

Date: March 13, 2006

To: T10 Committee (SCSI)

From: George Penokie (IBM)

Subject: SPC-4: Statistics and Performance Log Pages

## **1 Overview**

In large networks it is becoming important, and in some cases necessary, to monitor the performance of logical units to determine that the customer is receiving the level of throughput they have contracted. This is best handled by setting up log pages that would return a set of specified performance parameters.

This proposal fills an important gap in the ability to monitor block storage resources. It provides a mechanism, not previously available, by which block devices are able to measure and report their utilization. In addition, it is, currently, the only way that a block storage device has to report the activity of an identified I/O group. (i.e., by group number).

This proposal defines a log page that will return a set of overall performance parameters and a set of sublog pages, for that log page, that would return performance parameters based on group number.

### **1.1 Symbols and acronyms**

Add the following acronyms:

s - seconds

ns - nanoseconds

#### **1.1.1 Log page structure and page codes for all device types**

Add the Statistics and Performance log page to the Log page codes table (Table 212).

#### **1.1.2 Statistics and Performance log pages**

##### **1.1.2.1 Statistics and Performance log page overview**

The Statistics and Performance log pages consist of a General Statistics and Performance log page and up to 31 Group Statistics and Performance subpage logs. Each Group Statistics and Performance subpage log only collects statistics and performance information for the group number specified in a read CDB or a write CDB (see table 8).

The General Statistics and Performance log page (see 1.1.2.2) provides the following statistics and performance results associated to the addressed logical unit:

- a) number of read commands;
- b) number of write commands;
- c) number of read logical blocks transmitted by a target port;
- d) number of write logical blocks received by a target port;
- e) read command processing time;
- f) write command processing time;
- g) sum of the command weights of the read commands plus write commands;
- h) sum of the weighted command time of the read commands plus write commands;
- i) idle time; and
- j) time interval.

The Group Statistics and Performance subpage logs (see 1.1.2.3) provide the following statistics and performance results associated to the addressed logical unit and the GROUP NUMBER field:

- a) number of read commands;
- b) number of write commands;
- c) number of read logical blocks transmitted by a target port;

- d) number of write logical blocks received by a target port;
- e) read command processing time; and
- f) write command processing time.

In the General Statistics and Performance log page and the Group Statistics and Performance subpage logs a read command is one of the following commands:

- a) READ(6) command;
- b) READ(10) command;
- c) READ(12) command;
- d) READ(16) command;
- e) READ(32) command;
- f) READ CD command;
- g) READ CD MSF command;
- h) READ REVERSE(16) command;
- i) XDREAD(10) command; or
- j) XDREAD(32) command.

In the General Statistics and Performance log page and the Group Statistics and Performance subpage logs a write command is one of the following commands:

- a) WRITE(6) command;
- b) WRITE(10) command;
- c) WRITE(12) command;
- d) WRITE(16) command;
- e) WRITE(32) command;
- f) WRITE AND VERIFY(10) command;
- g) WRITE AND VERIFY(12) command;
- h) WRITE AND VERIFY(16) command;
- i) WRITE AND VERIFY(32) command;
- j) XDWRITE(10) command; or
- k) XDWRITE(32) command.

The General Statistics and Performance log page command weight is calculated as follows:

$$\text{command weight} = (360\ 360 / \text{task priority})$$

where:

task priority is as defined in SAM-3. However, if the result is a task priority of zero, then task priority shall be set to seven.

The General Statistics and Performance log page weighted command time is calculated as follows:

$$\text{weighted command time} = (\text{time increments processing the command} \times \text{time interval}) \times (360\ 360 / \text{task priority})$$

where:

task priority is as defined in SAM-3. However, if the result is a task priority of zero, then task priority shall be set to seven;

time increments processing a command is the number of time intervals from the time the task manager places the command into a task set until the device server sends a SCSI transport protocol service response for the command; and

time interval is the value represented in the TIME INTERVAL DESCRIPTOR field of the Time Interval log parameter (see table 5).

In the General Statistics and Performance log page the idle time is calculated as follows:

$$\text{idle time} = (\text{time increments not processing commands} \times \text{time interval}).$$

where:

time increments not processing commands is the number of time intervals when there are no commands in the task set and the device server has sent a SCSI transport protocol service response for all commands being processed (i.e., there are no commands to be processed or being processed).

time interval is the value represented in the TIME INTERVAL DESCRIPTOR in the Time Interval log parameter (see table 5).

**1.1.2.2 General Statistics and Performance log page**

Table 1 specifies the General Statistics and Performance log page parameters.

