X3T9.2/87-203

To: X3T9.2 Membership

From: Gregory G. Floryance
IBM Corporation
System Products Division
Hwy 52 & 37th Street NW
Rochester, Mn 55901
(507) 253-1869

Date: 12/06/87

Subject: LOAD SKIP MASK COMMAND

Their is a class of systems in which data storage (DASD) is managed in a non-contiguous fashion. In this environment, the scattering of data into seemingly random logical block addresses produces a result in which only a subset of the Logical Blocks over a given range are of interest to the system. The following figure and discussion attempt to illustrate the effects on system performance.

LBA ON MEDIUM	1	2	3	4	5	6	7	8
LBA OF INTEREST	X					X		X

The figure shown here has eight logical blocks numbered one thru eight in which only blocks 1, 6 and 8 are of interest to the initiator. Currently, SCSI provides the initiator with 2 alternate methods of acquiring the data. The first requires that the initiator issue three separate read commands (1 for each block). The overhead within the initiator, target and on the SCSI bus itself to transfer the extra two commands prevents them from doing other tasks and adversely affect performance. This may even lead to more disastrous performance consequences if additional revolutions of the medium occur. In addition, the overhead to manage a single request for data (made by an application program) and break it into three separate commands also uses up operating system resources which can be better used servicing other tasks.

The second method currently offered by SCSI is for the initiator to read all eight blocks (even though there is no interest in blocks 2-5,7) in order to insert blocks 1, 6 and 8 into the data stream on the subsequent write operation. This second method has a problem of wasting system memory to buffer the unwanted blocks as well as the additional problem of using up SCSI bus bandwidth in transferring those unnecessary blocks.

These performance problems resulting from the only currently available SCSI implementation alternatives are greatly compounded in a heavily utilized system in which system resources, including peripheral devices, the SCSI bus, system memory and system processing power are heavily loaded.

This proposal defines a function similar to that found in the INTELLIGENT PERIPHERAL INTERFACE (IPI) in which a skip mask is transferred from the initiator to the target which alters the action of the Read or Write command such that data blocks are selectively transferred under mask control. Unlike IPI, which supports the mask as a 32 byte (max) parameter of the Request Packet, this SCSI proposal extends the length (up to 256 bytes). A 256 byte map provides a 2K bit mask which in a typical 512 byte block mode provides addressability to 1 Mbyte of data. The second difference to IPI is that SCSI can not support two data transfers (the mask and the read or write data) for a single command. Therefore, the mask transfer must be accomplished with a new set of commands which are found in the following proposal.

87

7.1.7. LOAD SKIP MASK (Group 2) Commands

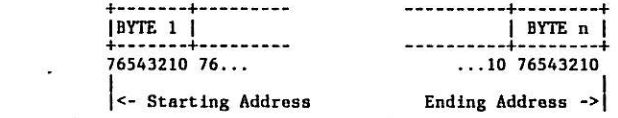
Peripheral Device Type: All
Operation Code Type: Optional

Table 8-??: LOAD SKIP MASK (Group 2) Command

Bit / Byte	7	6	5	4	3	2	1	0
0	Operation Code (58h,5Ah)							
1	Logical Unit Number			DPO	FUA	Reserved		RelAdr
2	(MSB)							
- - -	- -			Logical Block Address				- -
5								(LSB)
6	Parameter List Length (ie Mask Length)							
7	(MSB)							
8	---			Transfer Length				(LSB)
9	Control Byte							

The LOAD SKIP MASK commands (Table 8-??) transfer a skip mask during the Data Out phase to the Logical Unit to be used on a following linked Read (08h,28h) or Write (0Ah,2Ah) command. The Skip Mask alters the action of the Read or Write command such that data blocks are selectively transferred under mask control.

Each bit within the mask corresponds to a sequentially addressed data block within a range of logical block addresses that start from the Logical Block Address specified in bytes 2-5 of the CDB. The range extends to include the block at (Logical Block Address + (8 x mask length in bytes) - 1). Data blocks are transferred if their corresponding mask bit is '1', and are skipped if the bit value is '0'. Bit 7 of the first byte of the mask corresponds to the starting block address in the Command Extent parameter. Succeeding lower order Mask bits correspond to numerically higher valued addresses as illustrated in the following figure.



The LOAD SKIP MASK commands are intended to improve performance by reducing the number of commands and/or the amount of data transferred over the SCSI interface. Additional performance gains may also be realized since the SKIP MASK COMMANDS also provide sufficient information to begin the read or write operation without waiting for a subsequent linked READ or WRITE command to obtain the target logical block address.

The Operation Code field of 58h indicates that the following linked command shall be a READ (08h, 28h) command. An Operation Code field of 5Ah indicates that the following linked command shall be a WRITE (0Ah, 2Ah) command. A CHECK CONDITION status shall be returned on the subsequent linked command and the link shall be broken if the subsequent linked command following OpCode 58h is not a READ or following OpCode 5Ah is not a WRITE.

The Logical Block Address field specifies the logical block at which the read or write operation shall begin. This is provided to allow a seek or similar operation to be started immediately. The Logical Block Address for both the SKIP MASK and the subsequent linked READ or WRITE command shall be the same or the subsequent linked command shall end with CHECK CONDITION status and the link shall be broken.

The Parameter List Length field specifies the number of bytes in the skip mask which will be transferred in the subsequent data out phase. A value of zero shall be interpreted as indicating that 256 bytes shall be transferred.

The Transfer Length field specifies the number of logical blocks of data that shall be transferred. A transfer length of zero indicates that no logical blocks shall be transferred. This condition shall not be considered an error. Any other value indicates the number of logical blocks that shall be transferred. Although the Transfer Length field value shall be equal to the number of '1' bits in the Skip Mask, for performance reasons, the target is not required to verify this count. The Transfer Length for both the SKIP MASK and the subsequent linked READ or WRITE command shall be the same or the subsequent linked command shall end with CHECK CONDITION status and the link shall be broken.

Any unused bits in the last byte of the skip mask shall be set to zeros.

The sum of the Data Address and the bit string length from mask start to the last '1' bit in the Skip Mask parameter must not exceed the number of blocks in the addressable area. If the sum is greater than the number of blocks in the addressable area, CHECK CONDITION status shall be returned and the link broken.

See section 6.2.4 for a description of the cache control bits (DPO and FUA).
See section 6.2.5 for a description of the relative address bit (RelAdr).

This command shall be terminated with a RESERVATION CONFLICT status if any reservation access conflict (see section 8.1.16) exists.

If any of the following conditions occur, this command shall return a CHECK CONDITION status and the sense key shall be set as indicated in the following table. This table does not provide an exhaustive enumeration of all conditions that may cause the CHECK CONDITION status.

Condition	Sense Key
Invalid logical block address	ILLEGAL REQUEST (see note 1)
Target reset or medium change since last command from this initiator	UNIT ATTENTION

NOTES:

(1) The sense information bytes shall be set to the logical block address of the first invalid address.