HONEYWELL INTEROFFICE CORRESPONDENCE

PHOENIX OPERATIONS – HONEYWELL INFORMATION SYSTEMS

DATE 770222 PHONE MAIL ZONE COPIES

TO Architecture File

FROM Lou Krasny LouK

section: 14.4
doc# : 0131A-0
page : 1

COMPONENT LADC

SUBJECT CP-6 Use of IOM Paging Mechanism

1.0 INTRODUCTION

This document presents a brief overview of the IOM paged mode of operation followed by a proposal describing a subset of the IOM paging facilities that I believe will be satisfactory for CP-6. This document deals with the IOM modes of operation and and not with software data structures such as Device Control Tables, I/O Enqueuing Tables, etc. Hopefully, the selection of IOM modes will be independent of the specific implementation of the I/O system supervisor software.

This document is written from the perspective of the I/O system supervisor. For example, the term "user's buffer" as used frequently in this document means the buffer which the caller of the I/O system supervisor has identified as the data source or destination for the physical I/O operation. This buffer is usually not the same buffer as the one specified by the user program to the operating system (e.g., a M$READ service call to file management for a record from a keyed or consecutive file).

2.0 REFERENCES

The IOM information in this document has been obtained from the following HIS documents :

EPS-1 6000B IOM Central 43A239654

EPS-2 IOM-B 43A177745

3.0 BRIEF DESCRIPTION OF IOM PAGING MECHANISM

The IOM's that we will be using for CP-6 are capable of operating in any of three modes as selected by a manual mode select switch on the configuration panel of each IOM. The three modes of operation are: 1) Standard GCOS, 2) Extended GCOS, and 3) Paged. For CP-6, all IOM's will be operated with the mode select switch in the Paged position and we will be concerned no further in this document with other operating modes.

With the IOM mode switch in the Paged position, each IOM channel can operate in either one of two sub-modes under program control, the paged channel mode and the extended channel mode. The Paged Channel Mode allows paging via

the use of a page table in memory. The page table address is individually specified for each channel operating in this mode. The extended Channel Mode does not allow paging, however, both DCW lists and data can be located anywhere within 2**24 words of real memory via the use of a 6-bit address extension register. There is a unique address extension register for each channel operating in this mode.

Each of these two sub-modes is further divided into many additional sub-modes. Figure 3.0 (taken from the IOM EPS-1) is a state diagram of these IOM operating sub-modes showing the allowable transitions which are controlled by specific bits in PCWs, LPWs, and TDCWs. The terms real, extended, paged, and segmented are briefly defined in the following paragraphs:

- Real addressing

  Addresses are treated as absolute addresses suppling a bias from real memory location zero. Only the lower 256K words of real memory can be addressed in this mode.

- Extended addressing

  Addresses are treated as absolute addresses suppling a bias from real memory location zero. With the six-bit address extension register, any location within 2**24 words of real memory can be addressed. A single I/O operation can access data within any 256K word block that starts on a 256K word boundary in real memory. This mode is available only for channels operating in the extended channel mode. As shown in Figure 3.0, once a channel is in this mode, no transition is allowed to any mode that utilizes paging.

- Paged addressing

  Addresses are treated as a bias from the beginning of a virtual portion of memory mapped by a 256 word page table in memory. The address of the page table is specified separately for each channel via a field in the PCW. Since each page table word maps a 1024 word page, this mode allows addressing of up to 256K words of memory distributed in real memory as specified by the contents of the page table words. This mode is available only for channels operating in the paged channel mode.