**Table 1 — General Statistics and Performance log page**

Bit Byte	7	6	5	4	3	2	1	0
0	DS (0b)	SPF (0b)	PAGE CODE (xxh)					
1	SUBPAGE CODE (00h)							
2	(MSB)	PAGE LENGTH (005Ch)						
3							(LSB)	
General Statistics and Performance log parameters								
4	Statistics and Performance log parameter							
71								
72	Idle Time log parameter							
83								
84	Time Interval log parameter							
95								

The DS bit, SPF bit, PAGE CODE field, SUBPAGE CODE field and PAGE LENGTH field are described in 7.2.1.

Table 2 shows the format of Statistics and Performance log parameter.

**Table 2 — Statistics and Performance log parameter format**

Bit Byte	7	6	5	4	3	2	1	0
0	(MSB) _____							
1	PARAMETER CODE (0001h) _____							
2	DU	Obsolete	TSD	ETC	TMC	FORMAT AND LINKING		
3	PARAMETER LENGTH (40h) _____							
4	(MSB) _____							
11	NUMBER OF READ COMMANDS _____							
12	(MSB) _____							
19	NUMBER OF WRITE COMMANDS _____							
20	(MSB) _____							
27	NUMBER OF LOGICAL BLOCKS RECEIVED BY A TARGET PORT _____							
28	(MSB) _____							
35	NUMBER OF LOGICAL BLOCKS TRANSMITTED BY A TARGET PORT _____							
36	(MSB) _____							
43	READ COMMAND PROCESSING TIME _____							
44	(MSB) _____							
51	WRITE COMMAND PROCESSING TIME _____							
52	(MSB) _____							
59	WEIGHTED NUMBER OF READ COMMANDS PLUS WRITE COMMANDS _____							
60	(MSB) _____							
67	WEIGHTED READ COMMAND PROCESSING PLUS WRITE COMMAND PROCESSING _____							
	(LSB)							

The PARAMETER CODE field set to 0001h identifies the log parameter being transferred as the Statistics and Performance log parameter.

The values of the log parameter control bits and fields for Statistics and Performance log parameters are as specified in table 3.

**Table 3 — Log parameter control bits and fields**

Field	Value	Description
DU bit	0b	Value provided by device server
TSD bit	0b	Device server manages saving of parameter
ETC bit	0b	No threshold comparison is made on this value
TMC field	xxb	Ignored when the ETC bit is set to zero
FORMAT AND LINKING field	10b	If another parameter reported in this log page reaches its maximum value, then this parameter shall not stop incrementing. This parameter may be reinitialized by a LOG SELECT command.

The PARAMETER LENGTH field specifies the length in bytes of the statistics and performance fields that follow.

**T10/05-248 revision 4**

The NUMBER OF READ COMMANDS field contains the number of read commands (see 1.1.2.1) received by the addressed logical unit.

The NUMBER OF WRITE COMMANDS field contains the number of write commands (see 1.1.2.1) received by the addressed logical unit.

The NUMBER OF LOGICAL BLOCKS RECEIVED BY A TARGET PORT field contains the number of logical blocks received from the service delivery subsystem for the device server of the addressed logical unit as a result of write commands (see 1.1.2.1).

The NUMBER OF LOGICAL BLOCKS TRANSMITTED BY A TARGET PORT field contains the number of logical blocks delivered to the service delivery subsystem by the device server of the addressed logical unit as a result of read commands (see 1.1.2.1).

The READ COMMAND PROCESSING INTERVALS field contains the cumulative number of time intervals (see table 5) spent processing read commands addressed to the logical unit (see 1.1.2.1).

The WRITE COMMAND PROCESSING INTERVALS field contains the cumulative number of time intervals (see table 5) spent processing write commands addressed to the logical unit (see 1.1.2.1).

If task priority is supported (see SAM-4), then the WEIGHTED NUMBER OF READ COMMANDS PLUS WRITE COMMANDS field contains the cumulative weighted number of read commands and write commands addressed to the logical unit (see 1.1.2.1).

If task priority is not supported, then the WEIGHTED NUMBER OF READ COMMANDS PLUS WRITE COMMANDS field shall be set to zero.

If task priority is supported (see SAM-4), then the WEIGHTED NUMBER OF READ COMMANDS PLUS WRITE COMMANDS field contains the cumulative weighted number of time intervals (see table 5) spent processing read commands and write commands addressed to the logical unit (see 1.1.2.1).

If task priority is not supported, then the WEIGHTED READ COMMAND PROCESSING PLUS WRITE COMMAND PROCESSING field shall be set to zero.

Table 4 shows the format of the Idle Time log parameter.

