

To: scsi reflector  
cc: Jim Coomes, Mike Miller, Bharat Kothari  
From: Gerry Houlder  
Date: 02/28/94 05:00:20 PM  
Subject: Data Recovery on Deferred Errors

#### PROBLEM STATEMENT

We have a customer that is concerned about recovering write data after a deferred error. When Write Caching is used, the WRITE command will return GOOD status before the data is written to the disc. If an unrecoverable error occurs during that write operation, a later command from the same initiator will end with CHECK CONDITION status. The sense bytes for that error will report Deferred Error and the Logical Block Address of the erroring block. When such an error occurs, the customer wants to read the unwritten data back and attempt to retry the write (possibly at a different location). Existing direct access device SCSI doesn't have enough information/commands to do this to our customer's satisfaction.

#### POSSIBLE SOLUTION

The Auto Contingent Allegiance (ACA) feature of SCSI-3 is almost enough. The ACA procedure allows for retrieving sense data, then issuing a READ command to "retrieve" the unwritten data from the target's buffer. Of course, the target must be careful to retain all write data for failed write commands (and allow that data to be used to satisfy subsequent READ requests) until after the Contingent Allegiance is cleared. The returned sense data will contain the first LBA that was not written to disc in the information bytes. This can be used by the following READ command to retrieve the unwritten data. The only item that is still unknown is the number of blocks of unwritten data that is available.

The host is not interested in retaining the starting LBA/number of blocks information for all write commands. They consider this to be almost as bad as having to retain the write data. The problem is that there is no definite time when the data is guaranteed to be written to disk and the information doesn't need to be retained any more. Therefore they want the returned sense data to indicate the number of blocks of unwritten data that are available as well as the LBA of the first unwritten block.

I believe the "Command Specific Information Bytes" could be used for this purpose. This is a 4 byte field. Existing command descriptor blocks have a maximum of 2 bytes for the number of blocks to be transferred, although there has been talk of 12 and 16 byte CDBs that may have a 4 byte transfer length field. In reality, the number is limited to much lower values by the buffer space available on the target device. This use of the command specific information bytes could be required only for the "deferred error" sense type. This would limit most of the "backwards compatibility" problems.

I will introduce this idea at the March 16 SCSI Working Group Meeting at Newport Beach. If anyone would like to make comments or suggestions for improvements before then, my Internet address is:

gerry\_houlder@notes.seagate.com.