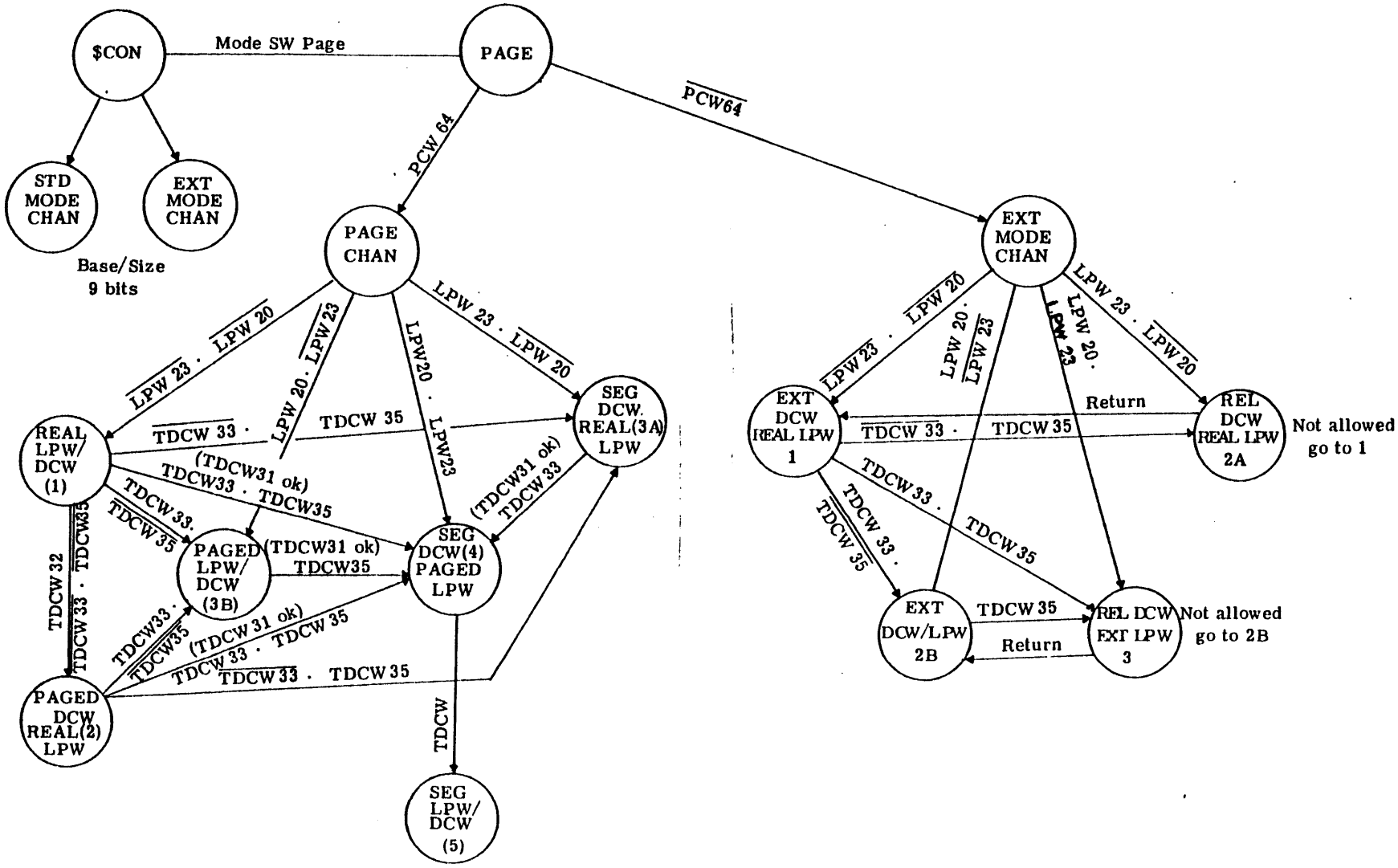
$CON
Mode SW Page
PAGE
STD MODE CHAN
EXT MODE CHAN
Base/Size 9 bits
PCW 64
PCW64
PAGE CHAN
EXT MODE CHAN
LPW 23 · LPW 20
LPW 20 · LPW 23
LPW20 · LPW23
LPW 23 · LPW 20
REAL LPW/ DCW (1)
TDCW 33 · TDCW 35
SEG DCW. REAL(3A) LPW
(TDCW31 ok) TDCW33 · TDCW35
TDCW33 · TDCW35
TDCW 32
TDCW33 · TDCW35
PAGED LPW/ DCW (3B)
(TDCW31 ok) TDCW35
SEG DCW(4) PAGED LPW
(TDCW31 ok) TDCW33
TDCW33 · TDCW35
(TDCW 31 ok) · TDCW 35
TDCW 33 · TDCW 35
PAGED DCW REAL(2) LPW
TDCW
SEG LPW/ DCW (5)
EXT DCW REAL LPW 1
Return
REL DCW REAL LPW 2A
Not allowed go to 1
EXT DCW/LPW 2B
TDCW 35
Return
REL DCW EXT LPW 3
Not allowed go to 2B

