

Accredited Standards Committee
X3, Information Processing Systems

Doc: X3T10/94-179R0
Date: 7 September, 1994
Project:N/A
Ref Doc.:
Reply to: John Scheible

To: X3T10 Membership
From: John Scheible, IBM

Subject: Protecting against out of order writes to the media proposal

BACKGROUND

UNIX based systems timestamp the beginning and end of a multisector file, and assume the file is correct if the timestamps agree. However, if the disk drive does an out of order write this can cause corrupted data to go undetected. If the drive writes the end of the file (ending timestamp), then the beginning of the file (beginning timestamp), but then fails to write the middle of the record (due to hardware error), the data is corrupted but UNIX will not detect this case. Out of order writes could be prohibited, but they do cause significant performance advantages for those systems that can handle the case shown above (most systems).

Current systems can control out of order writes across the interface, but the disk drive is not prohibited from writing the data out of order to the media. The disk drive is only responsible for giving Check Condition Status at the end of the command (or during the Synchronize Cache command).

The X3T10.1 (SSA) committee entertained a proposal to add a bit to control out of order writes to the media, but decided that this problem is independent of the physical layer and recommended that a change be made to the Write CDB to handle this problem independent of the physical layer. This proposal is the result of that decision.

CHANGES

This proposal recommends a change to the Write (10) command descriptor block to add a bit that controls the ability for the Target to write the data out of order. I took the SCSI-2 description of the Write (10) command, so I may not have all the SCSI-3 changes.

Add bit 1 in byte 1 of the Write (10) CDB and add the text that follows:

| Byte | Bit | | | | | | | |
|------|-----|---|---|-----|-----|----------|------|--------|
| | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
| 1 | LUN | | | DPO | FUA | Reserved | OOWM | RelAdr |

A OOWM (Out of Order Write to the Media) bit of 1 indicates the Target may write the data to the media in an order other than that which was transmitted over the interface (The Target may perform a split write operation). A OOWM bit of 0 indicates the Target must write the data to the media in the same order it was transferred over the SCSI interface.

I welcome any comments.

Sincerely,

John Scheible

Voice: (512) 823-8208

FAX: (512) 823-0758

Email: Scheible@vnet.ibm.com