

To: T10 Technical Committee
 From: Rob Elliott, HP (elliott@hp.com)
 Date: 8 November 2007
 Subject: 07-447r1 SBC-3 Read-Write Error Recovery clarifications

Revision history

Revision 0 (24 October 2007) First revision

Revision 1 (8 November 2007) Incorporated comments from November 2007 CAP WG

Related documents

sbc3r11 - SCSI Block Commands - 3 (SBC-3) revision 11

07-451 SBC-3 SPC-4 WRITE LONG COR_DIS and WR_UNCOR interaction (Rob Elliott, HP)

Overview

The combined error recovery bit description table is missing two combinations - 1010b and 1011b (both are invalid combinations). These entries were lost from sbc2r08 to sbc2r09 during the conversion from Microsoft Word to Adobe FrameMaker.

Several other rewording changes are proposed, normalizing usage of “[un]recovered [read] errors”.

Suggested changes

3.1.xx pseudo read data: [Data that is transferred by the device server based on the setting of the RC bit or the TB bit in the Read-Write Error mode page \(see 6.3.5\) in order to maintain a continuous flow of data or to transfer the amount of data requested for a command even though an unrecovered read error \(see 3.1.xx\) occurred while the device server was processing the command. Psuedo read data may be erroneous or fabricated by the device server \(e.g., data already in a buffer or any other vendor-specific data\).](#)

3.1.xx unrecovered error: [An error in which the device server is unable to read or write a logical block within the recovery limits specified in the Read-Write Error Recovery mode page \(see 6.3.5\) and the Verify Error Recovery mode page \(see 6.3.6\).](#)

3.1.xx unrecovered read error: [An unrecovered error \(see 3.1.xx\) during a read operation.](#)

3.1.xx recovered error: [An error in which the device server is able to read or write a logical block within the recovery limits specified in the Read-Write Error Recovery mode page \(see 6.3.5\) and the Verify Error Recovery mode page \(see 6.3.6\).](#)

4.9 Medium defects

Any medium has the potential for defects that cause data to be lost. Therefore, each logical block may contain additional information that allows the detection of changes to the user data and protection information, if any, caused by defects in the medium or other phenomena, and may also allow the data to be reconstructed following the detection of such a change (e.g., ECC bytes).

Direct-access block devices may allow the application client to examine and modify the additional information by using the READ LONG commands and the WRITE LONG commands (see 5.16, 5.17, 5.35, and 5.36). The application client may use the WRITE LONG commands to induce a defect to test the defect detection logic of the direct-access block device or to emulate an **unrecoverable** logical block [with an unrecovered read error](#) when generating a mirror copy.

Direct-access block devices may allow the application client to disable error correction and automatic reallocation on specific logical blocks by using the WRITE LONG command (see 5.35 and 5.36). This allows an application client to prevent logical blocks from being included in the algorithm used for information exception conditions, there by, preventing unwarranted information exception condition trips and unnecessary reassignments.

...

4.14.1 Error reporting overview

If any of the conditions listed in table 4 occur during the processing of a command, the command shall be terminated with CHECK CONDITION status with the sense key set to the specified value and the additional sense code set to the appropriate value for the condition. Some errors may occur after the completion status has already been reported. For such errors, SPC-4 defines a deferred error reporting mechanism. Table 4 lists some error conditions and the applicable sense keys. The list does not provide an exhaustive enumeration of all conditions that may cause CHECK CONDITION status.

Table 1 — Example error conditions

Condition	Sense key
Invalid LBA	ILLEGAL REQUEST
Unsupported option requested	ILLEGAL REQUEST
Logical unit reset, I_T nexus loss, or medium change since last command from this application client	UNIT ATTENTION
Self diagnostic failed	HARDWARE ERROR
Unrecovered read error	MEDIUM ERROR or HARDWARE ERROR
Recovered read error	RECOVERED ERROR
Over-run or other error that might be resolved by repeating the command	ABORTED COMMAND
Attempt to write on write-protected medium	DATA PROTECT

...

