To: T10 Technical Committee
From: Rob Elliott, HP (elliott@hp.com)
Date: 6 November 2003
Subject: T10/03-335r1 SBC-2 SPC-3 Obsolete block device linked commands

Revision History

Revision 0 (25 September 2003) first revision
Revision 1 (6 November 2003) incorporated comments from November CAP WG.

Related Documents

99-259r4 - 2 Terabyte Changes
sbc2r10 - SCSI Block Commands - 2 revision 10

Overview

An 8-byte LBA version of the SET LIMITS command, only used for linked commands, does not exist in SBC-2. This means linked commands are not really compatible with 8-byte LBAs.

Features related to linked commands can be made obsolete in SBC-2, leaving SBC-1 as the last reference to them.

This includes:

- obsolete SET LIMITS
- obsolete the RELADR bit in LOCK UNLOCK CACHE, PRE-FETCH, READ, READ CAPACITY, READ LONG, SYNCHRONIZE CACHE, VERIFY, WRITE, WRITE AND VERIFY, WRITE LONG, WRITE SAME, ERASE, READ GENERATION, READ UPDATED BLOCK, and UPDATE BLOCK
- change the RELADR bit to Reserved in any of the 8-byte LBA commands. There was never a SET LIMITS command with an 8-byte LBA to support them.
- in 4.2.3.5.2 remove recommendation to use linked commands to avoid a buffer full
- in note 21 remove recommendation to link WRITE and VERIFY commands
- keep the descriptions of PREFETCH and MEDIUM SCAN returning INTERMEDIATE-CONDITION MET when the LINK bit is set. Until linked commands are completely obsoleted, these rules are still true. SBC-2 software could still link them - it just could not use relative addressing.

The SPC-3 INQUIRY bit reporting support for relative addressing should be made obsolete.

Suggested Changes to SPC-3

6.4.2 Standard INQUIRY data

~~A relative addressing (RELADR) bit set to one indicates that the device server supports the relative addressing mode (see SBC-2). If this bit is set to one, the linked command (LINKED) bit shall also be set to one; since relative addressing is only allowed with linked commands. A RELADR bit set to zero indicates the device server does not support relative addressing.~~

[Editor's note: change the RELADR bit in the CDB to "Obsolete"]

Annex C section C.3:

[Mark SET LIMITS (10) (33h) and SET LIMITS (12) (B3h) obsolete for D, W, and O columns. The command was mentioned in MMC-3 but is no longer mentioned in MMC-4, and Bill McFerrin has confirmed that the R column should also be marked obsolete.]

Suggested Changes to SBC-2

~~4.2.1.13 Relative addressing~~

~~Relative addressing is a technique that may be useful in accessing structured data in a uniform manner. Relative addressing may be used when commands are linked.~~

~~The SET LIMITS command (see 5.2.22) is provided to define the limits of a linked chain of relative addressing commands. This protects against exceeding the specified set of blocks. The SET LIMITS command has no effect on any other initiator.~~

4.2.3 Model for XOR commands

4.2.3.5 Additional array subsystem considerations

4.2.3.5.2 Buffer full status handling

When the storage array controller sends an XDWRITE or REGENERATE command to a device, the device has an obligation to retain the resulting XOR data until the storage array controller issues a matching XDREAD command to retrieve the data. This locks up part or all (depending on the size of the device's buffer and the size of the XOR data block) of the device's buffer space. When all of the device's buffer is allocated for XOR data, it may not be able to accept new media access commands other than valid XDREAD commands and it may not be able to begin execution of commands that are already in the task set.

When the device is not able to accept a new command because there is not enough space in the buffer, the device shall terminate that command with a CHECK CONDITION status and the sense key shall be set to ILLEGAL REQUEST with the additional sense code set to BUFFER FULL.

When a storage array controller receives this status, it may issue any matching XDREAD commands needed to satisfy any previous XDWRITE or REGENERATE commands. This results in buffer space being freed for other commands. If it is a multi-initiator system and the storage array controller has no XDREAD commands to send, the storage array controller may assume the buffer space has been allocated to another initiator. The storage array controller may retry the command in the same manner that a command ending with TASK SET FULL status would be retried including not retrying the command too frequently.

~~The storage array controller may use command linking to avoid a buffer full condition. For example, a storage array controller supervised update write operation would consist of an XDWRITE command linked to an XDREAD command.~~

The bidirectional XDWRITEREAD command avoids the buffer full condition. The storage array controller may issue multiple XDWRITEREAD commands, since the device controls when it accepts more write data and provides read data.

The PREFETCH command and MEDIUM SCAN command mention the LINK bit as a reason to return INTERMEDIATE-CONDITION MET. This would have to remain until linked commands are gone from SAM and SPC.

Table 12 — Commands for direct-access block devices

[Move SET LIMITS (10) and (12) to the obsolete list in the table notes]

5.2.3 LOCK UNLOCK CACHE (10) command

...

See 4.2.1.9 for reservation requirements for this command. A LOCK bit of zero indicates that all logical blocks in the specified range that are currently locked into the cache memory shall be unlocked, but may not be removed. A LOCK bit of one indicates that any logical block in the specified range that is currently present in the cache memory shall be locked into cache memory. Only logical blocks that are already present in the cache memory are actually locked.