Figure 3.0

o Segmented addressing

Addresses are treated as a bias from the address specified by a lower bound (LB) field contained in the payload channel LPW. LB is itself treated as a bias from the beginning of a virtual portion of memory mapped by a 512 word page table in memory. The address of the page table is specified separately for each channel via a field in the PCW. Segmented addressing allows addressing of a block of up to 256K words of memory that are accessed through any 256 consecutive page table words contained within a 512 word page table. This mode is available only for channels operating in the paged channel mode.

In Figure 3.0, LPW addressing sub-modes are defined as those used to access the channel's DCW list and DCW addressing sub- modes are those used to access the channel's data locations. For example, the mode identified as 3A which is almost in the center of Figure 3.0 is described as "SEG DCW, REAL LPW". This means that the addresses in the LPW and TDCW (identified as DCW Pointers) are interpreted by the hardware as real addresses and the addresses in DCW's (identified as Data Address) are interpreted as segmented addresses.

The remainder of this document describes a proposal for CP-6 that uses paged channel modes 1 and 3A for paged I/O operations and extended channel mode 1 for non-paged I/O operations.

Appendix B of the IOM EPS-1 indicates that GCOS was expected (by the hardware designers) to use paged channel modes 1, 2, and 4 for paged I/O operations.

4.0 FACTORS CONSIDERED IN SELECTION OF PROPOSED SCHEME

(1) We wish to make maximum use of the hardware paging capability. When physical I/O is performed for a user whose page table is in memory, we would like to do the I/O through that page table without having to pickup and move page table words and without having to inspect the address range of the user's buffer for page boundary crossings.

(2) It seems natural to communicate the address of the I/O buffer to the I/O system supervisor by means of a descriptor (or possibly a vector from which a descriptor can be easily generated).

(3) The method of utilizing the IOM paging mechanism should facilitate the disassembly of the descriptor

into an I/O queue entry and the eventual construction of an I/O command list from the I/O queue entry.

(4) Some physical I/O is not directly related to a user (e.g. file read-ahead). For this type of I/O, it seems natural to place the DCW list in the fixed monitor data area and use an IOM mode that accesses memory directly rather than have to construct a dummy page table. An extension of this reasoning suggests that since we must place DCW lists for non-paged I/O within the fixed monitor data area, we also place DCW lists for paged I/O within the fixed monitor data area. The proposal for paged I/O described in the next section assumes that DCW lists are always accessed using real addressing and are therefore always located in the lower 256K words of real memory. It should be noted, however, that it is possible to construct other scenarios that would allow the DCW list to be located in the user's virtual memory and accessed via the user's page table. Real-time and/or T&D requirements may require that we also support DCW lists within the user's virtual space.

5.0 PAGED I/O

The proposed use of the IOM facilities for paged I/O is shown in Figure 5.0. Figure 5.0 shows the transfer of information from a type 0 descriptor framing the user's buffer to the command list for the case where a working space quarter other than zero is specified by the contents of the working space register identified in the descriptor. Examples of the user's buffer for this type of I/O are a blocking buffer within the user's File and Co-op Buffer segment or a data buffer within the user's Instruction Segment.

The following assumptions are made for the purpose of simplifying Figure 5.0 and the description of paged I/O:

- User's buffer starts and ends on a word boundary
- User's buffer <= 4096 words
- User's virtual space limited to 512K words.

These assumptions are not hard and fast restrictions. After the simple case is presented, algorithm modifications to relax each of the assumptions will be described.

