T10/05-248 revision 2

Date: August 18, 2005

To: T10 Committee (SCSI)

From: George Penokie (IBM/Tivoli)

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.

1.0.1 Statistics and Performance log pages

1.0.1.1 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 2).

The General Statistics and Performance log page (see 1.0.1.2) 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 blocks sent;
d) number of write blocks received;
e) read command processing time;
f) write command processing time;
g) number read commands plus write commands weighted by priority;
h) time processing read commands plus write commands weighted by priority;
i) idle time; and
j) time interval.

The Group Statistics and Performance subpage logs (see 1.0.1.3) 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 blocks sent;
d) number of write blocks received;
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.

In the General Statistics and Performance log page the weighted priority of a command is calculated as follows:

command weight = (360 360 / priority of the command)

where:

priority of the command is the value of the PRIORITY field in the CDB or if the PRIORITY field is set to zero, then the INITIAL PRIORITY field in the Control Extension mode page (see SPC-3) or the last priority assigned to the I_T_L nexus using the SET PRIORITY command.

In the General Statistics and Performance log page the weighted time of a command is calculated as follows:

weighted command time = (time increments processing the command x time interval) x (360 360 / priority of the command).

where:

priority of the command is the value of the PRIORITY field in the CDB or if the PRIORITY field is set to zero, then the INITIAL PRIORITY field in the Control Extension mode page (see SPC-3) or the last priority assigned to the I_T_L nexus using the SET PRIORITY command;

time increments processing a command shall be the number of time intervals from the time the task manager places the command into a task until the device server receives SCSI Transport Protocol Service Confirmation that the status from the command has been sent; and

time interval is the value represented in the TIME INTERVAL DESCRIPTOR in the Time Interval log parameter (see table 6).

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

idle time= (time increments not processing commands x time interval).

where:

time increments not processing commands shall be the number of time intervals when there are no commands in the task set and the device server has received SCSI Transport Protocol Service Confirmation statuses 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 6).

1.0.1.2 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	Reserved		PAGE CODE (xxh)					
1	SUBPAGE CODE							
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 PAGE CODE and PAGE LENGTH fields are described in 7.2.1.

The SUBPAGE CODE field is as specified in table 2.

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

Subpage code	Log page name	Group number [a] [b]
00h	General Statistics and Performance	Not applicable
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

[a] The GROUP NUMBER field is from the read command CDB or the write command CDB (see SBC-3).

[b] 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 2 — General Statistics and Performance log page subpage codes (part 2 of 2)

Subpage code	Log page name	Group number [a] [b]
0Bh	Group Statistics and Performance (11)	01011b
0Ch	Group Statistics and Performance (12)	00100b
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
12h	Group Statistics and Performance (18)	10010b
13h	Group Statistics and Performance (19)	10011b
14h	Group Statistics and Performance (20)	10100b
14h	Group Statistics and Performance (21)	10101b
15h	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)	10100b
1Dh	Group Statistics and Performance (29)	11101b
1Eh	Group Statistics and Performance (30)	11110b
1Fh	Group Statistics and Performance (31)	11111b

[a] The GROUP NUMBER field is from the read command CDB or the write command CDB (see SBC-3).

[b] 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 3 shows the format of Statistics and Performance log parameter.

Table 3 — Statistics and Performance log parameter format

Bit / Byte	7	6	5	4	3	2	1	0
0	(MSB)							
1		PARAMETER CODE (0001h)						(LSB)
2	DU	DS	TSD	ETC	TMC		LBIN	LP
3	PARAMETER LENGTH (40h)							
4	(MSB)							
11		NUMBER OF READ COMMANDS						(LSB)
12								
19		NUMBER OF WRITE COMMANDS						
20								
27		NUMBER OF BLOCKS RECEIVED						
28								
35		NUMBER OF BLOCKS SENT						
36								
43		READ COMMAND PROCESSING TIME						
44								
51		WRITE COMMAND PROCESSING TIME						
52								
59		WEIGHED NUMBER OF READ COMMANDS PLUS WRITE COMMANDS						
60								
67		WEIGHED READ COMMAND PROCESSING PLUS WRITE COMMAND PROCESSING						

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 for self test results log parameters is specified in table 4.

Table 4 — Parameter control bits for self-test results log parameters

Bit	Value	Description
DU	0	Value provided by device server
DS	0	Device server supports saving of parameter
TSD	0	Device server manages saving of parameter
ETC	0	No threshold comparison is made on this value
TMC	xx	Ignored when the ETC bit is set to zero
LBIN	x	Ignored when the LP bit is set to zero
LP	0	The parameter is a list parameter

The PARAMETER LENGTH field shall contain 04h.

The NUMBER OF READ COMMANDS field contains the number of read commands (see 1.0.1.1) received by the addressed logical unit.

The NUMBER OF WRITE COMMANDS field contains the number of write commands (see 1.0.1.1) received by the addressed logical unit.

The NUMBER OF BLOCK RECEIVED field contains the number of blocks received from the service deliver subsystem for the device server of the addressed logical unit as a result of write commands (see 1.0.1.1).