~~A relative address (RELADR) bit of zero indicates that the LOGICAL BLOCK ADDRESS field specifies the first logical block of the range of logical blocks for this command.~~

~~A RELADR bit of one indicates that the LOGICAL BLOCK ADDRESS field is a two's complement displacement. This negative or positive displacement shall be added to the logical block address~~

~~last accessed on the block device to form the LOGICAL BLOCK ADDRESS for this command. This feature is only available with linked commands. This feature also requires that a previous command in the linked group has accessed a block of data on the block device.~~

The NUMBER OF BLOCKS field specifies the total number of contiguous logical blocks within the range. A NUMBER OF BLOCKS field of zero indicates that all remaining logical blocks on the block device shall be within the range.

Multiple locks may be in effect from more than one application client. Locks from different application clients may overlap. An unlock of an overlapped area does not release the lock of another initiator.

5.2.5 PRE-FETCH (10) command

...

If the IMMED bit is zero and the specified logical blocks were successfully transferred to the cache memory, the device server shall return CONDITION MET status. If the LINK BIT (see SPC-3) is one, the device server shall return INTERMEDIATE-CONDITION MET status.

If IMMED is one, and the unlocked cache memory has sufficient capacity to accept all of the specified logical blocks, the device server shall return CONDITION MET status. If the LINK bit is one, and the unlocked cache memory has sufficient capacity to accept all of the specified logical blocks, the device server shall return INTERMEDIATE-CONDITION MET status.

If IMMED is one, and the unlocked cache memory does not have sufficient capacity to accept all of the specified logical blocks, the device server shall return GOOD status. The device server shall transfer to cache memory as many logical blocks that fit. If the LINK bit is one, the device server shall return INTERMEDIATE status.

~~5.2.22 SET LIMITS (10) command~~

~~The SET LIMITS (10) command (see table 56) defines the range where subsequent linked commands may operate. A second SET LIMITS command shall not be linked to a chain of commands if a SET LIMITS (10) command has already been issued in the chain. If a second SET LIMITS (10) command within a linked list of commands is detected, the command shall be rejected with CHECK CONDITION status and the sense key shall be set to DATA PROTECT with the appropriate additional sense code for the condition.~~

~~See 4.2.1.9 for reservation requirements for this command.~~

~~A read inhibit (RDINH) bit of zero indicates that read operations within the range are not inhibited. A read inhibit (RDINH) bit of one indicates that read operations within the range shall be inhibited.~~

~~A write inhibit (WRINH) bit of zero indicates that write operations within the range are not inhibited. A write inhibit (WRINH) bit of one indicates that write operations within the range shall be inhibited.~~

~~The LOGICAL BLOCK ADDRESS field specifies the starting address for the range.~~

~~The NUMBER OF BLOCKS field specifies the number of logical blocks within the range. A number of blocks of zero indicates that the range shall extend to the last logical block on the block device.~~

~~Any attempt to access outside of the restricted range or any attempt to perform an inhibited operation within the restricted range shall cause the function to not be performed. The command shall be terminated with CHECK CONDITION status and the sense key shall be set to DATA PROTECT with the appropriate additional sense code for the condition.~~

~~5.2.23 SET LIMITS (12) command~~

~~The SET LIMITS (12) command (see table 57) defines the range where subsequent linked commands may operate.~~

~~See 4.2.1.9 for reservation requirements for this command. See the SET LIMITS (10) command (5.2.22) for a description of the fields in this command.~~

5.2.34 WRITE AND VERIFY (10) command

The WRITE AND VERIFY (10) command (see table 69) requests that the device server write the data transferred from the application client to the medium and then verify that the data is correctly written. The data is only transferred once from the application client to the device server.

...

~~NOTE 21 - The WRITE AND VERIFY command specifically states that the data are not to be transferred twice (i.e., once for the write pass, and once for the verify pass) when performing a byte compare. If there is a need for two transfers to occur (e.g., to ensure the integrity of the path to the media), then the application client should issue a WRITE command with a LINK bit of one followed by a VERIFY command with a BYTCHK bit of one, transferring the same data on each command.~~

5.3.4 MEDIUM SCAN command

...

A LINK BIT of zero in the CONTROL byte (see SAM-3) indicates a non-linked command; if the scan is satisfied, the command shall be terminated with a CONDITION MET status. A REQUEST SENSE command may then be issued to determine the starting logical block address of the area that meets the request. If the scan is not satisfied and no error occurs, the command shall be terminated with GOOD status.

A LINK BIT of one in the CONTROL byte indicates that a command is linked to the MEDIUM SCAN command; if the search is satisfied, INTERMEDIATE-CONDITION MET status is returned and the next command is executed. ~~If the RELADR bit in the next command is one, the LOGICAL BLOCK ADDRESS of the next command is used as a displacement from the logical block address where the search was satisfied.~~ If a linked scan is not satisfied, the command is terminated with a CHECK CONDITION status. A REQUEST SENSE command may then be issued.

General SBC-2 changes (applies to many places in the standard)

Change the RELADR bit to “Obsolete” and delete the defining sentence in:

- LOCK UNLOCK CACHE (10)
- PRE-FETCH (10)
- READ (10)(12)
- READ CAPACITY (10)
- READ LONG
- SYNCHRONIZE CACHE (10)
- VERIFY (10)(12)
- WRITE (10)(12)
- WRITE AND VERIFY (10)(12)
- WRITE LONG
- WRITE SAME (10)
- ERASE (10)(12)
- MEDIUM SCAN
- READ GENERATION
- READ UPDATED BLOCK
- UPDATE BLOCK

Change the RELADR bit to “Reserved” and delete the defining sentence in:

- LOCK UNLOCK CACHE (16)
- PRE-FETCH (16)
- READ (16)
- READ CAPACITY (16)
- SYNCHRONIZE CACHE (16)
- VERIFY (16)
- WRITE (16)
- WRITE AND VERIFY (16)
- WRITE SAME (16)