Proposed **Serial ATA**
Draft **International Organization**

Version 05

January 28, 2013

TP_043v5_SATA31_Queuing Power Management

Title: Queuing Power Management

Proposed change, new functionality, or behavior to Serial ATA Revision 3.1

This is an internal working document of the Serial ATA International Organization. As such, this is not a completed standard and has not been approved. The Serial ATA International Organization may modify the contents at any time. This document is made available for review and comment only.

Author Information

Author Name	Company	e-mail address
Curtis E. Stevens	Western Digital	Curtis.Stevens@wdc.com

Workgroup Chair Information

Workgroup	Chairperson Name	e-mail address
Digital	Jim Hatfield	james.c.hatfield@seagate.com

Document History

Version	Date	Comments
00	2-Oct-2011	Initial draft
01	5-Dec-11	Updated with Feedback from conference calls Added NCQ NON-DATA for SET FEATURES commands
02	19-Mar-12	Accepted changes from previous version Added a note indicating that all set features are not appropriate when they are queued.
03	2-May-12	Added updates from 19-Mar telecon
04	4-May-12	Editor assigned numbers to clear TBD, added section numbers, and other editorial integration.
05	4-Feb-13	Ratified

1 Introduction

Storage technology has been evolving and has migrating to use queuing capability for read and write commands in a SATA connection. Unfortunately, SATA queuing only encompasses reading and writing user data. Any other operation requires the queue to be empty before continuing. This impacts system performance. There is a desire to manage mechanical parts in a hybrid device and also to monitor the devices condition. The EPC feature set as it is currently defined provides this capability.

2 Summary of the problem

There is a desire to issue EPC feature set commands, read log commands, and write log commands without emptying the queue. In hybrid technology, these functions need to be performed without impacting system performance.

3 Summary of the solution

Additions are proposed to add:

1) WRITE LOG DMA EXT to RECEIVE FPDMA QUEUED;
2) READ LOG DMA EXT to SEND FPDMA QUEUED; and
3) SET FEATURES to NCQ NON-DATA.

4 Proposed changes

Editor's note 1 – SATA rev 3.1 see 13.6.6.

1.1.1[Editor's note 13.6.6] RECEIVE FPDMA QUEUED

1.1.1.1[Editor's note 13.6.6.1] RECEIVE FPDMA QUEUED definition

The 512 Byte Block DMA IN subcommands make use of this transport command. The RECEIVE FPDMA QUEUED command supports LBA mode only and uses 48-bit addressing only. The format of the command is defined in ~~Figure 252~~Figure 242.

1.1.1.1 ~~<13.6.6.1>~~[Editor's note 13.6.6.2] Inputs

Register	7	6	5	4	3	2	1	0
Features~~-~~(7:0)	Sector Count~~-~~(7:0)							
Features~~-~~(15:8)	Sector Count~~-~~(15:8)							
Count~~-~~(7:0)	TAG					Reserved		
Count~~-~~(15:8)	PRIO~~-~~(1:0)		Res	Subcommand				
LBA~~-Low-~~(7:0)	LBA~~-~~(7:0)							
LBA~~-Low (31:24)~~(15:8)	LBA~~-31:24~~(15:8)							
LBA~~-Mid (15:8)~~(23:16)	LBA~~-15:8~~(23:16)							
LBA~~-Mid (39:32)~~(31:24)	LBA~~-39:32~~(31:24)							
LBA~~-High (23:16)~~(39:32)	LBA~~-23:16~~(39:32)							
LBA~~-High-~~(47:40)	LBA~~-~~(47:40)							
ICC(7:0)	ICC~~-~~(7:0)							
Auxiliary(7:0)	Auxiliary~~-~~(7:0)							
Auxiliary~~-~~(15:8)	Auxiliary~~-~~(15:8)							
Device	Res	1	Res	0	Reserved			
Command	65h							

Figure 242 – RECEIVE FPDMA QUEUED command definition

Sector Count The number of 512 byte blocks to be transferred, 0000h indicates that 65~~,~~ 536

512 byte blocks are to be transferred.