In additon to the channel number, the PCW pointed to by the connect channel LPW contains a page table pointer and a group of flags that control whether the channel is to be operated as a page mode channel or an extended mode

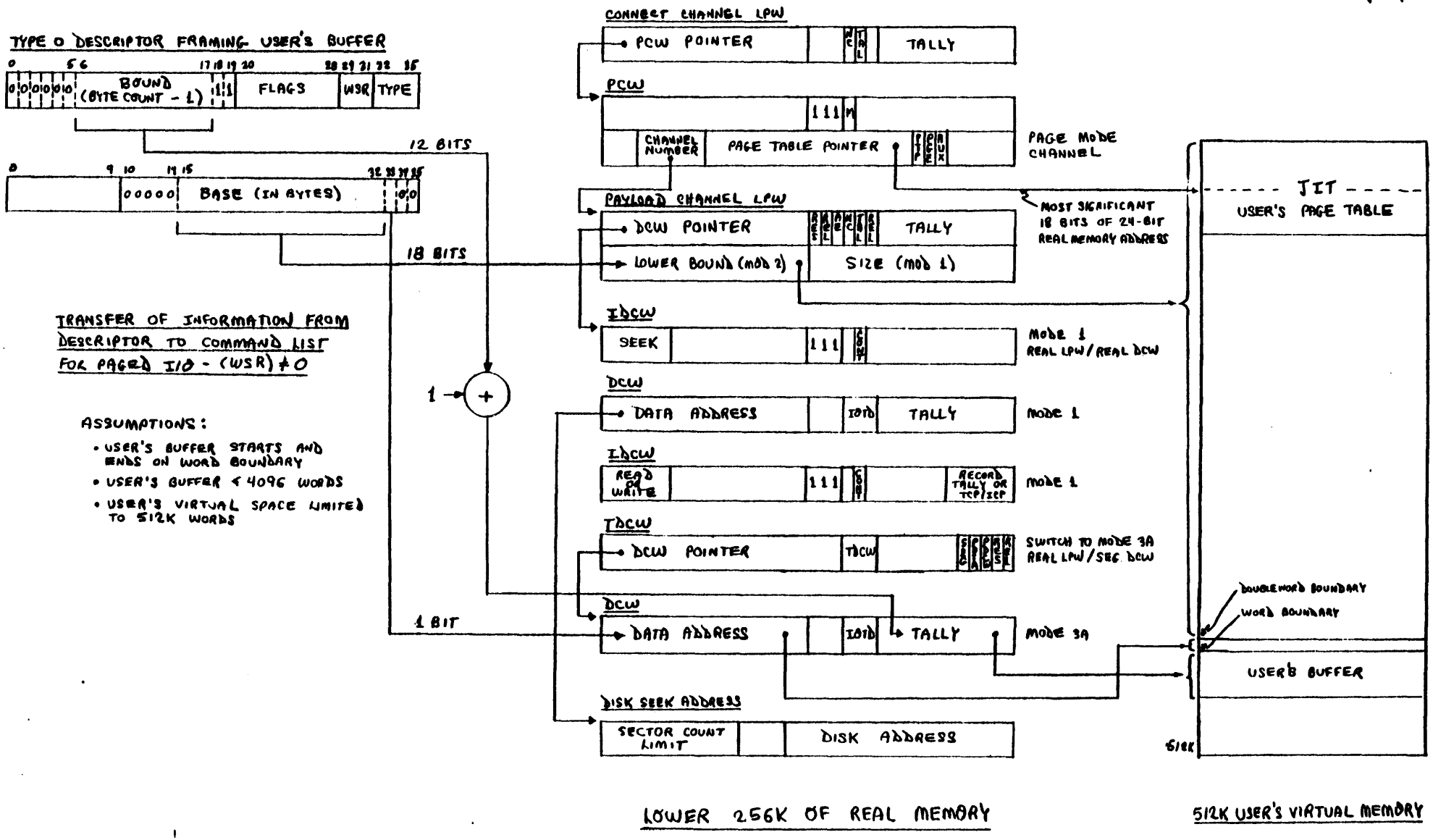
TYPE 0 DESCRIPTOR FRAMING USER'S BUFFER
BOUND (BYTE COUNT - 1)
FLAGS
WSR
TYPE
12 BITS
BASE (IN BYTES)
18 BITS
1 BIT
TRANSFER OF INFORMATION FROM DESCRIPTOR TO COMMAND LIST FOR PAGED I/O - (WSR) ≠ 0
ASSUMPTIONS:
• USER'S BUFFER STARTS AND ENDS ON WORD BOUNDARY
• USER'S BUFFER < 4096 WORDS
• USER'S VIRTUAL SPACE LIMITED TO 512K WORDS
CONNECT CHANNEL LPW
PCW POINTER
TALLY
PCW
CHANNEL NUMBER
PAGE TABLE POINTER
PAGE MODE CHANNEL
PAYLOAD CHANNEL LPW
DCW POINTER
TALLY
LOWER BOUND (MOD 2)
SIZE (MOD 1)
MOST SIGNIFICANT 18 BITS OF 24-BIT REAL MEMORY ADDRESS
IDCW
SEEK
MODE 1 REAL LPW/REAL DCW
DCW
DATA ADDRESS
MODE 1
IDCW
READ OR WRITE
RECORD TALLY OR TCP/ICP
MODE 1
TDCW
DCW POINTER
SWITCH TO MODE 3A REAL LPW/SEC DCW
DCW
DATA ADDRESS
TALLY
MODE 3A
DISK SEEK ADDRESS
SECTOR COUNT LIMIT
DISK ADDRESS
JIT
USER'S PAGE TABLE
DOUBLEWORD BOUNDARY
WORD BOUNDARY
USER'S BUFFER
SIZE
LOWER 256K OF REAL MEMORY
512K USER'S VIRTUAL MEMORY

