

Date: Oct. 29, 1999

To: T10 Committee (SCSI)

From: George Penokie (IBM)

Subject: Beyond 2TBytes

Overview

Subsystems connected to parallel SCSI and Fibre Channel are rapidly approaching sizes that will require SCSI commands that will address more than 2 TBytes of data. Many of the SCSI commands defined today for direct-access type SCSI devices are limited to addressing 2 TBytes when the block size is set to 512 bytes. Already some UNIX operating systems support an 8 byte address space, so where possible, this proposal will modify the LBA fields to 8 bytes.

This proposal will only address the direct-access type SCSI device command set.

Proposed changes

This proposal would make 16 byte commands out of any CDB that contains an LBA field. Those LBA fields would be made into 8 byte fields with the format of the CDB as shown in table 1.

Table 1 -Typical CDB for large LBA 16-byte commands

| Bit Byte | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
|-------------|-----------------------|---|---|-------------------------------------|---|---|---|-------|
| 0 | OPERATION CODE | | | | | | | |
| 1 | Reserved | | | MISC. CDB INFORMATION | | | | |
| 2 | (MSB) | | | | | | | |
| 3 | | | | | | | | |
| 4 | | | | | | | | |
| 5 | | | | | | | | |
| 6 | LOGICAL BLOCK ADDRESS | | | | | | | |
| 7 | | | | | | | | |
| 8 | | | | | | | | |
| 9 | | | | | | | | (LSB) |
| 10 | (MSB) | | | | | | | |
| 11 | | | | TRANSFER LENGTH (if required) | | | | |
| 12 | | | | PARAMETER LIST LENGTH (if required) | | | | |
| 13 | | | | ALLOCATION LENGTH (if required) | | | | |
| 14 | | | | Reserved | | | | |
| 15 | | | | CONTROL | | | | |

The commands that would use the above format are listed in table 2.

Table 2 -CDB and parameter list changes

| Command Name | Op code | Type | Standard | Comment |
|-----------------------|---------|------|----------|---|
| EXTENDED COPY | 83h | O | SPC-2 | Already in SPC-2 |
| FORMAT UNIT | 04h | M | SBC-2 | CDB OK - Need new Defect List Format (see table 6) |
| LOCK-UNLOCK CACHE(16) | | O | SBC-2 | Use format from table 1 for extended LBA and number of blocks. |
| PRE-FETCH(16) | | O | SBC-2 | Use format from table 1 for extended LBA and transfer length. |
| READ(16) | | O | SBC-2 | Use format from table 1 for extended LBA and transfer length. |
| READ CAPACITY(16) | | O | SBC-2 | Use format from table 1 for extended LBA and use table 3 for the parameter data's extended LBA. |
| READ LONG (16) | | O | SBC-2 | Use format from table 1 for extended LBA and byte transfer length. |
| REASSIGN BLOCKS | 07h | O | SBC-2 | CDB OK - New new option for defect list to add in 8-byte LBAs (see table 6). |
| REBUILD | 81h | O | SBC-2 | No room for larger LBA in CDB - No proposed change. |
| REGENERATE | 82h | O | SBC-2 | No room for larger LBA in CDB - No proposed change. |
| SET LIMITS(16) | | O | SBC-2 | Use format from table 1 for extended LBA and number of blocks. |
| SYNCHRONIZE CACHE(16) | | O | SBC-2 | Use format from table 1 for extended LBA and number of blocks. |
| VERIFY(16) | | O | SBC-2 | Use format from table 1 for extended LBA and verification length. |
| WRITE(16) | | O | SBC-2 | Use format from table 1 for extended LBA and transfer length. |
| WRITE AND VERIFY(16) | | O | SBC-2 | Use format from table 1 for extended LBA and transfer length. |
| WRITE LONG(16) | | O | SBC-2 | Use format from table 1 for extended LBA and transfer length. |
| WRITE SAME(16) | | O | SBC-2 | Use format from table 1 for extended LBA and number of blocks. |
| XDREAD(16) | | O | SBC-2 | Use format from table 1 for extended LBA and transfer length. |
| XDWRITE(16) | | O | SBC-2 | Use format from table 1 for extended LBA and transfer length. |
| XDWRITE EXTENDED | 80h | O | SBC-2 | No room for larger LBA in CDB - No proposed change. |
| XPWRITE(16) | | O | SBC-2 | Use format from table 1 for extended LBA and transfer length. |

In addition to the commands and parameters listed above the mode page header is another area where the LBA for direct-access SCSI device has only an eight byte field. I do not propose changing this as the capacity can be determined by the read capacity command.

Additional Read Capacity changes

The following statement needs to be added to the current Read Capacity command to cover the case were a target receives a Read Capacity command and the values that would be returned are too large to fit in the RETURNED LOGICAL BLOCK ADDRESS field of the Read Capacity data parameter list.

If the number of logical blocks exceeds the maximum value that may be specified in the RETURNED LOGICAL BLOCK ADDRESS field the device server shall transfer no data and return a CHECK CONDITION status and the sense key shall be set to ILLEGAL REQUEST and the additional sense code set to INVALID FIELD IN CDB.

The format for the Read Capacity (16) data parameter list is:

Table 3 -Read Capacity data

| Bit Byte | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | |
|----------|--------------------------------|---|---|---|---|---|---|-------|--|
| 0 | (MSB) _____ | | | | | | | | |
| | RETURNED LOGICAL BLOCK ADDRESS | | | | | | | | |
| 7 | | | | | | | | (LSB) | |
| 8 | (MSB) _____ | | | | | | | | |
| | BLOCK LENGTH IN BYTES | | | | | | | | |
| 11 | | | | | | | | (LSB) | |

Additional FORMAT UNIT defect descriptor

An additional Format unit defect descriptor will have to be added to allow returning a block format defect desiccator that can return the larger LBAs. The following additions will be needed to the FORMAT UNIT defect descriptor format and requirements table.

Table 4 -FORMAT UNIT defect descriptor format and requirements

| FMTDATA | CMPLST | Defect List Format | Defect List Length | Type | Comments |
|---------------|--------|--------------------|--------------------|------|-----------------------|
| Block Formats | | | | | |
| 1 | 0 | 011b | >0 | 0 | See notes (2) and (3) |
| 1 | 1 | 011b | >0 | 0 | See notes (2) and (4) |

The following FORMAT UNIT text would be added:

Each block format defect descriptor format specified as 000b (see table 5) specifies a four-byte defective block address that contains the defect. Each block format defect descriptor format specified as 110b (see table 6) specifies an eight-byte defective block address that contains the defect. Use of the Block format is vendor-specific.

Table 5 -DEFECT DESCRIPTOR - Block format (000b)

| Bit Byte | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | |
|----------|-------------------------|---|---|---|---|---|---|-------|--|
| 0 | (MSB) _____ | | | | | | | | |
| | DEFECTIVE BLOCK ADDRESS | | | | | | | | |
| 3 | | | | | | | | (LSB) | |

Table 6 -DEFECT DESCRIPTOR - Block format (110b)

| Bit Byte | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | |
|-------------|-------|---|---|---|---|---|---|-------------------------|-------|
| 0 | (MSB) | | | | | | | DEFECTIVE BLOCK ADDRESS | (LSB) |
| 7 | | | | | | | | | |