TAG The TAG value shall be assigned by host software to be different from all other TAG values corresponding to outstanding commands. The assigned TAG value shall not exceed the value specified in IDENTIFY DEVICE word 75.

PRIO The Priority (PRIO) value shall be assigned by the host based on the priority of the command issued. The device shall make a best effort to complete High priority requests in a more timely fashion than Normal and Isochronous priority requests. The device shall make a best effort to complete Isochronous priority requests prior to its associated deadline.

The Priority values are defined as follows:

a) 00b Normal Priority;

b) 01b Isochronous – deadline dependent priority;

c) 10b High priority; and

d) 11b Reserved.

Subcommand

~~The value of this field is Reserved.~~See [Editor's note 13.6.7.5].~~x.x.x~~

LBA ~~The value of this field is Reserved.~~See [Editor's note 13.6.7.5].~~x.x.x~~

ICC The Isochronous Command Completion (ICC) field shall be assigned by the host based on the intended deadline associated with the command issued. By default, ~~when~~if deadline is expired, the device shall continue to complete the command as soon as possible.

Auxiliary ~~The value of this field is Reserved.~~See [Editor's note 13.6.7.5].~~x.x.x~~

Upon accepting the command, the device shall clear the BSY bit by transmitting a Register Device to Host FIS to the host with the BSY bit cleared to zero in the Status field of the FIS. The ability for the device to quickly clear the BSY bit allows the host to issue another queued command without blocking on this bit.

The host shall check the BSY bit in the shadow Status register before attempting to issue a new command in order to determine that the device is ready to receive another command (and determine that the host has write access to the Shadow Register Block Registers). The device shall not trigger an interrupt in response to having successfully received the command, so the initial status return that clears BSY shall not have an interrupt associated with it.

1.1.1.2 [Editor's note 13.6.6.2] Success Outputs

Upon successful completion of one or more outstanding commands, the device shall transmit a Set Device Bits FIS with the Interrupt bit set to one and one or more bits set to one in the ACT field corresponding to the bit position for each command TAG that has completed since the last status notification was transmitted. The ERR bit in the Status register shall be cleared to zero and the value in the Error register shall be zero.

The ACT field occupies the last 32 bits of the Set Device Bits FIS as defined in Figure 243.

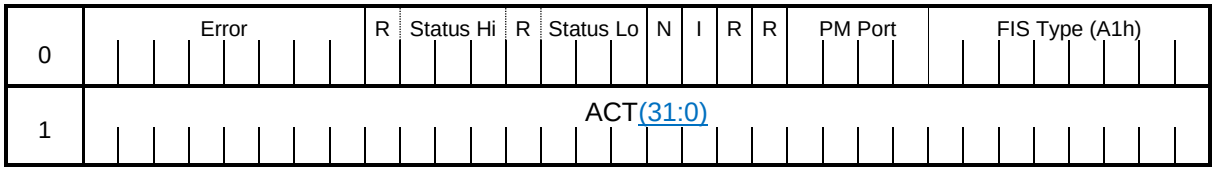

Figure 243 – Set Device Bits FIS for successful RECEIVE FPDMA QUEUED command completion

ACT The ACT field of the Set Device Bits FIS communicates successful completion notification for each of up to 32 queued commands. The field is bit-significant and the device sets bit positions to one for each command tag it is indicating successful completion notification for. The device may set more than one bit to one if it is explicitly aggregating successful status returns.

Error	The Error register shall be cleared to zero.
Status	As defined in ~~section~~ [Editor's note 10.3.6]. The ERR bit shall be cleared to zero indicating successful command completion. Bit 4 may be set to one.
I	Interrupt bit. The interrupt bit shall be set to one.

All other fields as defined in ~~section~~ [Editor's note 10.3.6].

