X3T9.2/87-43
Revision 1

COMMON READ LOG COMMAND
DOCUMENT NUMBER 87-43
FIRST DRAFT: 3/18/87
REVISION 1: 5/18/87

Gregory G. Floryance

IBM Corporation
Storage Products Development
System Products Division
Hwy 52 & 37 Street NW
Rochester, Mn 55901
(507) 253-1869

IBM 112

TABLE OF CONTENTS

IBM 112

1.0 INTRODUCTION

1.1 BACKGROUND

The Mode Select page 2 parameters provide the capability to tune device performance on the SCSI bus. These parameters include Buffer Full/Empty Ratios, Bus Inactivity Limit and Disconnect/Connect Time Limits. In order to determine the proper parameter values, the initiator must have information available on device usage such as the number of read/write operations performed, average transfer size, etc. In a single initiator environment, it is possible for the initiator to log most of the necessary device usage information. However, in a multi-initiator environment, the ability to monitor bus activity caused by another initiator is very complex and impractical and inefficient. In addition, some critical device usage information is only known internally by the device and can not be deduced by the initiator through bus activity regardless of the number of initiators on the bus. Therefore, a means of reporting internal device usage to an initiator is needed if the potential of the page 2 Mode Select performance tuning parameters is to be fully realized.

> Note: Bus performance tuning is not limited to development integration time. As the number and mix of devices are changed on a bus, the ability for a system to retune the operating characteristics of the attached devices (through occasional utilities or even in real time) becomes critically important. The ability for tuning a device is encouraged by the architected definition of the page 2 Mode Select parameters. However system performance tuning can not be fully utilized until an adequate reporting mechanism is architected and common parameter reporting formats are defined.

The Sequential access device READ LOG command provides the ability to report device internal information necessary for the Initiator to calculate the Mode Select parameter values. Currently, the command parameters are not defined which makes it's implementation vendor unique. In the case of DASD, the READ LOG command is not defined.

One piece of information critical to determining DASD performance characteristics is the device internal buffer overrun and underrun information. A buffer overrun on a read operation or a buffer underrun on a write operation will cause the transfer time for the command to be increased by a full revolution of the disk. Unlike read or write counters which can be maintained by a host, information on buffer over/under run counters is only available within the device itself. Therefore, a mechanism for reporting these values is necessary for DASD devices.

1.2 PROPOSAL SYNOPSIS

This proposal takes the existing Sequential Access device READ LOG command (1Fh) and makes it a common command for all device types. As currently specified, the contents of the data transferred is vendor unique with loose adherence to the CCS page concept. This proposal extends the page concept by defining page addressability within the CDB and also defines the structure of some pages which are needed to support the buffer over/under run counters in a common architected fashion.

1.3 NOTES

Note: Target or LUN addressability was considered and dropped for this command. Target/LUN addressability may be needed for devices which may have buffers in both the the target and LUNs. However, at this time, the thrust of the command is driven by mechanical activities which would effect buffer over/underruns and not the management of other resources within the target or LUN. In a wider view, the issue of Target addressability can be raised at a later time and should be address across all commands, possibly with a new message.

2.0 READ LOG COMMAND

Peripheral Device Type: All
Operation Code Type: Optional

bit byte	7	6	5	4	3	2	1	0
0	Op Code = (1Fh)							
1	LUN			PF	Reserved = 0			NLR
2	Page Code							
3 4	Allocation Length							
5	Vendor Uniq	Reserved = 0					Flag	Link

2.1 COMMAND DESCRIPTION

The READ LOG command is used to obtain statistical information maintained by the device about the device or attached medium. Logging information can also be returned to the initiator on a REQUEST SENSE command in the REQUEST SENSE data.

Log data may or may not be available to the initiator prior to the successful completion of device initialization. Device initialization is considered complete when GOOD status would be returned for the Test Unit Ready command. If a device requires a media access to obtain or clear the log data (similar to the saved parameters of MODE SENSE) and the media is not available (device is awaiting a START UNIT command), then a CHECK CONDITION shall be returned with an ILLEGAL REQUEST sense key. It is not considered an error for the device to return only a portion of the log data provided the NLR bit indicates that the log should not be cleared.

