natural language, as applied in linguistics, the transformational model of grammar, and the informational model of communication in general; the next section discusses the role of mathematics in this field, and the use of computers, followed by papers presenting the linguist's, logician's and statistician's views of language data-processing; the following two sections deal with the main areas of research, machine translation and information retrieval, and there is a general conclusion.

The book suffers from some of the defects of multiple authorship: some papers are more detailed and/or more sophisticated than others. Karush's paper on the use of mathematics in the behavioural sciences, for instance, is rather too general, and Ray's on programming is elementary compared with, say, Edmundson's on the statistician's approach. The derivation from lectures is apparent in the purely descriptive character of a few of the contributions; and some ambiguity of intention is indeed apparent in the book as a whole. The papers in Part 2, such as Ray's, are introductions for the non-specialist; Hayes' paper on mathematical models in information retrieval, on the other hand, would be a challenge to any beginner, and there are others which are much more suited to, and could be much more profitably read by, students with some background in the field than non-specialists.

The book is nevertheless successful as a whole: it covers almost the entire field in considerable detail, with extensive and useful references. It exhibits the connections between natural-language data-processing and other fields like mathematics, and shows how the problems of this field can be approached from quite different angles; at the same time, it brings out the characteristics of this field of research—the vast quantities and intractability of the material to be handled, the difficulties inherent in giving objective descriptions of something as close to the research worker as his own use of his own language, the need for "weak" rather than "strong" formal models—and emphasizes the point that natural-language data-processing should ultimately be treated as a subject in its own right and not as a branch of, say, symbolic logic. The book can be recommended both as a starting point for those interested in this field, and as a valuable reference work and stimulant for those already working in it. There are many suggestive analyses and discussions of different aspects of, and problems in, the field. To take two random examples: Garvin in the first section insists (and in my view rightly) on the fact that the linguistic model of language in terms of levels of units is a descriptive rather than a dynamic model; it acts as a framework for the analysis of language rather than as a foundation for a deductive system in which theorems about language can be proved. Similarly, Hayes' treatment of the documents, the index codes defining documents, the requests for documents, and the interest "profiles" of the users making requests, in an information-retrieval system as logically the same is a most interesting way of looking at the problem of obtaining the maximum communication between library and library user.

There are some omissions. Work being done in this field outside the United States is hardly mentioned: there is no reference, for instance, to Gardin's work on structured encodements of document specifications and retrieval requests. A much more fundamental point is that the fact that meaning is as important as, if not more important than, form in natural language, is not sufficiently emphasized: the uninformed reader would not realize either that the semantic problems which arise in this field are as serious as they are, or that no one has come anywhere near solving them. In machine translation especially, the problem of "multiple meaning"—the need to find the correct equivalent in the output language for a word with several meanings in the input language—remains as intractable as ever. This failure to discuss these semantic questions adequately is particularly unfortunate, since it is here that the real interest and challenge of this field lies.

KAREN S. JONES