Devices should be aware that if choosing to aggregate status to the point where many of the outstanding commands have actually completed successfully without notification to the host, that an error may cause the final completion status of those commands to be failure. The device should be selective ~~when~~if using status aggregation for outstanding queued commands to ensure the host is made aware of successful completion for outstanding commands in a way that an error ~~would not~~be unable to force a high number of unnecessary command retries.

1.1.1.3 [Editor's note 13.6.6.3] Error Outputs

1.1.1.3.1 [Editor's note 13.6.6.3.1] Upon Receipt of a Command

If the device has received a command that has not yet been acknowledged by clearing the BSY bit to zero and an error is encountered, the device shall transmit a Register Device to Host FIS (see Figure 244) with the ERR bit set to one and the BSY bit cleared to zero in the Status field, the ATA error code in the Error field.

Register	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**
Error	ERROR							
Count ~~-~~(7:0)	na							
Count ~~-~~(15:8)	na							
LBA ~~Low~~ (7:0)	na							
LBA ~~Low (31:24)~~(15:8)	na							
LBA ~~Mid (15:8)~~(23:16)	na							
LBA ~~Mid (39:32)~~(31:24)	na							
LBA ~~High (23:16)~~(39:32)	na							
LBA ~~High~~ (47:40)	na							
Device	na							
Status	BSY	DRDY	DF	na	DRQ	na	na	ERR

Figure 244 – RECEIVE FPDMA QUEUED error status result values on command receipt

ERROR ATA error code for the failure condition of the failed command

BSY 0

DRDY 1

DF 0

DRQ 0

ERR 1

Status bit 4 may be set to one.

Following transmission of the Register Device to Host FIS, the device shall stop processing any outstanding or new commands until the Queued Error Log (see [Editor's note 13.7.3]) has been read before continuing to abort all outstanding commands. See [Editor's note 13.6.3.3] for more details.

1.1.1.3.2 During Execution of a Command

If all commands have been acknowledged by clearing the BSY bit to zero and an error condition is detected, the device shall transmit a Set Device Bits FIS (see Figure 245) to the host with the ERR bit set to one in the Status field, the ATA error code in the Error field, and the Interrupt bit set to one. All outstanding commands at the time of an error shall be aborted as part of the error response and may be re-issued as appropriate by the host. For any commands that have not

completed successfully or have resulted in error, the device shall clear the corresponding ACT bits to zero in the Set Device Bits FIS.

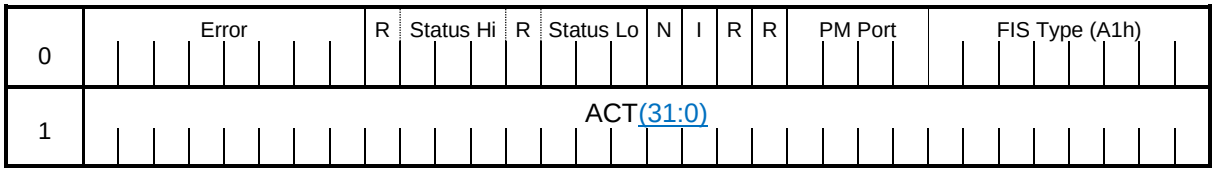

Figure 245 – Set Device Bits FIS with error notification, and command completions

ACT — The ACT field of the Set Device Bits FIS communicates successful completion notification for each of up to 32 queued commands. The field is bit-significant and the device sets bit positions to one for each command tag it is indicating successful completion notification for. The device may set more than one bit to one if it is explicitly aggregating successful status returns.

Error — The Error register shall contain the ATA error code.

Status — As defined in [Editor's note 10.5.7]. The ERR bit shall be set to one indicating an NCQ error has occurred. Status bit 4 may be set to one.

I — Interrupt bit. The interrupt bit shall be set to one.

All other fields as defined in ~~section~~ [Editor's note 10.5.7].