**Table 4 — Idle Time log parameter format**

Bit Byte	7	6	5	4	3	2	1	0
0	(MSB) _____							
1	PARAMETER CODE (0002h)						_____ (LSB)	
2	DU	Obsolete	TSD	ETC	TMC	FORMAT AND LINKING		
3	PARAMETER LENGTH (08h)							
4	(MSB) _____							
11	IDLE TIME INTERVALS						_____ (LSB)	

The PARAMETER CODE field set to 0002h identifies the log parameter being transferred as the Idle Time log parameter.

The values of the log parameter control bits and fields for Statistics and Performance log parameters are as specified in table 3.

The PARAMETER LENGTH field specifies the length in bytes of the IDLE TIME INTERVAL field that follows.

The IDLE TIME INTERVALS field contains the cumulative number of time intervals (see table 5) spent when there are no tasks in the task set and there are no tasks being processed by a logical unit.

Table 5 shows the format of the Time Interval log parameter.

**Table 5 — Time Interval log parameter format**

Bit Byte	7	6	5	4	3	2	1	0
0	(MSB) _____							
1	PARAMETER CODE (0003h) _____							
2	DU	Obsolete	TSD	ETC	TMC	FORMAT AND LINKING		
3	PARAMETER LENGTH (08h)							
4	_____							
11	TIME INTERVAL DESCRIPTOR _____							

The PARAMETER CODE field set to 0003h identifies the log parameter being transferred as the Time Interval log parameter.

The values of the log parameter control bits and fields for Time Interval log parameter is specified in table 3.

The PARAMETER LENGTH field specifies the length in bytes of the TIME INTERVAL DESCRIPTOR that follows.

The TIME INTERVAL DESCRIPTOR (see table 6) contains the time interval in seconds used in the Read Command Processing Time log parameter, the Write Command Processing Time log parameter, the Weighted Read Command Processing Plus Write Command Processing log parameter, and the Idle Time log parameter.

**Table 6 — TIME INTERVAL DESCRIPTOR format**

Bit Byte	7	6	5	4	3	2	1	0
0	(MSB) _____							
3	EXPONENT _____							
4	(MSB) _____							
7	INTEGER _____							
	(LSB)							

The EXPONENT field contains the negative power of 10 exponent to multiply with the INTEGER field (e.g., a value of 9 represents  $10^{-9}$ )

After the exponent has been applied, the INTEGER field contains the value that represents one time interval (e.g., a value of 5 in the INTEGER field and a value of 9 in the EXPONENT field represents a time interval of  $5 \times 10^{-9}$  s or 5 ns).

### 1.1.2.3 Group Statistics and Performance (n) log page

The Group Statistics and Performance (n) log pages (see table 7) are subpages of the General Statistics and Performance log page (see 1.1.2.2) and provide logging of statistics and performance of read and write operations based on group numbers. There are 31 Group Statistics and Performance (n) log pages one for each group number. The statistics and performance information associated with each group number is collected in the corresponding Group Statistics and Performance (n) log page (e.g., operations associated with group number 16 are logged in the Group Statistics and Performance (16) log page).

**Table 7 — Group Statistics and Performance (n) log pages**

Bit Byte	7	6	5	4	3	2	1	0
0	DS (0b)	SPF (1b)	PAGE CODE (xxh)					
1	SUBPAGE CODE (01h - 1Fh)							
2	(MSB)	PAGE LENGTH (34h)						(LSB)
3								
Group Statistics and Performance log parameters								
4	Group n Statistics and Performance log parameter <sup>a</sup>							
55								
<sup>a</sup> The log parameter associated with the specific group number as specified by the value of n is collected in the corresponding log parameter (e.g., the count of read commands with the GROUP NUMBER field set to 16 is logged in the Group 16 Number of Read Commands log parameter of the Group Statistics and Performance (16) log page).								

The DS bit, SPF bit, PAGE CODE field and PAGE LENGTH field are described in 7.2.1.

The SUBPAGE CODE field is as specified in table 8.