The NUMBER OF BLOCK SENT field contains the number of blocks delivered to the service deliver subsystem by the device server of the addressed logical unit as a result of read commands (see 1.0.1.1).

The READ COMMAND PROCESSING INTERVALS field contains the cumulative number of time intervals (see table 6) spent processing read commands addressed to logical unit (see 1.0.1.1).

The WRITE COMMAND PROCESSING INTERVALS field contains the cumulative number of time intervals (see table 6) spent processing write commands addressed to logical unit (see 1.0.1.1).

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 logical unit (see 1.0.1.1).

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 COMMAND PROCESSING PLUS WRITE COMMAND PROCESSING field contains the cumulative weighted number of time intervals (see table 6) spent processing read commands and write commands addressed to logical unit (see 1.0.1.1).

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

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

Table 5 — Idle Time log parameter format

Bit / Byte	7	6	5	4	3	2	1	0
0	(MSB)							
1		PARAMETER CODE (0002h)						(LSB)
2	DU	DS	TSD	ETC	TMC		LBIN	LP
3	PARAMETER LENGTH (08h)							
4	(MSB)							
11		IDLE TIME INTERVALS						(LSB)

The PARAMETER CODE field set to 0009h identifies the log parameter being transferred as the Idle Time log parameter.

The values of the log parameter control bits for self test results log parameters is specified in table 4.

The PARAMETER LENGTH field shall contain 04h.

The IDLE TIME INTERVALS field contains the cumulative number of time intervals (see table 6) spent when there are no tasks in the task set and there are no tasks being processed by a logical unit.

Table 6 shows the format of the Time Interval log parameter.

Table 6 — Time Interval log parameter format

Bit / Byte	7	6	5	4	3	2	1	0
0	(MSB)							
1	PARAMETER CODE (0003h)							(LSB)
2	DU	DS	TSD	ETC	TMC		LBIN	LP
3	PARAMETER LENGTH (08h)							
4	(MSB)							
11	TIME INTERVAL DESCRIPTOR							(LSB)

The PARAMETER CODE field set to 000Bh identifies the log parameter being transferred as the Time Interval log parameter.

The values of the log parameter control bits for self test results log parameters is specified in table 14.

The PARAMETER LENGTH field shall contain 08h.

The TIME INTERVAL DESCRIPTOR (see table 7) contains the time interval 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 7 — TIME INTERVAL DESCRIPTOR **format**

Bit / Byte	7	6	5	4	3	2	1	0
0	(MSB)							
3	EXPONENT							(LSB)
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 would represent 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 $5x10^{-9}$ or 5 nsec).

1.0.1.3 Group Statistics and Performance (n) log page

The Group Statistics and Performance (n) log pages (see table 8) are subpages of the General Statistics and Performance log page (see 1.0.1.2) 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 8 — Group Statistics and Performance (n) log pages

Bit / Byte	7	6	5	4	3	2	1	0
0	PAGE CODE (xxh)							
1	SUBPAGE CODE (01h - 1Fh)							
2	(MSB)	PAGE LENGTH (34h)						
3								(LSB)
	Group Statistics and Performance log parameters							
4	Group n Statistics and Performance log parameter [a]							
55								

[a] 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 PAGE CODE and PAGE LENGTH fields are described in 7.2.1.

The SUBPAGE CODE field is as specified in table 2.

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)	PARAMETER CODE (0001h)						
1								(LSB)
2	DU	DS	TSD	ETC	TMC		LBIN	LP
3	PARAMETER LENGTH (30h)							
4	(MSB)	GROUP N NUMBER OF READ COMMANDS						
11								(LSB)
12		GROUP N NUMBER OF WRITE COMMANDS						
19								
20		GROUP N NUMBER OF BLOCKS RECEIVED						
27								
28		GROUP N NUMBER OF BLOCKS SENT						
35								
36		GROUP N READ COMMAND PROCESSING TIME						
43								
44		GROUP N WRITE COMMAND PROCESSING TIME						
51								

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 for self test results log parameters is specified in table 4.

The PARAMETER LENGTH field shall contain 04h.

The GROUP N NUMBER OF READ COMMANDS field contains the number of read commands (see 1.0.1.1) received by the addressed logical unit.

The GROUP N NUMBER OF WRITE COMMANDS field contains the number of write commands (see 1.0.1.1) received by the addressed logical unit.

The GROUP N NUMBER OF BLOCK RECEIVED field contains the number of blocks received from the service deliver subsystem for the device server of the addressed logical unit as a result of write commands (see 1.0.1.1).

The GROUP N NUMBER OF BLOCK SENT field contains the number of blocks delivered to the service deliver subsystem by the device server of the addressed logical unit as a result of read commands (see 1.0.1.1).

The GROUP N READ COMMAND PROCESSING INTERVALS field contains the cumulative number of time intervals (see table 6) spent processing read commands addressed to logical unit (see 1.0.1.1).

The GROUP N WRITE COMMAND PROCESSING INTERVALS field contains the cumulative number of time intervals (see table 6) spent processing write commands addressed to logical unit (see 1.0.1.1).