J. C Griffith

June 28, 1957

FILE MEMO

SUBJECT: Streaming for the Basic Machine

BY: W. Buchholz

1. Introduction

There are a number of important operations in commercial data processing and scientific computing which call for the transfer in memory of substantial blocks of data in a progressive, though usually not entirely uniform, manner. Well-known examples are:

Re-arranging records within groups of records.

Re-arranging fields within a record.

Moving data from slower to faster memory, and the reverse.

Compressing data for more compact storage on tapes and in large, slow memory.

Expanding data for faster processing in faster memory.

Code conversion.

Editing input data.

Editing output data.

With the powerful variable field length and indexing operations already specified for the Basic Machine, these operations can be performed quite readily and certainly very much better than in existing machines.

A substantial further improvement can be obtained by adapting the HARVEST streaming concepts to the Basic Machine in a much simplified form. The advantage is a reduction in the number of memory cycles taken and the number of instructions needed for equivalent operations. It is possible to specify a simplified streaming mode for the Basic Machine with only modest additions to the equipment needed anyway for variable field length.

The justification for adding these features is two-fold. They are basic to commercial data processing, although it is hard to predict now how much the overall saving will be. Second, for scientific computing, a reduction in 2-microsecond memory cycles will help to relieve what is likely to be a major bottle-neck. Streaming will greatly facilitate the compression and expansion of data coupled with a transfer to 0.5 - microsecond memory for greater speed and less conflict with high-speed disks and tapes.

Another advantage of streaming is that it removes the address limitations of the previously defined TRANSMIT and RECEIVE instructions.

2. Characteristics of Streaming for Basic Machine

Streaming, as defined here, provides the following features:

(a) Two 26-bit starting addresses may be set up, one for the source of data and one for the destination.

(b) Two field lengths may be set up, one for the source and one for the destination. Field lengths may specify from 1 to 64 bits.

(c) Instructions may cause the source address or the destination address or both to advance by the amount of the corresponding field length. The corresponding information is passed serially from the Source register to the Destination register or to the Index Arithmetic Unit (for table look-up purposes).

(d) All advancing is from left to right. All fields are treated as if they were unsigned binary fields. "Byte Size" is not defined but is adjusted automatically to be 8 bits except for the last byte which may be less than 8 bits to complete the field length.

(e) Memory access is made only when the set-up is changed, when a source word is empty, when a destination word is full, or when the registers are needed for any other purpose. Access control is automatic.

(f) In contrast to HARVEST streaming with its extremely flexible instruction format, the streaming instructions for the Basic Machine resemble ordinary variable field length instructions. The operation is completely defined by an operation code. To reduce the number of set-up words needed, the instructions can define part of the set-up to the extent permitted by format restrictions. Parts of the set-up can be altered this way at any time without changing the remainder. All unspecified set-ups are used as they were at the end of the last operation. Change of address set-up is optional; a zero address field leaves the previous address set up unchanged.

(g) A table look-up instruction is defined which transforms a source field of one length into a new destination field of the same or different length. Table fields to be extracted can be of any size, spaced apart by any power of 2 up to 64 bits. To allow for exceptions, the instruction contains an 8-bit match byte which is compared with the low-order bits of the extracted table field; a match condition actuates the Program Interrupt feature, thus providing for branching. Tests by Branch instructions are also available.

(h) A Repeat Count feature is provided with some instructions. The Index Address specifies an index word whose limit field gives the number of times the operation is to be repeated. Program Interrupt can occur at any cycle of repetition.

3. Format

The attached diagram indicates relevant word formats. One change from the previously assumed instruction format is proposed for all instructions. The previous modifier bits are replaced by unique operation codes to make room for 7 more bits. The operation code is expanded to 8 bits to allow for the increase in the number of operations.

The 7 bits are used, in arithmetic operations, to specify the shift, that is, the starting point in the A B registers. (The register letters used are those on the attached diagram; A B is generally used as a double-length accumulator, C D as a memory register.) In other instructions, the 7 bits are combined with the Second Address Designator (SAD) to give an 8-bit Byte field which is used for matching or inserting bytes.

The index word is unchanged. The two set-up words follow the instruction format and can, in fact, be instructions. The Table Look-up (TLU) instruction also follows the regular format.

4. Instructions

The following is not intended to be a precise description of the instructions but is sufficient, with the previous explanation, to outline the streaming system.

4.1 SET UP SOURCE (SUS)

The effective address, after indexing and substitution, if any, and the field length become the new source set-up.

4.2 SET UP DESTINATION (SUD)

Similar to 4.1 for destination set-up.

4.3 TRANSMIT (TMT)

The address and field length refer to the destination, the source having been set up previously. The destination word or words are merely transmitted to memory, replacing all information previously stored in those memory words. This instruction is primarily intended for storing full words, although the address and field length permit fields of any length to be stored. (Bits adjacent to the field may be preserved by using TMT-VFL at the cost of more memory cycles.)

The Index Address specifies an index word which is used as a Repeat Counter. There is no address modification.

