

Date: September 04, 2003

To: T10 Committee (SCSI)

From: George Penokie (IBM/Tivoli), David Chambliss (IBM Almaden Research Center)

Subject: Per-Command Priority Tagging

1 Overview

The following proposed wording represents changes to SCSI Block Commands - 2 (SBC-2) to enable the transmission of priority information on a per-command basis.

This proposal standardizes the interface by which device servers can offer differentiated quality of service to different applications associated with the same initiator. Examples of its use would include offering lower priority on IO associated with background destage writes within a storage controller or on IO associated with background applications, so that response time may be reduced for those IO operations that directly affect the responsiveness offered to the end user.

SBC-2 additions

1.1 Affected commands

It is proposed that a 5-bit field (PRIORITY) be added to at least the following commands:

- a) READ (10)
- b) READ (12)
- c) READ (16)
- d) VERIFY (10)
- e) VERIFY (12)
- f) VERIFY (16)
- g) WRITE (10)
- h) WRITE (12)
- i) WRITE (16)
- j) WRITE AND VERIFY (10)
- k) WRITE AND VERIFY (12)
- l) WRITE AND VERIFY (16)

In addition the PRIORITY field should be added to the following commands:

- a) PRE-FETCH (10)
- b) PRE-FETCH (16)
- c) REBUILD (16)
- d) REBUILD (32)
- e) REGENERATE (16)
- f) REGENERATE (32)
- g) SEEK (10)
- h) SYNCHRONIZE CACHE (10)
- i) SYNCHRONIZE CACHE (16)
- j) WRITE SAME (10)
- k) WRITE SAME (16)
- l) XDREAD (10)
- m) XDREAD (32)
- n) XDWRITE (10)
- o) XDWRITE (32)
- p) XDWRITEREAD (10)
- q) XDWRITEREAD (32)
- r) XDWRITE EXTENDED (16)
- s) XDWRITE EXTENDED (32)
- t) XDWRITE EXTENDED (64)

- u) XPWRITE (10)
- v) XPWRITE (32)

The PRIORITY field would be in bits 4-0 in byte 6 of 10-byte commands, byte 10 of 12-byte commands, byte 14 of 16-byte commands and byte 11 of 32-byte commands.

3.1 Definitions

[3.1.1 priority task: A task to which a priority has been assigned by an application client or a device server \(see 3.1.3\).](#)

3.1.2 READ (10) command

The READ (10) command (see table 1) requests that the device server transfer data to the application client. The most recent data value written in the addressed logical block shall be returned.

Table 1 — READ (10) command

Byte\Bit	7	6	5	4	3	2	1	0
0	OPERATION CODE (28h)							
1	Reserved			DPO	FUA	Reserved		RELADR
2	(MSB) _____ LOGICAL BLOCK ADDRESS _____ (LSB)							
5								
6	Reserved			PRIORITY				
7	(MSB) _____ TRANSFER LENGTH _____ (LSB)							
8								
9	CONTROL							

[The PRIORITY field specifies the relative scheduling of this task in relation to other simple tasks already in the task set for service by the device server, as determined by the PRIOR_CTL field in the Control mode page \(see 3.1.3\).](#)

3.1.3 Control mode page

The Control mode page (see table 2) provides controls over several SCSI features that are applicable to all device types such as tagged queuing and error logging.

Table 2 — Control mode page

Bit Byte	7	6	5	4	3	2	1	0
0	PS	SPF (0b)	PAGE CODE (0Ah)					
1	PAGE LENGTH (0Ah)							
2	TST			Reserved		D_SENSE	GLTSD	RLEC
3	QUEUE ALGORITHM MODIFIER				Reserved	QERR		DQUE
4	TAS	RAC	UA_INTLCK_CTRL		SWP	Obsolete		
5	Reserved		PRIOR_CTL			AUTOLOAD MODE		
6	Obsolete							
7								
8	(MSB)	BUSY TIMEOUT PERIOD						
9								(LSB)
10	(MSB)	EXTENDED SELF-TEST COMPLETION TIME						
11								(LSB)

The PRIOR_CTL field (see table 3) indicates how the device server shall interpret the value contained in the PRIORITY field within the CDB.

Table 3 — PRIOR_CTL field

Value		Description
PRIOR_CTL	PRIORITY	
<u>000b</u>	<u>xxh</u>	<u>The contents of the PRIORITY field shall be ignored.</u>
<u>001b</u> ^b	<u>00h</u>	<u>If the task attribute for the task is SIMPLE then the task's priority shall be equivalent to the average PRIORITY value of tasks in the task set. If there are no priority tasks within the task set then the task shall become a priority task and be assigned a priority of 10h. If the task attribute for the task is not SIMPLE then the PRIORITY field shall be ignored.</u>
	<u>01h - 1Fh</u>	<u>If the task attribute for the task is SIMPLE then the task shall be assigned a priority based on the value of the PRIORITY field with 01h denoting the lowest priority and 1Fh denoting the highest priority. If the task attribute for the task is not SIMPLE then the PRIORITY field shall be ignored.</u> <u>NOTE 1 - A difference in priority between tasks does not necessarily override other task manager scheduling considerations, such as different times to access different logical block addresses. However, processing of a set of tasks with different priorities should cause the subset of tasks with the higher priority to return status sooner in aggregate than the same subset would if the same set of tasks were submitted under the same conditions but with all priorities equal.</u>
	<u>none</u> ^a	<u>If the task attribute for the task is SIMPLE then the task's priority shall be equivalent to the average PRIORITY value of tasks in the task set. If there are no priority tasks within the task set then the task shall become a priority task and be assigned a priority of 10h.</u>
<u>010b</u>	<u>xxh</u>	<u>The interpretation of the PRIORITY field is not defined by this standard.</u>
<u>011b-111b</u>	<u>xxh</u>	<u>Reserved</u>
^a <u>Any CDB that does not support the PRIORITY field.</u> ^b <u>Any tasks in the task set when the device server processes the control mode page shall become priority tasks and be assigned a priority of 10h</u>		