Only the registers that are updated as part of the Set Device Bits FIS are modified if the device signals an error condition ~~when~~if the BSY bit in the shadow Status register is cleared to zero, leaving the other Shadow Register Block Registers unchanged. If the device signals an error condition ~~when~~if the BSY bit in the shadow Status register is set to one, the device clears the BSY bit to zero with a Register Device to Host FIS which updates all registers in the Shadow Register Block.

Following transmission of the Set Device Bits FIS, the device shall stop processing any outstanding or new commands until the Queued Error Log (see [Editor's note 13.7]) has been read before continuing to abort all outstanding commands. See [Editor's note 13.6.4.4] for more details.

13.6.6 [Editor's note 13.6.7] SEND FPDMA QUEUED

[Editor's note 13.6.7.1] SEND FPDMA QUEUED definition

The 512 Byte Block DMA OUT subcommands make use of this transport command. The SEND FPDMA QUEUED command supports LBA mode only and uses 48-bit addressing only. The format of the command is defined in ~~Figure 257~~Figure 246.

1.1.1.4 ~~<13.6.7.1>~~[Editor's note 13.6.7.2] Inputs

Register	7	6	5	4	3	2	1	0
Features~~-~~(7:0)	Sector Count~~-~~(7:0)							
Features~~-~~(15:8)	Sector Count~~-~~(15:8)							
Count~~-~~(7:0)	TAG					Reserved		
Count~~-~~(15:8)	PRIO~~-~~(1:0)		Res	Subcommand				
LBA~~-Low-~~(7:0)	LBA~~-~~(7:0)							
LBA~~-Low (31:24)~~(15:8)	LBA~~-31:24~~(15:8)							
LBA~~-Mid (15:8)~~(23:16)	LBA~~-15:8~~(23:16)							
LBA~~-Mid (39:32)~~(31:24)	LBA~~-39:32~~(31:24)							
LBA~~-High (23:16)~~(39:32)	LBA~~-23:16~~(39:32)							
LBA~~-High-~~(47:40)	LBA~~-~~(47:40)							
ICC(7:0)	ICC~~-~~(7:0)							
Auxiliary(7:0)	Auxiliary~~-~~(7:0)							
Auxiliary~~-~~(15:8)	Auxiliary~~-~~(15:8)							
Device	Res	1	Res	0	Reserved			
Command	**64h**							

Figure 246 – SEND FPDMA QUEUED command definition

Sector Count The number of 512 byte blocks to be transferred, 0000h indicates that 65~~,~~ 536

512 byte blocks are to be transferred.

TAG The TAG value shall be assigned by host software to be different from all other TAG values corresponding to outstanding commands. The assigned TAG value shall not exceed the value specified in IDENTIFY DEVICE word 75.

PRIO The Priority (PRIO) value shall be assigned by the host based on the priority of the command issued. The device shall make a best effort to complete High priority requests in a more timely fashion than Normal and Isochronous priority requests. The device shall make a best effort to complete Isochronous priority requests prior to its associated deadline.

The Priority values are defined as follows:

a) 00b Normal Priority;

b) 01b Isochronous – deadline dependent priority;

c) 10b High priority; and

d) 11b Reserved.

Subcommand See [Editor's note 13.6.7.4] 1.1.1.7

LBA Subcommand dependent. See [Editor's note 13.6.7.4] 1.1.1.7 for subcommand definitions.

ICC The Isochronous Command Completion (ICC) field shall be assigned by the host based on the intended deadline associated with the command issued. By default, ~~when~~if deadline is expired, the device shall continue to complete the command as soon as possible.

Auxiliary Subcommand dependent. See [Editor's note 13.6.7.4] 1.1.1.7 for subcommand definitions.

Upon accepting the command, the device shall clear the BSY bit by transmitting a Register Device to Host FIS to the host with the BSY bit cleared to zero in the Status field of the FIS. The ability for the device to quickly clear the BSY bit allows the host to issue another queued command without blocking on this bit.

The host shall check the BSY bit in the shadow Status register before attempting to issue a new command in order to determine that the device is ready to receive another command (and determine that the host has write access to the Shadow Register Block Registers).