**Table 8 — General Statistics and Performance log page subpage codes (part 1 of 2)**

Subpage code	Log page name	Group number <sup>a b</sup>
01h	Group Statistics and Performance (1)	00001b
02h	Group Statistics and Performance (2)	00010b
03h	Group Statistics and Performance (3)	00011b
04h	Group Statistics and Performance (4)	00100b
05h	Group Statistics and Performance (5)	00101b
06h	Group Statistics and Performance (6)	00110b
07h	Group Statistics and Performance (7)	00111b
08h	Group Statistics and Performance (8)	01000b
09h	Group Statistics and Performance (9)	01001b
0Ah	Group Statistics and Performance (10)	01010b
0Bh	Group Statistics and Performance (11)	01011b
0Ch	Group Statistics and Performance (12)	01100b
0Dh	Group Statistics and Performance (13)	01101b
0Eh	Group Statistics and Performance (14)	01110b
0Fh	Group Statistics and Performance (15)	01111b
10h	Group Statistics and Performance (16)	10000b
11h	Group Statistics and Performance (17)	10001b
<sup>a</sup> The GROUP NUMBER field is from the read command CDB or the write command CDB (see SBC-3). <sup>b</sup> The statistics and performance information associated with a group number is collected in the corresponding Group Statistics and Performance (n) log page (e.g., operations associated with group number 16 are logged in the Group Statistics and Performance (16) log page).		

**Table 8 — General Statistics and Performance log page subpage codes** (part 2 of 2)

<b>Subpage code</b>	<b>Log page name</b>	<b>Group number</b> <sup>a b</sup>
12h	Group Statistics and Performance (18)	10010b
13h	Group Statistics and Performance (19)	10011b
14h	Group Statistics and Performance (20)	10100b
15h	Group Statistics and Performance (21)	10101b
16h	Group Statistics and Performance (22)	10110b
17h	Group Statistics and Performance (23)	10111b
18h	Group Statistics and Performance (24)	11000b
19h	Group Statistics and Performance (25)	11001b
1Ah	Group Statistics and Performance (26)	11010b
1Bh	Group Statistics and Performance (27)	11011b
1Ch	Group Statistics and Performance (28)	11100b
1Dh	Group Statistics and Performance (29)	11101b
1Eh	Group Statistics and Performance (30)	11110b
1Fh	Group Statistics and Performance (31)	11111b
<sup>a</sup> The GROUP NUMBER field is from the read command CDB or the write command CDB (see SBC-3). <sup>b</sup> The statistics and performance information associated with a group number is collected in the corresponding Group Statistics and Performance (n) log page (e.g., operations associated with group number 16 are logged in the Group Statistics and Performance (16) log page).		

Table 9 shows the format of Group n Statistics and Performance log parameter.



**Table 9 — Group n Statistics and Performance log parameter format**

Bit Byte	7	6	5	4	3	2	1	0
0	(MSB) _____							
1	PARAMETER CODE (0001h) _____							
2	DU	Obsolete	TSD	ETC	TMC	FORMAT AND LINKING		
3	PARAMETER LENGTH (30h) _____							
4	(MSB) _____							
11	GROUP N NUMBER OF READ COMMANDS _____							
12	(MSB) _____							
19	GROUP N NUMBER OF WRITE COMMANDS _____							
20	(MSB) _____							
27	GROUP N NUMBER OF LOGICAL BLOCKS RECEIVED BY A TARGET PORT _____							
28	(MSB) _____							
35	GROUP N NUMBER OF LOGICAL BLOCKS TRANSMITTED BY A TARGET PORT _____							
36	(MSB) _____							
43	GROUP N READ COMMAND PROCESSING TIME _____							
44	(MSB) _____							
51	GROUP N WRITE COMMAND PROCESSING TIME _____							
								(LSB)

The PARAMETER CODE field set to 0001h identifies the log parameter being transferred as the Group n Statistics and Performance log parameter.

The values of the log parameter control bits and fields for the Group n Statistics and Performance log parameters is specified in table 3.

The PARAMETER LENGTH field specifies the length in bytes of the group n statistics and performance parameters that follow.

The GROUP N NUMBER OF READ COMMANDS field contains the number of read commands (see 1.1.2.1) received by the addressed logical unit.

The GROUP N NUMBER OF WRITE COMMANDS field contains the number of write commands (see 1.1.2.1) received by the addressed logical unit.

The GROUP N NUMBER OF LOGICAL BLOCKS RECEIVED BY A TARGET PORT field contains the number of logical blocks received from the service delivery subsystem for the device server of the addressed logical unit as a result of write commands (see 1.1.2.1).

The GROUP N NUMBER OF LOGICAL BLOCKS TRANSMITTED BY A TARGET PORT field contains the number of logical blocks delivered to the service delivery subsystem by the device server of the addressed logical unit as a result of read commands (see 1.1.2.1).

The GROUP N READ COMMAND PROCESSING TIME field contains the cumulative number of time intervals (see table 5) spent processing read commands addressed to the logical unit (see 1.1.2.1).

The GROUP N WRITE COMMAND PROCESSING TIME field contains the cumulative number of time intervals (see table 5) spent processing write commands addressed to the logical unit (see 1.1.2.1).