date? Feb 1957

PROJECT STRETCH

Some Preliminary Thoughts on Input-Output Problems

by

Harwood G. Kolsky
Los Alamos Scientific Laboratory

I. Introduction

The question as to how the STRETCH I/O computer will work in detail with the many input-output units and the main computer appears to be one of the most complicated and perhaps the most serious set of logical problems to be solved in the design of STRETCH.

One thing which makes discussion of Input-Output hard is that one cannot make definite statements as to how much I/O data is needed to anything like the precision that one can say, for example, how many arithmetic steps will be needed to solve a given neutron diffusion problem. The amount of I/O material for such a problem can vary over at least a factor of 1000, depending on, (a) the numerical scheme used for solving the problem, (b) how much printed or graphical detail is wanted by the problem originator, and (c) how many intermediate results must be saved for later machine calculations.

Some of these items can be studied successfully in the light of experience on present and past problems, often however, they seem to become more questions of philosophy as to how STRETCH should be used. In the following paragraphs I shall attempt to make a list of some of the details of the input-output problem and indicate some probable conclusions as to what should be the STRETCH I/O Philosophy.

II. List of STRETCH Input-Output Devices and their Approximate Speeds.

One may divide the I/O devices into two main functional classes, those which are mainly for communication from one machine to another and those which are mainly for communication from the machine to humans.

I. Devices principally for communication between machines

	Name	time for 1st word	time for succeeding words	Size (no. words)	Other Remarks
1.	RAMAC	0-100 msec	4. μsec	1,048,576	
2.	Magnetic Tapes	10-50 msec	400. μsec	1,000,000 (one 2400 ft tape)	727 speeds assumed.
3.	Card Reader	over 100 msec	10,000 μsec	12,000. 1,000. (one hopper full)	500 cards/min for binary for dec/per card 155 cards/min
4.	Card Punch	over 100 msec	32,000 μsec	12,000 (one hopper full)	for binary

II. Devices for communication to humans

	Name	Speed in lines per minute	Characters per sec	Pages per min	Other Remarks
1.	Typewriter	3	3.3	~0.1	mainly for logical monitoring
2.	Mechanical Printer	1000.	2000.	16.7	
3.	Electronic Printer	9000.	20,000.	150.	as a printer
		----	----	150 or more	as a graph plotter

One should mention here the distinction between "off-line" and "on-line" operation. A magnetic tape may be written for the purpose of being fed into an off-line printer at a later time, however, it still communicates with another machine as far as the above dichotomy is concerned. It may, however, have a considerable effect on how the machine is operated if the results are not available until later. In this connection it should be noted that the electronic printer is essentially "off-line" in operation even though it may be tied directly to the machine simply because the film is not developed and printed until later.

III. Some Obvious Conclusions from the Above Speeds

A. Tapes

Note the huge time disparity between speed of the internal memory (RAMAC) and the fastest means of loading it from the outside (tape). One large tape reel contains about the same amount of information as a RAMAC, however, it takes 400. sec. or 6.6 minutes for a single tape to read this data in whereas the RAMAC could accept it in 4. sec.

Since anything like 6 minutes is a tremendously long time for STRETCH, the main conclusion one can draw is that unloading and reloading any large portion of the RAMAC from tape should be done only once every 2 or 3 hours as regular practice. It is exactly this type of consideration which was used in arriving at the acceptable error-free running times needed.

B. Cards

The time disparity between cards and other devices is even more marked. However, cards are still very useful because of their property of allowing the easy removal and insertion of data. They are an easily manipulated means of controlling the machine and entering small amounts of data. If one includes the time required to take a tape reel off and put on another as contrasted to the time required to put cards in the hopper, the results are startling. If it takes 1 minute to change a tape reel (a reasonable time for the 704) the card reader could have already read in more than 4000 words before the tape starts. It seems likely that for small amounts of data and for control instructions which can be contained in a few hundred words, punched cards will still be a preferred means of input.

The card punch as an output cannot compete with the tape as a means of writing large blocks of data, however, the convenience of cards for use in controlling later runs and for filing away in permanent storage will still make them very useful for records of a few hundred words or so.

The use of auxiliary card-to-tape and tape-to-card machines might alter the "break-even" point as to the number of cards one would punch, but should not eliminate their use altogether.

Even if the tape reel change time is made very fast by the use of cartridge loaders, it still seems likely that cards will be used for control instructions and small amounts of data because of their convenience.

C. Typewriter

The typewriter will probably be used mainly to transmit key words to the operator to inform him of the progress of the calculations. The amount of output data printed will necessarily have to be small because of the slowness of the typewriter, however the operator usually needs only a few carefully selected bits of information to enable him to make whatever decisions are necessary. The amount of direct human intervention should be quite small for a good STRETCH problem. Another use of the typewriter which I forsee is the keeping of a running "log book" of what problems are being run, when each started and ended (since presumably STRETCH will have a built in clock which it can read) and of course the key data words mentioned above. The time can serve as a convenient identification word in many cases. Whatever action the operator takes will be typed out exactly for future reference in duplicate. (NOTE: Had such exact records of operators' actions been available on the 701, they would have saved me a great deal of needless debugging and rerunning of problems in the past few years.)