Figure 5.0

channel. The page table pointer is the most significant 18 bits of the 24-bit real address of the beginning of the user's page table. The page table must be located on a 64-word boundary. The start of the user's page table is found by using the contents of the working space register addressed by the descriptor WSR field as an index into the working space page table directory. For paged I/O, the PTP and PGE flags in the PCW (bits 63-64) must both be coded as one's.

The first two words of the payload channel mailbox locations contain the payload channel LPW. The second word of this LPW contains an 18-bit lower bound field and an 18-bit size field. The lower bound field is a doubleword address and is used in CP-6 as the primary address for locating the user's buffer within the user's 512K virtual memory area. Bits 15 through 32 of the second word of the descriptor are placed in the lower bound field thereby addressing the doubleword in which the user's buffer starts. The lower bound field is not effective until a TDCW places the channel in mode 3A as described later. The size field in the second word of the payload channel LPW will not play an important role in CP-6, however, it must be initialized with a value that will not interfere with normal operation. The size field should be initialized with the smaller of 256K or 512K minus the lower bound address.

As shown in Figure 5.0, the remainder of the base and bound data from the descriptor goes into the final DCW in the command string. This is the DCW that actually transfers data to or from the user's buffer. Bit 33 of the second word of the descriptor goes into the 18-bit data address field of the final DCW. For CP-6 paged I/O, this field will be used solely to address the word boundary on which the user's buffer starts within the doubleword addressed by the lower bound field of the payload channel LPW. The remaining bits of the base field in the descriptor will be zero under the assumptions previously stated.

Bits 6-17 of the first word of the descriptor are incremented by one and then placed in the tally field of the final DCW. This field is the number of words in the user's buffer. Bits 0-5 of the first word of the descriptor will be zero if the size of the user's buffer is less than or equal to 4096 words. If the user's buffer starts and ends on a word boundary, bits 18-19 of the first word of the descriptor will both be one's.

The remainder of the command list in Figure 5.0 is largely self-explanatory. The DCW pointer in the payload channel LPW points at the beginning of the control word list. For

I/O to a disk, the first control word is a seek IDCW followed by a DCW pointing at the seek address. For CP-6 the entire control word list and the seek address will be located in the fixed data area of the monitor in the lower 256K words of real memory. The seek IDCW is chained to a read or write IDCW which is followed by a TDCW with the REL flag (bit 35) set to one. This TDCW switches the channel to mode 3A which means that the data address in all subsequent DCW's is added to the lower bound field and the result is mapped throught the user's page table for all subsequent data accesses. In mode 3A the LPW addressing remains real so the final DCW pointed to by the DCW pointer in the TDCW is also in the lower 256K of real memory.

5.1 User's Buffer Not Aligned on Word Boundary

It appears that the capability to have physical I/O buffers that are not on a word boundary will be available only for writing records on magnetic tape. In other words, hardware limitations restrict all records to and from disk or unit record and records read from magnetic tape to both start and end on a word boundary. For read operations from magnetic tape it appears that the software can determine the actual number of bytes or characters read from the magnetic tape by examining the status words returned by the IOM at the completion of the operation.