The device shall not trigger an interrupt in response to having successfully received the command, so the initial status return that clears BSY shall not have an interrupt associated with it.

1.1.1.5 ~~<13.6.7.2>~~[Editor's note 13.6.7.3] Success Outputs

Upon successful completion of one or more outstanding commands, the device shall transmit a Set Device Bits FIS with the Interrupt bit set to one and one or more bits set to one in the ACT field corresponding to the bit position for each command TAG that has completed since the last status notification was transmitted. The ERR bit in the Status register shall be cleared to zero and the value in the Error register shall be zero.

The ACT field occupies the last 32 bits of the Set Device Bits FIS as defined in Figure 247.

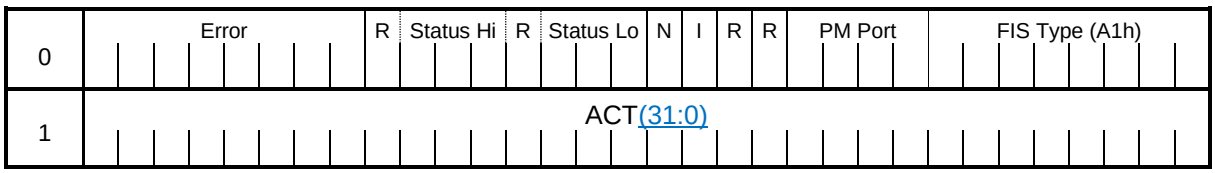

Figure 247 – Set Device Bits FIS for successful SEND FPDMA QUEUED command completion

ACT The ACT field of the Set Device Bits FIS communicates successful completion notification for each of up to 32 queued commands. The field is bit-significant and the device sets bit positions to one for each command tag it is indicating

successful completion notification for. The device may set more than one bit to one if it is explicitly aggregating successful status returns.

Error — The Error register shall be cleared to zero.

Status — As defined in ~~section~~ [Editor's note 10.3.6]. The ERR bit shall be cleared to zero indicating successful command completion. Bit 4 may be set to one.

I — Interrupt bit. The interrupt bit shall be set to one.

All other fields as defined in ~~section~~ [Editor's note 10.3.6].

Devices should be aware that if choosing to aggregate status to the point where many of the outstanding commands have actually completed successfully without notification to the host, that an error may cause the final completion status of those commands to be failure. The device should be selective ~~when~~if using status aggregation for outstanding queued commands to ensure the host is made aware of successful completion for outstanding commands in a way that an error ~~would not~~be able to force a high number of unnecessary command retries.

1.1.1.6 ~~<13.6.7.3>~~[Editor's note 13.6.7.4] Error Outputs

13.3.1.13.1 [Editor's note 13.6.7.4.1] Error Outputs status

If the device has received a command that has not yet been acknowledged by clearing the BSY bit to zero and an error is encountered, the device shall transmit a Register Device to Host FIS (see Figure 248) with the ERR bit set to one and the BSY bit cleared to zero in the Status field, the ATA error code in the Error field.

Register	7	6	5	4	3	2	1	0
Error	ERROR							
Count ~~-~~(7:0)	na							
Count ~~-~~(15:8)	na							
LBA ~~Low~~ (7:0)	na							
LBA ~~Low (31:24)~~(15:8)	na							
LBA ~~Mid (15:8)~~(23:16)	na							
LBA ~~Mid (39:32)~~(31:24)	na							
LBA ~~High (23:16)~~(39:32)	na							
LBA ~~High~~ (47:40)	na							
Device	na							
Status	BSY	DRDY	DF	na	DRQ	na	na	ERR

Figure 248 – SEND FPDMA QUEUED error status result values on command receipt

ERROR ATA error code for the failure condition of the failed command

BSY 0

DRDY 1

DF 0

DRQ 0

ERR 1

Status bit 4 may be set to one.