When a recovered read error is reported, the INFORMATION field of the sense data shall contain the LBA of the last [logical block accessed by the command on which a recovered read error occurred during the transfer](#).

When an unrecovered ~~read~~ error is reported, the INFORMATION field of the sense data shall contain the LBA of the ~~unrecovered~~ logical block [on which the unrecovered error occurred](#).

4.15.3.3 Access to an inconsistent stripe

...

If any of the XOR commands end with CHECK CONDITION status and an unrecovered error is indicated, an inconsistent stripe may result. It is the storage array controller's responsibility to identify the failing device, the identify the scope of the failure, then limit access to the inconsistent stripe. The recovery procedures that the storage array controller implements are outside the scope of this standard.

If any of the XOR commands end with CHECK CONDITION status and an unrecovered [read](#) error is indicated, an inconsistent stripe may result. It is the storage array controller's responsibility to identify the failing device, the identify the scope of the failure, then limit access to the inconsistent stripe. The recovery procedures that the storage array controller implements are outside the scope of this standard.

5.18 REASSIGN BLOCKS

If the REASSIGN BLOCKS command failed due to an unexpected ~~unrecoverable~~[unrecovered](#) read error that would cause the loss of data in a logical block not specified in the defective LBA list, the LBA of the ~~unrecoverable~~ logical block [with the unrecovered read error](#) shall be returned in the INFORMATION field of the sense data and the VALID bit shall be set to one.

6.3.5 Read-Write Error Recovery mode page

The Read-Write Error Recovery mode page (see table 125) specifies the error recovery parameters the device server shall use during any command that performs a read or write operation to the medium (e.g., a READ commands, [a WRITE commands](#), ~~and/or a WRITE AND VERIFY commands~~) ~~except on logical blocks that have correction disabled (see 5.35 and 5.36)~~.

...

A transfer block (TB) bit set to zero specifies that if an unrecovered read error occurs during a read operation, then the device server shall not transfer a-any data for the logical block to the data-in buffer-if the logical block is not recovered within the recovery limits specified. A TB bit set to one specifies that if an unrecovered read error occurs during a read operation, then the device server shall transfer a-pseudo read data-logical block to the data-in buffer before returning CHECK CONDITION status-if the logical block is not recovered within the recovery limits specified. The data returned in this case is vendor-specific. The TB bit does not affect the action taken for ~~recovered data~~recovered read errors.

A read continuous (RC) bit set to zero specifies that error recovery operations that cause delays during the data transfer are acceptable-during the data transfer. Data shall not be fabricated.

An RC bit set to one specifies the device server shall transfer the entire requested length of data without adding delays during the data transfer to perform error recovery procedures. ~~This implies that the~~The device server may ~~send transfer pseudo read data that is erroneous or fabricated~~ in order to maintain a continuous flow of data. The device server shall assign priority to the RC bit over conflicting bits within this byte.

NOTE 24 - ~~Fabricated data may be data already in a buffer or any other vendor-specific data.~~The RC bit may be ~~used~~set to one in image processing, audio, or video applications.

An enable early recovery (EER) bit set to one specifies that the device server shall use the most expedient form of error recovery first. An EER bit set to zero specifies that the device server shall use an error recovery procedure that minimizes the risk of error mis-detection or mis-correction. This bit only applies to data error recovery and it does not affect positioning retries.

NOTE 25 - An EER bit set to one may imply an increase in the probability of error mis-detection or mis-correction. An EER bit set to zero allows the specified retry limit to be exhausted prior to using ~~error-correction codes~~additional information (e.g., ECC bytes) to correct the error.