The No log Reset (NLR) bit, set to one, causes the requested log data to be transferred but not cleared. The NLR bit set to zero, causes the specified log data to be cleared after the data has been transferred. The allocation length being zero and the NLR bit zero causes the specified log data to be cleared without transferring any data to the initiator. When the NLR bit is set to zero and either PF=0 or PF=1 with Page Code 0 causes all supported log data for all pages to be cleared.

Note: A device may selectively choose to support the ability to clear all, some or none of the log data. When requested to clear a log (NLR=0), the device shall clear those portions of the log for which clearing is supported. For those portions where clearing is not supported, the request to clear shall be ignored and GOOD status shall be returned.

The Allocation length indicates the number of bytes the initiator has set aside for the READ LOG data. The target shall terminate the data transfer when either the allocation length has been transferred or when all available log data has been transferred to the initiator, whichever is less. An allocation length of zero is not considered an error.

Note: Specifying an allocation length smaller than the available log data will result in lost log data in the NLR bit is set to clear the log.

The Page Format (PF) bit of one indicates that the Read Log data conforms to page format. A value of zero indicates that the Read Log data is vendor unique and may or may not conform to page format.

The Page Code (PC) field specifies which page to return and is only valid when the PF bit is set to 1. The page codes are defined in Figure 1

Page Code	Description
00h	Summary list of supported READ LOG pages
01h	Device Specific Log Data
02	Medium Specific Log Data
03	Buffer Over/Under run counters
04	Read/Write Counters
05	Seek Length Counters
06 - 7Fh	Rsvd
80 - FFh	Vendor Unique

Figure 1. Valid Page Codes

2.2 READ LOG DATA

The format of the log data consists of a 4 byte header followed by 1 or more variable parameters. The 4 byte header is illustrated in Figure 2 on page 5. Following the 4 byte page header, are the data parameters associated with that page.

The Page Code field identifies the page of data which is being returned.

The Parameter List Length field defines the length, in bytes, of the remaining page data. The length does not include bytes zero thru three.

bit byte	7	6	5	4	3	2	1	0
0	Page Code							
1	Reserved							
2-3	Parameter List Length							

Figure 2. Page Code Header

2.3 LOG DATA PARAMETERS

The following rules apply for log parameter implementation.

After being reset, a log counter will contain the value of zero. A counter will continue to increment until the maximum value has been reached. Upon reaching the maximum value, the counter will no longer increment (will not wrap) but rather retain the maximum value as an indication that the counter has overflowed.

Parameters shall be returned in ascending order. A device may indicate non support of a parameter either by not including the parameter in the list or by returning a zero length in the parameter length field.

Parameters can either be cleared (volatile) or their values retained (non-volatile) during a power-down or a reset. Each parameters definition identifies which applies to the particular parameter.

Note: The reserved area allows for an extension to the ParmCode field. One possible use of the reserved area is to report whether the counter is cleared during a power-down or reset. Another use would be to indicate if the counter can be reset or not.

2.4 SUMMARY LIST FORMAT (PF=1, PAGE CODE=0)

bit byte	7	6	5	4	3	2	1	0
0	Page Code = 0							
1	Reserved							
2-3	Parameter List Length							
4	1st supported page number							
.	.							
.	.							
n	Last supported page number							

This page requests that a summary list of supported READ LOG pages for the device be sent to the Initiator. If any other pages are supported by the device, then support of page zero is mandatory.

Note: Typically, an initiator should first request page zero to determine the list of other supported pages within the device.

The supported page number list begins at byte 4 and contains a list of the supported pages for the device.

2.5 BUFFER OVERRUN/UNDERRUN COUNTERS PAGE (PF=1, PAGE CODE=3)

bit byte	7	6	5	4	3	2	1	0
0	Page Code = 3							
1	Reserved							
2-3	Parameter List Length (n)							

0	Reserved	ParmCode = 1	
1	Overrun Counter Length		
2	Overrun Counter		(MSB)
.			
n			(LSB)

0	Reserved	ParmCode = 2	
1	Underrun Counter Length		
2	Underrun Counter		(MSB)
.			
n			(LSB)

2.5.1 DESCRIPTION OF BUFFER OVERRUN/UNDERRUN

This page returns the device buffer overrun and underrun counters to the Initiator. A buffer overrun or underrun can occur when the initiator does not transmit the data to/from the buffer fast enough to keep up with the media. This can be caused either by a slow transfer rate on the SCSI interface or because of heavy SCSI bus usage preventing reconnection by the target.