If the address is zero, the destination address left by the previous address is used.

This instruction replaces the TMT instruction in the HARVEST Manual.

4.4 TMT-VFL

This is identical with TMT except the previous contents of each destination word are loaded into the destination register before new bits are inserted. Thus all bits not specifically altered are preserved, as in the VFL type of STORE instruction.

4.5 DISTRIBUTE (DIST)

The address, as modified by indexing, specifies the destination. The field length applies to both source and destination. The source address is used as left at the end of the previous operation. The operation cannot be automatically repeated. Unaltered bits are preserved in memory (as in TMT-VFL). This instruction is convenient for distributing successive fields of different lengths to scattered memory locations.

4.6 RECEIVE (RCV)

4.7 RCV - VFL

4.8 COLLECT (COLL)

These are similar to TMT, TMT-VFL, and DIST, except that the words "source" and "destination" are interchanged. They are used to assemble fields into a single output stream with a minimum of set-up.

4.9 INSERT (INS)

The Byte field specifies a byte to be sent to the destination register. The address and field length refer to the destination. Unaltered bits are preserved. The Index Address specifies a Repeat Counter.

4.10 TABLE LOOK-UP (TLU)

The instruction fields are defined on the diagram. The source field, as defined by the source address and field length, is added to the Base Address given in the instruction starting at a bit position defined by the Offset. The Offset is the left-most bit position, relative to 000000 which is the left end of the Base Address. The sum is the effective 26-bit address in memory of a table field whose length is defined by the Field Length given in the instruction.

The low-order 8-bits, or less for shorter fields, are compared with the low-order bits or all of the Match Byte specified in the instruction. If all bits are equal (match), a Program Interrupt occurs. If the bits do not match, the table field is sent to the destination register.

The Index Address specifies a Repeat Counter.

4.11 TLU-NM

This is the same as TLU, except that no matching takes place and all fields are sent to the destination.

4.12 BRANCH ON MATCH (BR-M)

The next 8 bits or less of the source stream are tested against the Match Byte given in the instruction , but the source address is not advanced. The field length specifies the number of bits, up to 8, to be tested starting at the right. The address and index specify the branch address in the usual manner. Branching takes place ff the specified bits are equal.

4.13 BRANCH ON NO MATCH (BR-NM)

Similar to above, except that branching takes place if the bits are not all equal.

4.14 INCREMENT (INC)

The previously defined INC instruction is used to advance or backspace either the source or the destination address by giving the address of the appropriate address register. Memory accesses are initiated automatically if needed.

Attachments:
Appendix I
Appendix 2

APPENDIX I

Comparison of Simplified Streaming with Corresponding Programmed Operations

The following comparisons apply only to some specific operations which streaming was intended to do and gives no indication of the overall benefit on actual jobs. A time estimate is made roughly in terms of memory cycles. The effect of memory cycle overlap and the possible use of more than class of memory are ignored because they will benefit both kinds of operations.

A. Code translation for 5-character field, no exceptions.

a. Streaming

	EMC *
SUS	2
SUD	2
TLU (repeat 5 times)	11 (1+5x2)
	15

* Estimated Memory Cycles

b. Programmed

		Address	Index	Shift	EMC
RCV 5	1. RCV	(Set up two index words I_1 and I_2)			5
Load 2	2. LOD	Source Area	Byte (I_1)	Offset	15 (5x3)
Add to IR 2	3. LOD	Table Base	Accumulator		10 (5x2)
Connect 2	4. STO	Destination Area	Byte (I_2)		20 (5x4)
INC 2	5. INC	(Advance I_1)			15 (5x3)
INC 2	6. INC	(Advance I_2)			15 (5x3)
Tr 1	7. BIF	2.			5 (5x1)
					85

9xC=45

B. Compression of Floating-Point Words

Assume a standard word having a 10-bit exponent with sign, a 48-bit mantissa with sign, and 4 tag bits is to be compressed into a 5-bit exponent with sign, a 20-bit mantissa with sign, and no tag bits.

a. Streaming

	EMC
1. RCV (set up index word)	-----
2. SUD (set up destination)	-----
3. LOD (full source word, indexed)	3
4. COLL (Address accumulator; transfer 26 bits)	2
5. COLL (Address accumulator; transfer mantissa sign)	1
6. INC (Advance index)	3
7. BIF (Branch to 3)	1
	10 per word

b. Programmed

			EMC
RCV		1. RCV (Set up 2 index words)	-----
LOD	3	2. LOD (Full source word, indexed)	3
COM	3		
STO	4	3. STO (Shift out 26 bits)	4
INC	3	4. STO (Shift out mantissa sign bit)	4
BIF	1		
	14	5. INC (Source)	3
		~~6. INC (Destination)~~ use differential	
		7. BIF (Branch to 2)	1
			18 per word