A post error (PER) bit set to one specifies that ~~the device server shall report recovered errors if a recovered read error occurs during a command performing a read or write operation, then the device server shall terminate the command with CHECK CONDITION status with the sense key set to RECOVERED ERROR.~~ A PER bit set to zero specifies that ~~the device server shall not report recovered errors if a recovered read error occurs during a command performing a read or write operation, then the device server shall terminate the command with CHECK CONDITION status~~ and ~~the device server~~ shall perform error recovery procedures within the limits established by the error recovery parameters. If the DTE bit is set to one, then the PER bit shall be set to one.

A data terminate on error (DTE) bit set to one specifies that the device server shall terminate the data-in or data-out buffer transfer of a command performing a read or write operation upon detection of a recovered error. A DTE bit set to zero specifies that the device server shall not terminate the data-in or data-out buffer transfer of a command performing a read or write operation upon detection of a recovered error.

A disable correction (DCR) bit set to one specifies that ~~ECC (see 4.4)~~additional information (e.g., ECC bytes) (see 4.4) shall not be used for data error recovery. A DCR bit set to zero allows the use of ~~ECC~~additional information (e.g., ECC bytes) for data error recovery. If the EER bit is set to one, the DCR bit shall be set to zero.

The combinations of the error recovery bits (i.e., the EER bit, the PER bit, the DTE bit, and the DCR bit) are explained in table 126.

Table 126 — Combined error recovery bit descriptions (part 1 of 4)

EER	PER	DTE	DCR	Description
0	0	0	0	<p>The device server shall perform the full number of retries as specified in the READ RETRY COUNT field for read operations, the WRITE RETRY COUNT field for write operations, and the VERIFY RETRY COUNT field (see 6.3.6) for verify operations and shall perform error correction in an attempt to recover the data.</p> <p>The device server shall not report recovered read errors. The device server shall terminate a command performing a read or write operation with CHECK CONDITION status before the transfer count is exhausted only if an unrecoverable unrecovered error is detected.</p> <p>If an unrecoverable data unrecovered read error occurs during a read operation, the transfer block (TB) bit determines whether the data infor the logical block with the unrecoverable unrecovered read error may or may not be transferred to the data-in buffer depending on the setting of the transfer block (TB) bit.</p>
				<p>The device server shall perform the full number of retries as specified in the READ RETRY COUNT field for read operations, the WRITE RETRY COUNT field for write operations, and the VERIFY RETRY COUNT field (see 6.3.6) for verify operations but shall not perform error correction in an attempt to recover the data.</p> <p>The device server shall not report recovered read errors. The device server shall terminate a command performing a read or write operation with CHECK CONDITION status before the transfer count is exhausted only if an unrecoverable unrecovered error is detected.</p> <p>If an unrecoverable data unrecovered read error occurs during a read operation, the transfer block (TB) bit determines whether the data infor the logical block with the unrecoverable unrecovered read error may or may not be transferred to the data-in buffer depending on the setting of the transfer block (TB) bit.</p>
		1	0	Invalid mode. The PER bit shall be set to one if the DTE bit is set to one. ^a
		1	1	

^a If an invalid combination of the error recovery bits is sent by the application client, the device server shall terminate the MODE SELECT command with CHECK CONDITION status with the sense key set to ILLEGAL REQUEST and the additional sense code set to INVALID FIELD IN PARAMETER LIST.

Table 126 — Combined error recovery bit descriptions (part 2 of 4)