A BUFFER OVERRUN condition occurs during a read operation when a buffer full condition prevents continued transfer of data from the media to the buffer. In the case of a disk drive, a mechanical delay is introduced equivalent to the time required for the data to reappear under the head.

A BUFFER UNDERRUN condition occurs during a write operation when a buffer empty condition prevents continued transfer of data to the media from the buffer. In the case of a disk drive, a mechanical delay is introduced equivalent to the time required for the target sector of the drive to reappear under the head.

2.5.2 FORMAT OF RETURNED DATA

For ParmCode 1 and 2, the counter will be cleared whenever a power-down or reset occurs or when the READ LOG command is issued with the NLR bit set to zero.

The OVERRUN COUNTER LENGTH and UNDERRUN COUNTER LENGTH fields indicates the length of each counter in bytes. It is recommended that the total parameter length (including the counter) be a multiple of 2. A length value of zero indicates that the counter is not supported.

The OVERRUN COUNTER and UNDERRUN COUNTER fields contain the number of times a buffer overrun or underrun has occurred since the last time the counter was reset to zero. The count will be incremented for each occurrence of an underrun or overrun condition and can be incremented more than once for multiple occurrences during the execution of a single command.

2.6 READ/WRITE COUNTERS PAGE (PF=1, PAGE CODE=4)

bit / byte	7	6	5	4	3	2	1	0
0	Page Code = 4							
1	Reserved							
2-3	Parameter List Length							

0	Reserved	ParmCode = 1	
1	Read Ops Counter Length (a)		
2	Total Read Ops Counter		(MSB)
.			
n			(LSB)

0	Reserved	ParmCode = 2	
1	Write Ops Counter Length (b)		
2	Total Write Ops Counter		(MSB)
.			
n			(LSB)

0	Reserved	ParmCode = 3	
1	Blocks Read Counter Length (c)		
2	Total Blocks Read Counter		(MSB)
.			
n			(LSB)

0	Reserved	ParmCode = 4	
1	Blocks Written Counter Length (d)		
2	Total Blocks Written Counter		(MSB)
.			
n			(LSB)

20

This page returns read/write usage counters to the Initiator.

For ParmCodes 1 thru 4, the counter will be cleared whenever a power-down or reset occurs or when the READ LOG command is issued with the NLR bit set to zero. These counters are only intended to reflect data transfer operations between the device and the initiator. These counters should not be effected by such operations as error recovery, data recovery, reassignment algorithms, etc.

2.7 SEEK LENGTH COUNTERS PAGE (PF=1, PAGE CODE=5)

bit / byte	7	6	5	4	3	2	1	0
0	Page Code = 5							
1	Reserved							
2-3	Parameter List Length							

0	Reserved	ParmCode = 1
1	Zero Length Seek Counter Length (a)	
2	Total Zero Length Seeks Counter	(MSB)
.		
n		(LSB)

0	Reserved	ParmCode = 2
1	Multi-Cyl Seek Counter Length (b)	
2	Cyls per Seek Counter (c)	(MSB)
3		(LSB)
4	Seek Group 1 Counter	(MSB
.		
.		(LSB)
.	.	.
.	.	.
.	.	.
.	Seek Group (x) Counter	(MSB
.		
n		(LSB)

This page returns information on seek lengths to the Initiator. A seek is considered a physical movement of the access mechanism.

2.7.1 ZERO SEEK LENGTH COUNTER

The Number Of Zero Length Seeks field indicates that no seek was required for the command execution. This is the case where the logical block is on the same physical track or the same cylinder such that only a head switch is required.

2.7.2 MULTI-CYL SEEK COUNTER

The Multi_Cyl Seek Counter Length field indicates the number of bytes for each seek counter.

The Cyls Per Seek Counter field defines the seek distance (in cylinders) associated with each seek group counter. The cyl per seek counter value is multiplied by the seek group counter number to calculate the high end of the range value. The low end of the range is defined by the high end value of the previous seek group counter range + 1. For seek group 1, the low end is zero.

The Seek Group (x) Counter field indicates the number of seeks which have occurred within that range of seek lengths.

For ParmCodes 1 and 2, the counter will be cleared whenever a power-down or reset occurs or when the READ LOG command is issued with the NLR bit set to zero.