D. Mechanical and Electronic Printers

It takes only a glance at the number of lines per minute which these devices can produce to emphasize the question which has been raised so often concerning big computers, "Who is going to look at all those numbers?" There has been perhaps a factor of ten difference of opinion between different machine users as to what the proper amount of output data is for a given type of problem. There are actually reasonably good arguments for both extremes. However, here we are not arguing whether there should be 30 pages or 300 printed for a four hour problem on the 701, but whether there should be this many pages printed in <u>two minutes</u>. Clearly one wants neither,--even 3 pages might be somewhat excessive. Note, however, that considerably more thought must go into the few numbers which are to be printed on these 3 pages. One of the main arguments of the 300 page per problem advocates (myself included) in the above example was that they wanted to make sure that whatever numbers that might be needed later would be available in case someone just happened to want to see them and thereby avoid the trouble of rerunning the problem later to print more. The conclusion for STRETCH is that one can no longer afford this luxury of not considering in advance what the numbers will be used for. Quite detailed summarizing should be planned in advance to keep from swamping the customers in irrelevent details.

The use of the graph-plotting feature of the electronic printer should help considerably in cutting down necessary printing. It is an often observed fact that the first thing which most physicists do with the results of a calculation is to plot graphs of the various quantities to see if they "look reasonable." Now we can get some of the graphs directly without the intermediate time consuming steps. The main limitation will probably be that one may not know all the graphs which are needed until after the problem is run.

This field of printed (or plotted) output data is an area which is very greatly influenced by the philosophy of the individual user and strongly dependent on the particular problems being solved. However, I think it is safe to conclude that all STRETCH users are going to have to take a serious look at doing at least a factor of ten more condensing, compiling and analyzing in the machine before printing so that we can keep the output down to a reasonable average of a few pages per minute for regular production running.

These remarks should not be construed as meaning that a printer and plotter of the speeds described is not necessary. The average number of pages per minute may be low, but the peak rate will be essential at times when the problem is being held up until the output is completed.

IV. The Problems of Actually Operating STRETCH

A. Present usage of 704s.

Let us look at the present use of the 704s during the day shifts at Los Alamos. The average person asks for a 15 or 20 minute period. During this period he either makes runs of short problems, sets up long problems to be ready for running during the night shifts, or does debugging of new codes. If the same mode of operating were carried over to STRETCH, the average person would have perhaps 20 seconds on the machine. He clearly could not change tape reels and switches in this time, but even assuming there were spare tape units to use ahead of time, there would still be a tremendous crowd of 20 or 30 people buzzing around the machine trying to get their tapes on and off, looking for their listings, getting their cards mixed up, etc., at any given time.

One conclusion might be that there should be no 20 second periods allowed on STRETCH--that the customer save up his 20 second periods until he gets, say, 5 minutes worth before going on the machine. There will probably be a natural trend in this direction, however, having a fixed rule of this kind certainly would encourage the waste of machine time. In describing a possible method of operation let us list the types of problems which will probably be encountered.

B. Types of Problems

The following types of problems are presently encountered on existing machines and will probably exist on STRETCH with the differences noted.

(1) Assembly of Codes: It is possible that the actual assembly (or encoding by an "automatic" coding system) may be done on a smaller computer such as the 704, but I will assume for the present that this will prove to be less convenient than using STRETCH itself.

The original input to the assembly program will be in the form of decimally and alphabetically punched cards, which will probably be put on tape first using an auxiliary card-to-tape machine. The output of the assembly will be a tape with the code writtern on it and either a tape to be printed on the auxiliary printer or preferably printed directly on the mechanical or electronic printers.

The tape input and output will be quite reasonable. The printed output may be large, but should only take a few minutes on the electronic printer even for an enormous code.

One possible means of avoiding the shuffling of tapes which occurs to me is to have the code formed by the assembly program be assigned a permanent position of one of several reserved tape units. The assembly program will punch out a binary card which will serve the operator as a control card to call in his code whenever he wants to use it; much as utility programs are called in on the 704 now. In this way he need never remove his code tape from the machine. The tapes containing small problems could be removed and saved as a regular procedure before the period when big problems are to be run if absolutely necessary. (See below.)

(2) Debugging

With the improved machine-assisted coding systems under consideration, the debugging of codes should take on quite a different aspect. The troubles will be more logical mistakes and incorrectly formulated problems rather than bookkeeping errors. Standard debugging will probably involve furnishing control cards for traps, etc., which will result in a printed record of selected operations or regions in memory. These will then be examined at his leasure by the programmer who need not have been present when they were run off.

Debugging at the machine keyboard should be limited to two types, (1) very small changes which can be entered quickly, and (2) very serious difficulties which defy ordinary methods. In spite of advanced methods these latter will probably always be with us and will have to be done the hard way. Note: There will be the need for an improved class of debugging programs which look for the logical type difficulties which will plague the STRETCH user.

