Date: October 03, 2005

To: T10 Committee (SCSI)

From: George Penokie (IBM/Tivoli)

Subject: SBC-3: SPC-4: Disabling Reassign on Write Long Logical Blocks

1 Overview

Some SCSI/SAS hosts controllers make use of the Write Long command to intentionally create unrecoverable errors on the media. Currently there is no way to differentiate these pseudo uncorrectable errors (intentionally created) with a typical unrecoverable error caused by media defects. There is also a problem as these errors are included in the algorithm used for Information Exception Conditions and, as a result, may cause information exception condition trips, as well as unnecessary reassignments.

The solution being proposed is to add a bit to the Write Long command that adds a bit to the WRITE LONG command that would indicate to a logical unit to not count the logical block being written in any IEC or automatic reassignment calculations. Also, a bit is proposed to be added to the extended inquiry data VPD page to indicate if a logical unit supports this new feature.

2 SBC-3 changes

2.1 WRITE LONG (10) command

The WRITE LONG (10) command (see table 1) requests that the device server transfer data for a single logical block from the data-out buffer and write it to the medium. The data written shall be the same length and shall be in the same order as the data returned by the READ LONG (10) command (see 5.14). The device server shall write the logical block to the medium, and shall not return GOOD status until the logical block has actually been written on the medium.

Table 1 — WRITE LONG (10) command

Byte\Bit	7	6	5	4	3	2	1	0
0	OPERATION CODE (3Fh)							
1	Reserved							Obsolete
2	(MSB)			LOGICAL BLOCK ADDRESS				
5								(LSB)
6	Reserved							COR_DIS
7	(MSB)			BYTE TRANSFER LENGTH				
8								(LSB)
9	CONTROL							

See the PRE-FETCH (10) command (see 5.3) for the definition of the LOGICAL BLOCK ADDRESS field.

A correction disabled (COR_DIS) bit set to zero specifies when the indicated logical block is read that the device server shall perform normal error recovery on that logical block. A COR_DIS bit set to one specifies when the indicated logical block is read that the device server shall:

  a) perform no error recovery on that logical block:
  b) perform no automatic reallocation of that logical block;
  c) not use the errors indicated on a read of that logical block to control the reporting and the operations of specific informational exception conditions as defined by the INFORMATION EXCEPTIONS CONDITIONS mode page (see SPC-4): and

d) return CHECK CONDITION status with the sense key set to MEDIUM ERROR and the additional sense code set to READ ERROR - LBA MARKED BAD BY APPLICATION CLIENT.

The condition established by the COR_DIS bit being set to one shall remain in effect until the logical block is written by any means (e.g., any WRITE command, WRITE SAME command, or FORMAT command), other than another WRITE LONG command to the same logical block with the COR_DIS bit set to one.

The BYTE TRANSFER LENGTH field specifies the number of bytes of data that the device server shall transfer from the data-out buffer and write to the specified logical block. If the BYTE TRANSFER LENGTH field is not set to zero and does not match the data length that the device server returns for a READ LONG command, then the device server shall terminate the command with CHECK CONDITION status with the sense key set to ILLEGAL REQUEST and the additional sense code set to INVALID FIELD IN CDB. In the sense data (see 4.13 and SPC-4), the ILI and VALID bits shall be set to one and the INFORMATION field shall be set to the difference (i.e., residue) of the requested length minus the actual length in bytes. Negative values shall be indicated by two's complement notation. A BYTE TRANSFER LENGTH field set to zero specifies that no bytes shall be written. This condition shall not be considered an error.

2.2 WRITE LONG (16) command

The WRITE LONG (16) command (see table 2) requests that the device server transfer data for a single logical block from the data-out buffer and write it to the medium. The data written shall be the same length and shall be in the same order as the data returned by the READ LONG (16) command (see 5.15). The device server shall write the logical block to the medium, and shall not return GOOD status until the logical block has actually been written on the medium. This command is implemented as a service action of the SERVICE ACTION OUT operation code (see A.2).

Table 2 — WRITE LONG (16) command

Byte\Bit	7	6	5	4	3	2	1	0
0	OPERATION CODE (9Fh)							
1	Reserved			SERVICE ACTION (11h)				
2	(MSB)			LOGICAL BLOCK ADDRESS				
9								(LSB)
10				Reserved				
11								
12	(MSB)			BYTE TRANSFER LENGTH				
13								(LSB)
14	Reserved							COR_DIS
15	CONTROL							

See the WRITE LONG (10) command (see 2.1) for the definitions of the fields in this command.

3 SPC-4 changes

3.1 New ASC/ASCQ

11h 14h READ ERROR - LBA MARKED BAD BY APPLICATION CLIENT

3.1.1 Extended INQUIRY Data VPD page

The Extended INQUIRY Data VPD page (see table 3) provides the application client with a means to obtain information about the logical unit.

Table 3 — Extended INQUIRY Data VPD page

Bit Byte	7	6	5	4	3	2	1	0
0	PERIPHERAL QUALIFIER			PERIPHERAL DEVICE TYPE				
1	PAGE CODE (86h)							
2	Reserved							
3	PAGE LENGTH (3Ch)							
4	Reserved				RTO	GRD_CHK	APP_CHK	REF_CHK
5	Reserved			GROUP_SUP	PRIOR_SUP	HEADSUP	ORDSUP	SIMPSUP
6	DCORR_SUP	Reserved					NV_SUP	V_SUP
7	Reserved							
63								

The PERIPHERAL QUALIFIER field and the PERIPHERAL DEVICE TYPE field are as defined in 6.4.2.

The PAGE LENGTH field specifies the length of the following VPD page data and shall be set to 60. The relationship between the PAGE LENGTH field and the CDB ALLOCATION LENGTH field is defined in 4.3.4.6.

....

A disable correction supported (DCORR_SUP) bit set to zero indicates that the device server does not support application clients disabling read error checking on a logical block written using a WRITE LONG command (see SBC-3). A DCORR_SUP bit set to one indicates that the device server supports application client disabling read error checking on a logical block written using a WRITE LONG command.