Following transmission of the Register Device to Host FIS, the device shall stop processing any outstanding or new commands until the Queued Error Log (see [Editor's note 13.7.3]) has been read before continuing to abort all outstanding commands. See [Editor's note 13.6.3.3] for more details.

1.1.1.6.1 ~~<13.6.7.3.1>~~[Editor's note 13.6.7.4.2]During Execution of a Command

If all commands have been acknowledged by clearing the BSY bit to zero and an error condition is detected, the device shall transmit a Set Device Bits FIS (see [Editor's note Figure 260]) to the host with the ERR bit set to one in the Status field, the ATA error code in the Error field, and the Interrupt bit set to one. All outstanding commands at the time of an error shall be aborted as part of the error response and may be re-issued as appropriate by the host. For any commands that have not completed successfully or have resulted in error, the device shall clear the corresponding ACT bits to zero in the Set Device Bits FIS.

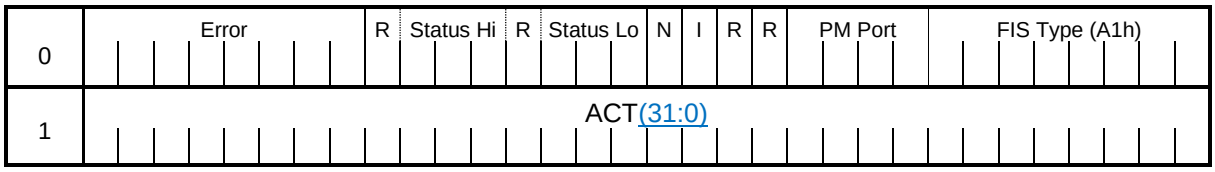

Figure 249 – Set Device Bits FIS with error notification, and command completions

ACT — The ACT field of the Set Device Bits FIS communicates successful completion notification for each of up to 32 queued commands. The field is bit-significant and the device sets bit positions to one for each command tag it is indicating successful completion notification for. The device may set more than one bit to one if it is explicitly aggregating successful status returns.

Error — The Error register shall contain the ATA error code.

Status — As defined in [Editor's note 10.3.6]. The ERR bit shall be set to one indicating an NCQ error has occurred. Status bit 4 may be set to one.

I — Interrupt bit. The interrupt bit shall be set to one.

All other fields as defined in section [Editor's note 10.3.6].

Only the registers that are updated as part of the Set Device Bits FIS are modified if the device signals an error condition ~~when~~if the BSY bit in the shadow Status register is cleared to zero, leaving the other Shadow Register Block Registers unchanged. If the device signals an error condition ~~when~~if the BSY bit in the shadow Status register is set to one, the device clears the BSY bit to zero with a Register Device to Host FIS which updates all registers in the Shadow Register Block.

Following transmission of the Set Device Bits FIS, the device shall stop processing any outstanding or new commands until the Queued Error Log (see [Editor's note 13.7.3]) has been before continuing to abort all outstanding commands. See [Editor's note 13.6.3.3.] for more details.

1.1.1.7 ~~<13.6.7.4>~~[Editor's note 13.6.7.5] SEND FPDMA QUEUED Subcommands

Subcommands for the SEND FPDMA QUEUED commands are contained within the Count (13:8) field. The allowed values are defined in Table 92.

Table 92 – Subcommands for SEND FPDMA QUEUED

Value	Subcommand
00h	Data Set Management (see ~~section 10.3.6.1~~ [Editor's note 13.6.7.5])
01h	Editor's Note see TPR042

02h	WRITE LOG DMA EXT (see [Editor's note 13.6.7.new])
~~01h-~~ 03h..1Fh	Reserved

1.1.1.8 [Editor's note 13.6.7.new]WRITE LOG DMA EXT

1.1.1.9 [Editor's note 13.6.7.new.1]WRITE LOG DMA EXT overview

The WRITE LOG DMA EXT subcommand functionality and behavior is dependent on all requirements of the WRITE LOG DMA EXT command and the IDENTIFY DEVICE command defined in ACS-2.