EER	PER	DTE	DCR	Description
0	1	0	0	<p>The device server shall perform the full number of retries as specified in the READ RETRY COUNT field for read operations, the WRITE RETRY COUNT field for write operations, and the VERIFY RETRY COUNT field (see 6.3.6) for verify operations and shall perform error correction in an attempt to recover the data.</p> <p>The device server shall terminate a command performing a read or write operation with CHECK CONDITION status before the transfer count is exhausted only if an unrecoverableunrecovered error is detected.</p> <p>If an unrecoverable dataunrecovered read error occurs during a read operation, the transfer block (TB) bit determines whether the data infor the logical block with the unrecoverableunrecovered read error may or may not beis transferred to the data-in buffer depending on the setting of the transfer block (TB) bit.</p> <p>The device server shall return CHECK CONDITION status with the sense key set to RECOVERED ERROR at the completion of a command performing a read or write operation during which any recoverablerecovered read error occurs. The INFORMATION field in the sense data shall contain the LBA of the last recovered error that occurred during the command.</p>
		0	1	<p>The device server shall perform the full number of retries as specified in the READ RETRY COUNT field for read operations, the WRITE RETRY COUNT field for write operations, and the VERIFY RETRY COUNT field (see 6.3.6) for verify operations but shall not perform error correction in an attempt to recover the data.</p> <p>The device server shall terminate a command performing a read or write operation with CHECK CONDITION status before the transfer count is exhausted only if an unrecoverableunrecovered error is detected.</p> <p>If an unrecoverable dataunrecovered read error occurs during a read operation, the transfer block (TB) bit determines whether the data infor the logical block with the unrecoverableunrecovered read error may or may not beis transferred to the data-in buffer depending on the setting of the transfer block (TB) bit.</p> <p>The device server shall return CHECK CONDITION status with the sense key set to RECOVERED ERROR at the completion of a command performing a read or write operation during which any recoverablerecovered error occurs. The INFORMATION field in the sense data shall contain the LBA of the last recovered error that occurred during the command.</p>
<p>^a If an invalid combination of the error recovery bits is sent by the application client, the device server shall terminate the MODE SELECT command with CHECK CONDITION status with the sense key set to ILLEGAL REQUEST and the additional sense code set to INVALID FIELD IN PARAMETER LIST.</p>				

Table 126 — Combined error recovery bit descriptions (part 3 of 4)

EER	PER	DTE	DCR	Description
0	1	1	0	<p>The device server shall perform the full number of retries as specified in the READ RETRY COUNT field for read operations, the WRITE RETRY COUNT field for write operations, and the VERIFY RETRY COUNT field (see 6.3.6) for verify operations and shall perform error correction in an attempt to recover the data.</p> <p>The device server shall terminate a command performing a read or write operation with CHECK CONDITION status before the transfer count is exhausted if any error, either recoverablerecovered or unrecoverableunrecovered, is detected. The INFORMATION field in the sense data shall contain the LBA of the block-in error.</p> <p>If an unrecoverable-dataunrecovered read error occurs during a read operation, the transfer block (TB) bit determines whether the data infor the logical block with the unrecoverableunrecovered read error may or may not beis transferred to the data-in buffer depending on the setting of the transfer block (TB) bit.</p>
		1	1	<p>The device server shall perform the full number of retries as specified in the READ RETRY COUNT field for read operations, the WRITE RETRY COUNT field for write operations, and the VERIFY RETRY COUNT field (see 6.3.6) for verify operations but shall not perform error correction in an attempt to recover the data.</p> <p>The device server shall terminate a command performing a read or write operation with CHECK CONDITION status before the transfer count is exhausted if any error, either recoverablerecovered or unrecoverableunrecovered, is detected. The INFORMATION field in the sense data shall contain the LBA of the block-in error.</p> <p>If an unrecoverable-dataunrecovered read error occurs during a read operation, the transfer block (TB) bit determines whether the data infor the logical block with the unrecoverableunrecovered read error may or may not beis transferred to the data-in buffer depending on the setting of the transfer block (TB) bit.</p>
1	0	0	0	<p>The device server shall perform the fewest possible number of retries and perform error correction in an attempt to recover the data.</p> <p>The device server shall not report recovered errors. The device server shall terminate a command performing a read or write operation with CHECK CONDITION status before the transfer count is exhausted only if an unrecoverableunrecovered error is detected.</p> <p>If an unrecoverable-dataunrecovered read error occurs during a read operation, the transfer block (TB) bit determines whether the data infor the logical block with the unrecoverableunrecovered read error may or may not beis transferred to the data-in buffer depending on the setting of the transfer block (TB) bit.</p>
		0	1	Invalid mode. The DCR bit shall be set to zero if the EER bit is set to one. ^a
		<u>1</u>	<u>0</u>	Invalid mode. The PER bit shall be set to one if the DTE bit is set to one. ^a
		<u>1</u>	<u>1</u>	
^a If an invalid combination of the error recovery bits is sent by the application client, the device server shall terminate the MODE SELECT command with CHECK CONDITION status with the sense key set to ILLEGAL REQUEST and the additional sense code set to INVALID FIELD IN PARAMETER LIST.				

