

To: T10 Technical Committee
From: Rob Elliott, Compaq Computer Corporation (Robert.Elliott@compaq.com)
Date: 8 September 2000
Subject: SBC-2 issues

These are issues for SBC-2 based on SBC rev 08b and the accepted proposal 99-259r4 (Beyond 2 TB).

Revision 0: initial revision

Revision 1: added comments noting conclusions on each item from the July 2000 CAP WG

Revision 2: fixed a few typos

1. Support LUN addressing

[Ed Gardner has a proposal for large transport IDs that should also fix this problem]

The "3rd party" XOR commands (those where the target becomes a "temporary initiator") REBUILD, REGENERATE, and XDWRITE EXTENDED currently assume they only need to access LUN 0 of the specified target. In SBC-2, these should be expanded to support LUN addressing. This will allow them to talk to disks that may sit behind a SCSI LUN Bridge, for example. This will also facilitate adding Access Control proxy support.

2. Support full target identifiers

[Ed Gardner has a proposal for large transport IDs that should also fix this problem]

The SECONDARY ADDRESS field in the XDWRITE EXTENDED command is only one byte, with a TABLE ADDRESS bit provided to allow indexing into a lookup table of full-sized addresses. This table was not further defined in SBC. With a variable length CDB version of XDWRITE EXTENDED, the full address can be included in the CDB and the unused table reference concept can be dropped.

3. Support Access Control proxies

[once item #1 is fixed, this should easily follow]

REBUILD, REGENERATE, and XDWRITE EXTENDED should be expanded to support Access Control proxies. This will help the commands work in a large SAN environment. See 99-245r9 (Access Controls) Appendix D.2 for details on how this was done for the EXTENDED COPY command.

4. Increase defect list lengths

[July SCSI CAP WG questioned whether defect lists should be obsoleted or left as is (supporting only short lists, which may be all that is needed for testing) rather than fixed.

Disk drive qualification does use these lists, so they need to remain and be expanded. For example, the complete list is cleared and the bad blocks are accessed to test that the disk drive detects the error again.

]

Defect list lengths used in various commands should match that supported by REPORT DEFECT LIST(12) - a 4 byte list length. A shorter length makes it impossible to use the "complete list" bit

(CMPLST) with new devices and more difficult to use defect lists in general - you can't just feed the data from REPORT DEFECT LIST back into FORMAT UNIT. FORMAT UNIT, REASSIGN BLOCKS, SEND DIAGNOSTIC and RECEIVE DIAGNOSTIC RESULTS all may be affected.

5. Expand cylinder, head, and sector fields

[July SCSI CAP WG questioned whether any of these fields and associated pages could be made obsolete rather than fixed. These were suggested for obsolescence:

Page 03h – format device page

Page 04h - rigid disk geometry page

Page 05h – flexible disk page

It was also suggested that the order requirements for defects be removed.

Disk drive qualification does use all these pages, so they need to remain and be fixed.

]

Do the cylinder head or sector fields of various parameter lists and descriptors need to be expanded to go with the larger LBAs?

6. Expand Initialization Pattern size

[July SCSI CAP WG recommends using the least significant 4 bytes of the LBA]

Should the FORMAT UNIT Initialization Pattern options 01b and 10b write 8 bytes worth of the logical block address rather than just 4 bytes when long LBAs are being used? If only 4 bytes are being used, which 4 - the most or least significant portion?

7. Add the T10-approved definitions for "reserved", "may", and "may not"

[accepted as editorial by July SCSI CAP WG]

Current SBC definitions:

may: Indicates flexibility of choice with no implied preference.

reserved: Refers to bits, bytes, words, fields, and code values that are set aside for future standardization. Their use and interpretation may be specified by future extensions to this or other standards. A reserved bit, byte, word, or field shall be set to zero, or in accordance with a future extension to this standard. The recipient may not check reserved bits, bytes, words, or fields. Receipt of reserved code values in defined fields shall be treated as an error.

Definitions from SPC-2:

may: A keyword that indicates flexibility of choice with no implied preference (equivalent to "may or may not").

may not: A keyword that indicates flexibility of choice with no implied preference (equivalent to "may or may not").

reserved: A keyword referring to bits, bytes, words, fields and code values that are set aside for future standardization. A reserved bit, byte, word or field shall be set to zero, or in accordance with a future extension to this standard. Recipients are not required to check reserved bits, bytes, words or fields for zero values. Receipt of reserved code values in defined fields shall be reported as error.

8. Update Note 6 on migration to READ(10)

[accepted as editorial by July SCSI CAP WG]

Note 6 on page recommends "migrating to READ(10)." This should be extended to mention the new long LBA version READ(16).

Note 6 - Although the READ(6) command is limited to directly addressing logical blocks up to a capacity of 2 Giga-bytes, for logical block sizes of 512 bytes, this command has been maintained as mandatory since some system initialization routines require that the READ(6) command be used. Application clients should migrate from the READ(6) command to the READ(10) command which may address 2 Tera-bytes with logical block sizes of 512 bytes, *or the READ(16) command to address more than 2 Terabytes.*

9. Add guidelines for commands used by REBUILD, REGENERATE, and XDWRITE EXTENDED

[accepted as editorial by July SCSI CAP WG]

Third-party commands REBUILD and REGENERATE should have text indicating when they will use READ(16) vs. READ(10) vs. READ(6). The first paragraph could be extended in this manner:

The target, acting as a temporary initiator, issues READ commands to retrieve the specified data. *READ(10) should be used for accesses less than 2 Terabytes, and READ(16) should be used for accesses greater than or equal to 2 Terabytes.*

Similarly, the XDWRITE EXTENDED command should indicate when it uses XPWRITE(16) vs. XPWRITE(32). The first paragraph could be extended in this manner:

The target, acting as a temporary initiator, issues XPWRITE commands to retrieve the specified data. XPWRITE(16) should be used for access below 2 Terabytes, and XPWRITE(32) should be used for accesses greater than or equal to 2 Terabytes.

10. Other minor changes

[accepted as editorial by July SCSI CAP WG]

In two places in section 5.1.4 (pages 10 and 11), "power condition code 5h" should be "power condition code of 5h."

Two references in 7.1.1.1 (page 82) need to be clearer. Change:

"Valid values for this field are defined in the FORMAT UNIT command (see 6.1.1)."

To:

"Valid values for this field are defined in the DEFECT LIST FORMAT field of the FORMAT UNIT command (see 6.1.1)."

The SEND DIAGNOSTIC/RECEIVE DIAGNOSTIC RESULTS translate address pages each include this text that is no longer true for the new 8 byte logical block address format (pages 83 and 84):

If the logical block format is specified the block address shall be in the first four bytes of the field with the remaining bytes set to zero.