7

A NON-SEQUENTIAL BUFFERING SYSTEM USING A VARIABLE NUMBER OF TAPES.

by

D.W. Barron (1), D.F. Hartley (1) and C.R. Spooner (2)

(1) University Mathematical Laboratory, Cambridge, England.

(2) Ferranti Ltd.

1. INTRODUCTION

It is now almost standard practice to use magnetic tapes to buffer the input and output of large computers. The normal practice is to use separate tapes for input and output and to operate them on a strict first-on-first-off principle. The Atlas 1 supervisor (Kilburn, Payne and Howarth, 1961) broke with this tradition by abandoning the sequential operation of the input and output tapes. More recently a smaller version of Atlas 1, called Atlas 2, has been designed, and the Atlas 2 supervisor departs even further from tradition in that separate identity of the input and output tapes is lost, and the available tapes are amalgamated into a multi-purpose buffering system.

The advantage of a non-sequential buffer is that it has a sorting ability, so that it can accept information in parallel streams at an irregular rate and deliver the streams in sequence, at a rate to suit the receiving program. The advantage of a multi-purpose buffer over conventional separate input and output buffers is that in this way one can get better performance out of a small number of tapes. This is important to Atlas 2 since the basic machine has neither drums nor discs nor satellites. In fact, the buffering system described in this paper is sufficiently powerful and efficient for satellites to be unnecessary, even for the largest installation.

In this paper we first describe the characteristics of the buffering system, and the "external" view, then we describe in some detail the way it is implemented. Details of the Atlas 2 magnetic tape system are only given to the extent that they are needed for an understanding of the buffering scheme.

2. CHARACTERISTICS OF THE BUFFERING SYSTEM

The general-purpose buffer system is known, within the Atlas 2 supervisor, as the WELL, and this convenient short name will be used in the remainder of the paper. The basic unit handled by the well is the DOCUMENT, which is a self-contained piece of program or data. The well is designed to cope with the following general situation.

(a) Documents are small to medium large in size (not more than about 80,000 characters).

(b) Documents arrive at random times, in parallel, at a not too great average rate (preferably not more than 4,000 characters/second, certainly not more than 7,000 characters/second).

(c) Requests for documents from the well can be scheduled ahead, fairly accurately over a short period (about 5 seconds), less accurately over a longer period (about 30 seconds).

The well consists mainly of magnetic tape storage, but uses some core store in addition, being designed to give the maximum ratio of capacity to core store used, consistent with the input and output rates required. It operates with a dynamically variable number of tapes (between 1 and 4), but to the other parts of the system it appears as a "black box", with the same external appearance whatever the number of tapes being used.

3. DETAILED DESCRIPTION OF THE WELL

3.1 The Magnetic Tape System

Atlas 2 uses pre-addressed tapes with fixed block length: tapes are written forwards, to be read forwards or backwards. The well is designed to work most efficiently with this sort of tape system, though none of the features is absolutely essential. However, it does rely on selective reading from records, which is difficult without pre-addressed tapes, and without backward reading the "red-tape" would be formidable. The block length on Atlas 2 is 512 words, with a scatter facility which allows eight 64 word BLOCKLETS to make up a block. 512 words is an inconveniently large unit for the well, which works exclusively with blocklets.

3.2 The Swinging Tape

The well is based on a variant of the swinging tape principle described by Kilburn et al (loc.cit.) A swinging tape operates in conjunction with two core store buffers, the READING and WRITING buffers, and at any stage information on the tape is contained in a certain region, the space beyond being as yet unused. The tape swings over this region and each cycle consists of three phases: READING, TRAVERSING and WRITING. The reading phase is initiated when the tape is at the upper end of the information region and moving backwards. As the information is scanned any items that have been requested are read and placed in the reading buffer. At the end of the scan (when the tape has reached the lower end of the information region) the tape is reversed and is run forward to the upper end: this is the traversing phase. The contents of the writing buffer are now written to the tape, thereby extending the information region; the tape is then reversed again and the reading phase started. During the reading and traversing phases material for the tape accumulates in the writing buffer; during the traversing and writing phases the reading buffer is emptied. The buffers must be large enough to deal with the information flowing to and from the tape during one swing (assuming an average rate). Each writing phase uses a new section of the tape and since items are normally read only once, the information region gradually creeps forward. The maximum length of the information region is fixed by the buffer size, and therefore it is necessary that items are not required to stay on the tape for longer than a critical time, determined by the swing time and the average rate of arrival of information. If this condition is not satisfied, special steps have to be taken.