For writing records on magnetic tape, bits 30-35 of the write IDCW contain a field called "Record Tally or TCP/ICP". Three bits of this field specify the initial character positon and three bits specify the terminate character position. For write operations to magnetic tape, the TCP/ICP field of the write IDCW will be initialized using bits 18-19 of the first word of the descriptor and bits 34-35 of the second word of the descriptor.

5.2 User's Buffer Larger Than 4096 Words

Since the tally field in a DCW occupies 12 bits, each DCW can transfer no more than 4096 words. If the user's buffer is larger than 4096 words, bits 0-5 of the first word of the descriptor will be non-zero. In this case, multiple DCW's must be used to transfer the data. All DCW's except the last will be coded in bits 21-23 as IOTP's and the last DCW will be coded as an IOTD. The DCW list need not be contiguous; one or more TDCW's may be inserted as long as the TDCW PDCW flag (bit 33) is set to zero thereby allowing the channel to remain in mode 3A.

5.3 User's Virtual Space Larger Than 512K Words

If we wished to allow user's virtual space to expand beyond 512K words on some later release of CP-6, the proposed I/O addressing scheme could still be used with a minor increase in complexity. It is assumed that no segment within the user's working space quarter would extend over a 512K word boundary. In this case the page table pointer in the PCW must be initialized to point to the beginning of a 512 word portion of the user's page table containing the user's buffer. The lower bound field in the payload channel LPW will still be loaded from bits 15-32 of the second word of the descriptor which will be the doubleword address of the beginning of the user's buffer relative not to the beginning of the user's virtual memory but to the 512K word portion of the user's virtual memory containing the user's buffer.

6.0 NON-PAGED I/O

The proposed use of IOM facilities for non-paged I/O is shown in Figure 6.0. Non-paged I/O may occur for transfers such as file read-ahead where it may be more convenient for the requestor (file management) to provide the real address of the memory buffer rather than a virtual address.

The following assumptions are made for the purpose of simplifying Figure 6.0 and the description of non-paged I/O:

- Real memory buffer starts and ends on a word boundary
- Real memory buffer <= 4096 words.

The comments under paged I/O in sections 5.1 and 5.2 relative to these assumptions apply equally well to non-paged I/O.

For non-paged I/O the channel will be operated in the extended channel mode. This is accomplished coding a zero in the PGE flag (bit 64) of the PCW. The page table pointer field in the PCW and the lower bound and size fields in the payload channel LPW are not used.

As shown in Figure 6.0, bits 16-33 of the second word of the descriptor are placed in the data address field of the final DCW. Bits 10-15 of the second word of the descriptor are placed in the address extension field of the read or write IDCW. This address extension field addresses a 256K word block within the 2**24 word real memory. The data