Table 126 — Combined error recovery bit descriptions (part 4 of 4)

EER	PER	DTE	DCR	Description
1	1	0	0	<p>The device server shall perform the fewest possible number of retries and perform error correction in an attempt to recover the data.</p> <p>The device server shall terminate a command <u>performing a read or write operation</u> with CHECK CONDITION status before the transfer count is exhausted only if an unrecoverable<u>unrecovered</u> error is detected.</p> <p>If an unrecoverable-data<u>unrecovered read</u> error occurs during a read operation, <u>the transfer block (TB) bit determines whether</u> the data in<u>for</u> the <u>logical</u> block with the unrecoverable<u>unrecovered read</u> error may or may not be<u>is</u> transferred to the data-in buffer depending on the setting of the transfer block (TB) bit.</p> <p>The device server shall return CHECK CONDITION status with the sense key set to RECOVERED ERROR at the completion of a command <u>performing a read or write operation</u> during which any recoverable<u>recovered</u> error occurs. The INFORMATION field in the sense data shall contain the LBA of the last recovered error that occurred during the command.</p>
			1	Invalid mode. The DCR bit shall be set to zero if the EER bit is set to one. ^a
		1	0	<p>The device server shall perform the fewest possible number of retries and perform error correction in an attempt to recover the data.</p> <p>The device server shall terminate the a command <u>performing a read or write operation</u> with CHECK CONDITION status before the transfer count is exhausted if any error, either recoverable<u>recovered</u> or unrecoverable<u>unrecovered</u>, is detected. The INFORMATION field in the sense data shall contain the LBA of the block-in error.</p> <p>If an unrecoverable-data<u>unrecovered read</u> error occurs during a read operation, <u>the transfer block (TB) bit determines whether</u> the data in<u>for</u> the <u>logical</u> block with the unrecoverable<u>unrecovered read</u> error may or may not be<u>is</u> transferred to the data-in buffer depending on the setting of the transfer block (TB) bit.</p>
			1	Invalid mode. The DCR bit shall be set to zero if the EER bit is set to one. ^a
^a If an invalid combination of the error recovery bits is sent by the application client, the device server shall terminate the MODE SELECT command with CHECK CONDITION status with the sense key set to ILLEGAL REQUEST and the additional sense code set to INVALID FIELD IN PARAMETER LIST.				

The READ RETRY COUNT field specifies the number of times that the device server shall attempt its recovery algorithm during read operations.

The WRITE RETRY COUNT field specifies the number of times that the device server shall attempt its recovery algorithm during write operations.

The RECOVERY TIME LIMIT field specifies in milliseconds the maximum time duration that the device server shall use for data error recovery procedures. The device server may round this value as described in SPC-4. The limit in this field specifies the maximum error recovery time allowed for any individual logical block. A RECOVERY TIME LIMIT field set to zero specifies that the device server shall use its default value.

When both a retry count and a recovery time limit are specified, the field that specifies the recovery action of least duration shall have priority.

NOTE 26 - To disable all types of correction and retries the application client should set the EER bit to zero, the PER bit to one, the DTE bit to one, the DCR bit to one, the READ RETRY COUNT field to 00h, the WRITE RETRY COUNT field to 00h, and the RECOVERY TIME LIMIT field to 0000h.

6.3.6 Verify Error Recovery mode page

The Verify Error Recovery mode page (see table 127) specifies the error recovery parameters the device server shall use during the VERIFY command and the verify operation of the WRITE AND VERIFY command.