3.3 Combination of tapes in the well

The single swinging tape as described above has two drawbacks. Suppose we want a well capacity of about 800,000 characters, then we need (on Atlas 2) a swing of about 30 seconds. If the mean rate of flow to and from the well is 3,000 characters/second (which seems typical for a medium-large installation), we must provide core-bufferage for the order of 200,000 characters, which requires about 25K of core store, an inconveniently large amount! Also we have to schedule requests for documents 30 seconds ahead, and this may be difficult to do accurately. We therefore use a subsidiary swinging tape, with a swing time of about 5 seconds, to buffer information on and off the main tape. This reduces the core-space required to the order of 10K, and means that we only have to do accurate scheduling 5 seconds ahead, with a less accurate scheduling 30 seconds ahead. The time taken in additional tape transfers is negligible.

The figures for core store use are for the average case, and are rather larger in the worst possible case. For efficient working, the well is used in conjunction with a dynamic storage-allocation scheme which can make space available to meet the worst case, and can absorb the unused space when the well is not being used to full capacity.

If more tapes are available first the main swinging tape, and then the subsidiary swinging tape, can be provided with a feeder. Whilst a swinging tape is in the reading or traversing phase, information destined for it is written to the feeder, which moves progressively forwards from a fixed base line. When the swinging tape starts the writing phase the feeder is read backwards, and its contents transferred. Material arriving for the swinging tape during the writing phase goes straight to it, the transfer from the feeder tape being temporarily halted. The use of feeder tapes reduces the core store requirement for buffering; also, if a feeder is provided for the main swinging tape there is always a tape available to accept information coming to the well, and this serves to smooth out peaks in the information rate. It may appear that the feeder tape is used inefficiently, but it should be remembered that it would otherwise be idle.

To change the number of tapes used by the well is easy: it is only necessary to divert the flow of information to or from a tape, and it will fill up or drain within a few seconds by the normal working of the well. Thus it is possible to use any tapes that are available, so that an otherwise idle tape can temporarily increase the efficiency of the well system.

4. PERFORMANCE OF THE WELL

The best index of well performance is the ratio of well capacity to core-store used, which is given in the table below.

Number of tapes	Rate of flow (characters/second) 500	1000	1500	3000	6000
1	25	13	8	4	2
2	> 5000	140	50	12	3
3		400	150	35	9
4		700	250	60	16

The figures show that at low rates of flow the subsidiary well tape makes most difference: the feeder tapes become important at moderate rates of flow. The performance falls markedly at very high rates of flow, but these are not expected to arise in the applications for which the well is designed.

5. ORGANIZATION

Documents are stored in the well as chains of blocklets, the first two words of each blocklet being used for chain links and identifying information. It is not necessary to keep a record of what is on a feeder tape, since it must inevitably be transferred to the associated swinging tape within a short time, but a dictionary is needed for each swinging tape, giving the position of every document currently on the tape. These dictionaries are updated whilst information is being written to the tape, and are consulted by the scheduling routines at the beginning of each reading sweep. The dictionaries, and the associated routines, are not required during the rest of the swing, and they are therefore stored on the appropriate swinging tape, being written at the end of the information region after the writing phase, and brought back to core at the beginning of the subsequent writing phase.

6. CONCLUSION

The well was designed initially as an input-output buffer, but it can be used for many other purposes, for example, holding the inter-pass streams of multi-pass compilers. It provides, in an economical manner, efficient input-output with many convenient user-facilities, in particular multiple input-output streams, and it enables the system to provide all the advantages of buffered input-output with the convenience that to the programmer the peripheral devices appear to be on-line.

This work is part of the Atlas 2 project undertaken jointly by the University Mathematical Laboratory, Cambridge and Ferranti Ltd., whose permission to publish this paper is gratefully acknowledged.

REFERENCE

Kilburn, T., Payne, R.B. and Howarth, D.J. "The Atlas Supervisor", Proc. E.J.C.C. 1961, p.279.