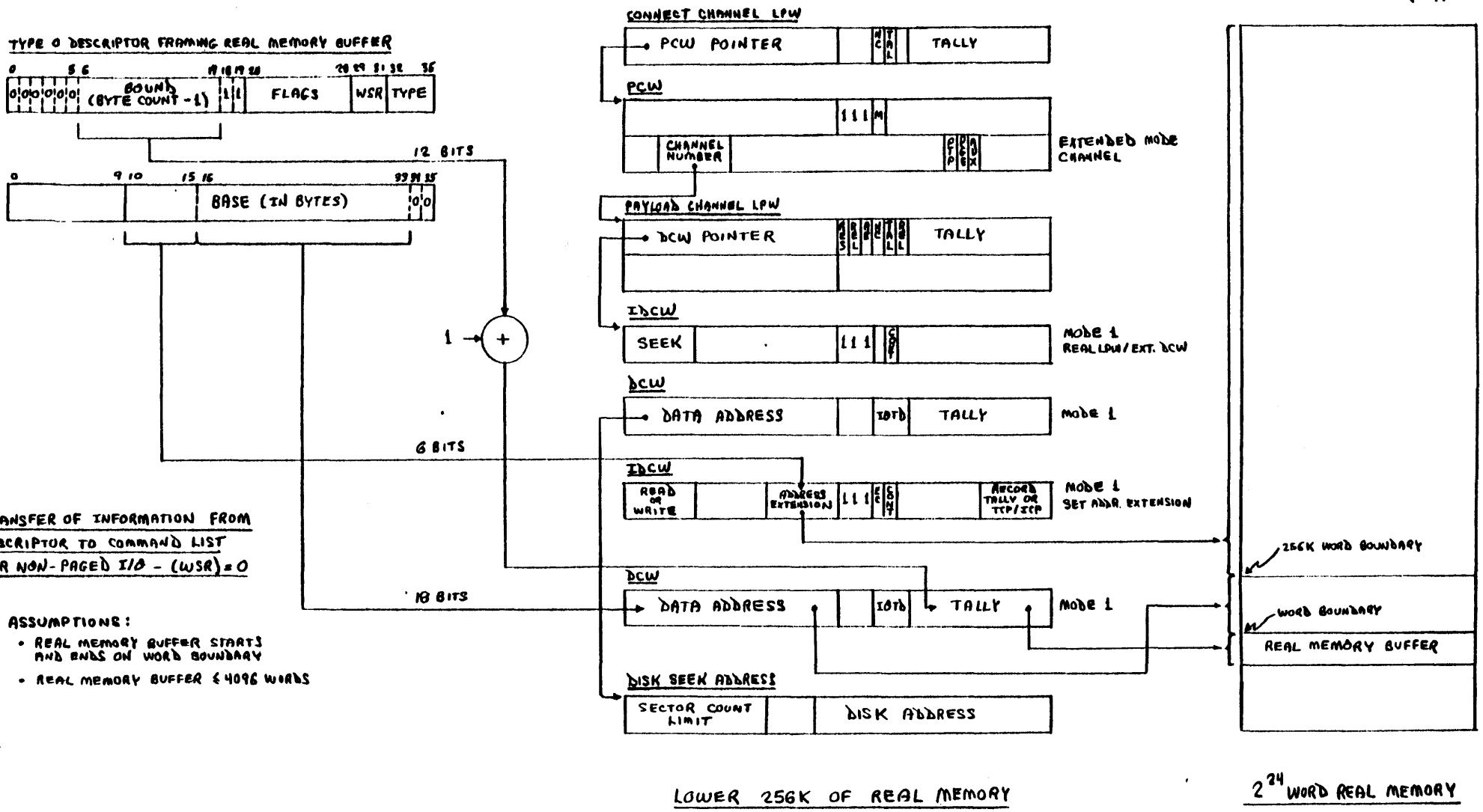
TYPE 0 DESCRIPTOR FRAMING REAL MEMORY BUFFER
BOUND (BYTE COUNT - 1)
FLAGS
WSR
TYPE
12 BITS
BASE (IN BYTES)
1
6 BITS
18 BITS
TRANSFER OF INFORMATION FROM DESCRIPTOR TO COMMAND LIST FOR NON-PAGED I/O - (WSR) = 0
ASSUMPTIONS:
• REAL MEMORY BUFFER STARTS AND ENDS ON WORD BOUNDARY
• REAL MEMORY BUFFER ≤ 4096 WORDS
CONNECT CHANNEL LPW
PCW POINTER
TALLY
PCW
CHANNEL NUMBER
EXTENDED MODE CHANNEL
PAYLOAD CHANNEL LPW
DCW POINTER
TALLY
IDCW
SEEK
MODE 1
REAL LPW / EXT. DCW
DCW
DATA ADDRESS
TALLY
MODE 1
IDCW
READ OR WRITE
ADDRESS EXTENSION
RECORD TALLY OR TCP/ICP
MODE 1
SET ADDR. EXTENSION
DCW
DATA ADDRESS
TALLY
MODE 1
DISK SEEK ADDRESS
SECTOR COUNT LIMIT
DISK ADDRESS
256K WORD BOUNDARY
WORD BOUNDARY
REAL MEMORY BUFFER
LOWER 256K OF REAL MEMORY
2^24 WORD REAL MEMORY

Figure 6.0

address within the final DCW provides the word address of the real memory buffer within the 256K word block specified by the address extension register. Bits 6-17 of the first word of the descriptor incremented by one are placed in the tally field of the final DCW and indicate the size (in words) of the real memory buffer.