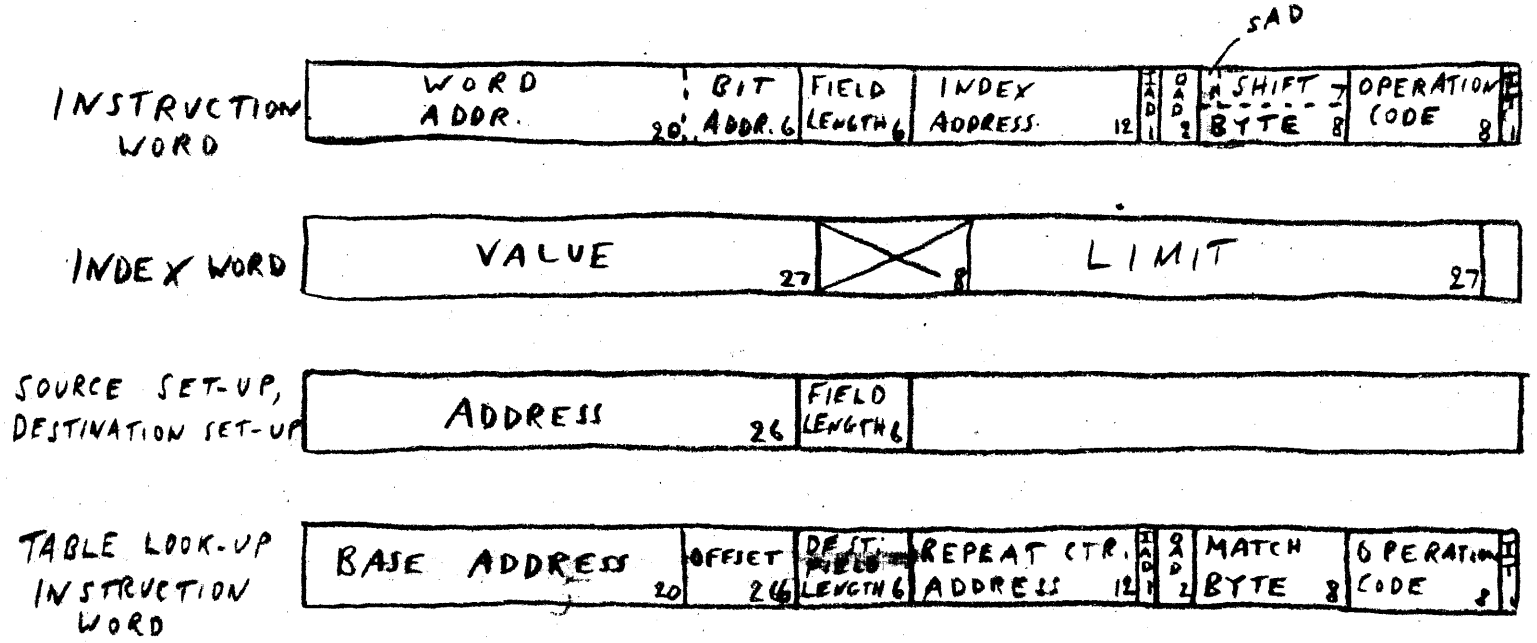
APPENDIX 2

Editing for Printing

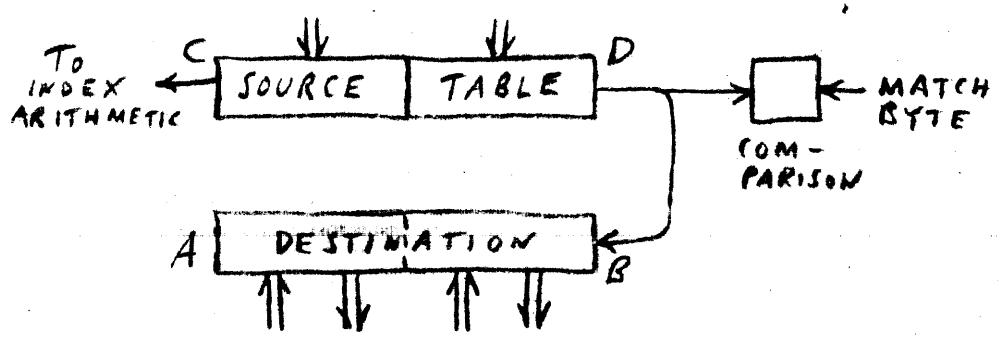
A file memo by H. C. Montgomery, dated 11 June 1957, with Revision No. 1 of June 18, 1957, outlines a scheme of editing for printing using a streaming approach. The present scheme is intended to be applied to this job, among others, in an equivalent fashion.

The following is a rough table of equivalent instructions.

H. C. M.	W. B.
1. Substitute on Match	BR - NM, followed by INS
2. Branch on Match	BR - M
3. Insert Character	INS
4. Pass Byte	TMT, RCV, or variants
5. Adjust Output Field	SUD or INC
6. Skip Byte	INC
7. Table Look-up	TLU or TLU-NM



WORD FORMATS



ASSIGNMENT OF FOUR DATA REGISTERS FOR TABLE LOOK-UP

W. B. 6-28-57