Date: December 09, 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 FCP-3, SAS 1.1, and SRP-2 to enable the transmission of priority information on a per-command basis.

It also request the removal of the priority field from any OSD command that contains one.

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.

The method defined in this proposal to accomplish this involves changes to the protocol standards to accommodate an extension to the task attribute field to allow different priorities to be assigned to simple task attributes. There should be no changes required to SAM-3 as a simple task with a priority is still architecturally handled the same as a simple task.

2 OSD changes

Remove the priority field and OSD service action that contains it including APPEND, READ, and WRITE service actions.

3 FCP-3, SAS 1.1, and SRP-2 additions

The changes below will be applied to the FCP-3, SAS 1.1, and SRP-2 standards. Only the section numbers are different.

.

Table 1 — COMMAND information unit

Byte\Bit	7	6	5	4	3	2	1	0
0	(MSB)							
7				LOGICAL UNIT NUMBER				(LSB)
8	Reserved							
9	Reserved			PRIORITY			TASK ATTRIBUTE	
10	Reserved							
11	ADDITIONAL CDB LENGTH (n dwords)						Reserved	
12								
27				CDB				
28								
27+n×4				ADDITIONAL CDB BYTES				

The PRIORITY field specifies the relative scheduling of this task in relation to other tasks already in the task set for processing by the device server. The relationship between the PRIORITY field and the TASK ATTRIBUTE field is defined in table 2. A priority of zero indicates the device server schedules the task using the default priority of the initiator associated with the I_T nexus of the received task. If no default priority has been assigned to the initiator using the SET DEFAULT PRIORITY command (see SPC-3) the priority assigned to the task is vendor specific. Priority 1h is the highest priority, with increasing priority values indicating lower priorities.

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.

The TASK ATTRIBUTE field is defined in table 2.

Table 2 — TASK ATTRIBUTE **field**

Task Attribute Code	Task attribute	Priority Code	Description
000b	SIMPLE	0h	Requests that the task be managed according to the rules for a simple task attribute (see SAM-3). The default priority code may be used by the task manager to determine an ordering to process simple tasks in addition to it's vendor specific ordering rules.
		1h-Fh	Requests that the task be managed according to the rules for a simple task attribute (see SAM-3). The priority code may be used by the task manager to determine an ordering to process simple tasks in addition to it's vendor specific ordering rules.
001b	HEAD OF QUEUE	Reserved	Requests that the task be managed according to the rules for a head of queue task attribute (see SAM-3).
010b	ORDERED	Reserved	Requests that the task be managed according to the rules for an ordered task attribute (see SAM-3).
011b	Reserved		
100b	ACA	Reserved	Requests that the task be managed according to the rules for an automatic contingent allegiance task attribute (see SAM-3).
101b-111b	Reserved		

4 Additions to SPC

4.1 REPORT DEFAULT PRIORITY command

The REPORT DEFAULT PRIORITY command (see table 3) requests the current default priority that has been assigned to the initiator port through which this command was sent.

The REPORT DEFAULT PRIORITY command is a service action of the MAINTENANCE IN command. Additional MAINTENANCE IN service actions are defined in SCC-2 and in this standard. The MAINTENANCE IN service actions defined in SCC-2 apply only to logical units that return a device type of 0Ch or the SCCS bit equal to one in their standard INQUIRY data (see 6.4.2).

Table 3 — REPORT DEFAULT PRIORITY command

Bit / Byte	7	6	5	4	3	2	1	0
0	OPERATION CODE (A3h)							
1	Reserved			SERVICE ACTION (xxh)				
2	Reserved							
5								
6	(MSB)			ALLOCATION LENGTH (4h or larger)				
9								(LSB)
10	Reserved							
11	CONTROL							

The ALLOCATION LENGTH field indicates the number of bytes that have been allocated for the returned parameter data. The allocation length shall be at least four. If the allocation length is less than for the command shall be terminated with a CHECK CONDITION status, the sense key shall be set to ILLEGAL REQUEST, and the additional sense code shall be set to INVALID FIELD IN CDB.

The format of the parameter data returned by the REPORT DEFAULT PRIORITY command is shown in table 4.

Table 4 — REPORT DEFAULT PRIORITY parameter data

Bit / Byte	7	6	5	4	3	2	1	0
0	Reserve				CURRENT DEFAULT PRIORITY			
1	Reserved							
3								

The CURRENT DEFAULT PRIORITY field contains the current default priority (see FCP-3, SAS 1.1, or SRP-2) for the initiator port associated with the I_T nexus of this command.

4.2 SET DEFAULT PRIORITY command

The SET DEFAULT PRIORITY command (see table 5) requests that the default priority for the SCSI initiator port associated with the I_T nexus for this command be set to the value received in the DEFAULT PRIORITY field of the CDB.

The SET DEFAULT PRIORITY command is a service action of the MAINTENANCE OUT command. Additional MAINTENANCE OUT service actions are defined in SCC-2 and in this standard. The MAINTENANCE OUT service actions defined only in SCC-2 apply only to logical units that return a device type of 0Ch or the sccs bit equal to one in their standard INQUIRY data

Table 5 — SET DEFAULT PRIORITY command

Bit / Byte	7	6	5	4	3	2	1	0
0	OPERATION CODE (A4h)							
1	Reserved			SERVICE ACTION (xxh)				
2	S_ALL_PRY	Reserved			DEFAULT PRIORITY			
3	Reserved							
4	Reserved							
5	Reserved							
6	(MSB)							
7				Reserved				
8								
9								(LSB)
10	Reserved							
11	CONTROL							

The DEFAULT PRIORITY field indicates the default priority. The value in the DEFAULT PRIORITY field shall be returned in subsequent REPORT DEFAULT PRIORITY commands.

A set all priority (S_ALL_PRY) set to zero indicates the default priority value shall only apply to the initiator associated with the I_T nexus from which the SET DEFAULT PRIORITY command was sent (see FCP-3, SAS 1.1, or SRP-2). A S_ALL_PRY set to one indicates the default priority value shall be used for all initiators regardless of any prior default priority.

If the S_ALL_PRY bit is set to one, on successful completion of a SET DEFAULT PRIORITY command a unit attention shall be generated for all initiators except the one that issued the service action. When reporting the unit attention condition the additional sense code shall be set to DEFAULT PRIORITY CHANGED.

Editor's Note 1: The DEFAULT PRIORITY CHANGED is a new ASCQ. I would suggest it be coded as 0Fh 12h.