(3) Short Problems

On the 704 a short problem may be defined as one which runs about 15 minutes or less at a time. On STRETCH such a problem would take 10 seconds. Although there will probably be a natural tendency away from such short problems as was mentioned before, there will probably remain a fair number of problems which will take a minute or less for completion. Of course one will be willing to accept rather low machine efficiency for such problems in exchange for human convenience.

Here, as in debugging, the problem could be handled entirely by control cards without the necessity of the programmer being present.

The control cards can call the code off a tape, and read in the conditions or new variables for the particular problems, they can also stipulate what is to be printed and how. When the problem is finished, the card reader can call in the next problem automatically, while the typewriter is recording what was run on the typewriter "log."

Neither debugging nor short problems should make very heavy demands on the I/O units, with the possible exception of some data processing jobs. For all such jobs one will not worry about efficiency as much as convenience.

(4) Big Problems

These are the problems for which STRETCH is mainly needed--the ones which can be done only in rudementary fashion on present machines, but which we really need to do. If past experience is a guide, we will probably find that a big problem on STRETCH will take about two or three hours and use all parts of the machine thoroughly. Problems taking longer than this will probably be parameter studies on shorter problems. Although there will certainly be a few real gigantic multidimensional calculations taking several days being considered at any given time, the chances are that they will not be representative of the bulk of STRETCH's work.

The big problem runs may be isolated in time as they are on the 704's now. However, there is a distinct possibility that it might prove considerably easier for the human operators if several short problems could be sandwiched in between each long problem or if they could be put in natural breaking points during a long problem. This would enable the operator to get his deck of control cards assembled more or less leisurely while waiting for the completion of the next part of a long problem. The switching from problem to problem would then be done automatically by STRETCH. The last control card on the deck would be to resume the long problem where it left off. The operator could then proceed to get the next set ready. Perhaps two operators alternating would be needed for tight schedules.

The real difficulty in such a procedure is one of providing enough tape space for all the active codes without either hampering the big problems's need of tapes or making a lot of tape-reel shuffling necessary.

The big problems will make the hard demands on STRETCH's internal input and output system, but not so much on the external equipment.

The load on the card reader and punch will probably be quite light. The printers will also probably not be used too heavily by big problems simply because of the human impossibility of digesting the reams of data mentioned earlier. A few pages or graphs a minute on the average will probably be all that are needed during regular running.

The internal data stored on the RAMAC or tapes for analysis later in the problem can be tremendous. In fact one of the natural breaking points mentioned above in a big problem, will be when these external memory devices are getting full so that the problem will have to be stopped and a processing code called in to distill out the information needed before proceeding.

In this connection it is clear that the 727 type tapes are too slow to be used very effectively as an active external memory. The RAMAC is quite good for block data transferring but is limited to a fixed million words and cannot be read or written on at random without a big time loss. If one wants to both read and write large amounts at high speed, probably a <u>pair</u> of RAMAC's will be needed.

The input-output pattern for bit problems will probably be as follows: The initial input will be relatively small; a few thousand words of code and a few thousand input data words. The problem will then generate large amounts of data as it proceeds; part of it being saved on the RAMAC for later analysis. When an appropriate place is reached in the calculation, a processing code is called in which will take the data from the RAMAC and calculate the information needed by the problem originator. When the problem is completed or interrupted for later restarting it should be done, if possible, at a time when the data on the RAMAC can be abandoned in order to avoid the lengthy read off time.

The exact way in which the RAMAC should be used during the course of large calculations depends critically on the nature of the calculation. I will not go into this subject now since this is supposed to be just a general discussion, but I do hope to do so in a forthcoming paper on two-dimensional hydrodynamics.

V. Summary

The tentative conclusions which I have drawn from this casual study of STRETCH input-output problems are: (1) Punched cards will be very useful for controlling the course of small problems and for switching quickly from one problem to another, (2) One means of cutting down on the input-output human bottleneck is to keep all active codes on tapes permanently attached to the machine, calling them in by means of control cards when needed, assuming enough tapes were available, (3) The 727 tape speed is so slow that one cannot consider using them as active external storage during a problem or for loading and unloading more than a small part of the RAMAC, (4) The electronic and mechanical printers will probably have a fairly low duty cycle during big problems, although they will be called on to operate at top speed in "spurts," particularly on smaller problems. (5) An auxiliary tape-to-printer probably will not be essential, although a card-to-tape, tape-to-card equipment probably will. (6) The typewriter will be very useful as a monitor and as a log keeper for the details of operator decisions. (7) STRETCH's main load of work will be problems which run for two or three hours, but special consideration must be made to handle the large number of small problems lasting a minute or less. Less worry is needed for the giant problems requiring over 10 hours provided the machine is reliable. (8) Debugging and assembly could be handled much like small problems, i.e., using control cards. (9) A convenient means of operation might be to have small problems sandwiched in between breaks in large problems so that the human operators can keep up with the input and output devices.

Distribution:

S. Dunwell
F. Beckman
J. Griffith
L. Sarahan
D. Sweeney
B. Carlson
R. Frank
Max Goldstein
H. G. Kolsky
R. B. Lazarus
N. C. Metropolis
E. A. Voorhees
D. F. Woods
T-5 Files