1.1.1.9.1 [Editor's note 13.6.7.new.2] Inputs

Register	7	6	5	4	3	2	1	0
Auxiliary(7:0)	Reserved							
Auxiliary(15:8)	Reserved							

Figure [Editor's note 249+1] – SEND FPDMA QUEUED, Subcommand = 02h

The Page Count field of the WRITE LOG DMA EXT command (see ACS-3) is the sector count placed in the feature field of the SEND FPDMA QUEUED command.

See ACS-2 for the definition of LBA(47:0) of the WRITE LOG DMA EXT command.

1.1.1.9.2 [Editor's note 13.6.7.new.3]Success Outputs

See 1.1.1.5[Editor's note 13.6.7.2]

1.1.1.9.3 [Editor's note 13.6.7.new.4]Error Outputs

See 1.1.1.6[Editor's note 13.6.7.3]

1.1.1.10 [Editor's note 13.6.7+1]RECEIVE FPDMA QUEUED Subcommands

Subcommands for the RECEIVE FPDMA QUEUED commands are contained within the Count(13:8) field. The allowed values are defined in [Editor's note Table 92+1].

Table [Editor's note 92+1] – Subcommands for RECEIVE FPDMA QUEUED

Value	Subcommand
00h	Reserved
01h	READ LOG DMA EXT (see [Editor's note 13.6.7+2]~~x.x.x.~~)
02h..1Fh	Reserved

1.1.1.11 [Editor's note 13.6.7+2]READ LOG DMA EXT

13.3.1.13.2 [Editor's note 13.6.7+2.1]READ LOG DMA EXT overview

The READ LOG DMA EXT subcommand functionality and behavior is dependent on all requirements of the READ LOG DMA EXT command and the IDENTIFY DEVICE command defined in ACS-2.

1.1.1.11.1 [Editor's note 13.6.7+2.2]Inputs

Register	7	6	5	4	3	2	1	0
Auxiliary(7:0)	Reserved							
Auxiliary(15:8)	Reserved							

Figure [Editor's note 249+2] – RECEIVE FPDMA QUEUED, Subcommand = 01h

The Page Count field of the READ LOG DMA EXT command (see ACS-3) is the sector count placed in the feature field of the RECEIVE FPDMA QUEUED command.

The feature field in the READ LOG DMA EXT command shall be set to zero.

See ACS-2 for the definition of LBA(47:0) of the READ LOG DMA EXT command.

1.1.1.11.2 [Editor's note 13.6.7+2.3]Success Outputs

See 1.1.1.5 [Editor's note 13.6.7.2]

1.1.1.11.3 [Editor's note 13.6.7+2.4]Error Outputs

See 1.1.1.6 [Editor's note 13.6.7.3]

1.1.1.11.4 [Editor's note 13.6.7+2.5]NCQ NON-DATA Subcommands

Table [Editor's note 92+2] – Subcommand Field

Subcommand	Description	Reference
0h	Abort NCQ queue	
1h	Deadline Handling	
5h~~TBDx~~	SET FEATURES	

1.1.1.12 [Editor's note 13.6.7+2.6]SET FEATURES

13.3.1.13.3 [Editor's note 13.6.7+2.6.1]SET FEATURES overview

The SET FEATURES subcommand functionality and behavior is dependent on all requirements of the SET FEATURES command and the IDENTIFY DEVICE command defined in ACS-2.

Note 58+1 - **Some SET FEATUREs subcommands may return command aborted if they are issued at a time when the device is unable to process them (e.g., changing the non-zero buffer offsets setting while commands are in the queue).**

1.1.1.12.1[Editor's note 13.6.7+2.6.2]Inputs

Register	7	6	5	4	3	2	1	0
Auxiliary(7:0)	Reserved							
Auxiliary(15:8)	Reserved							

Figure [Editor's note 249+3] – SET FEATURES QUEUED, Subcommand = 2h~~TBDx~~

The Feature field of the SET FEATURES command is placed in the FEATURES(15:8) field of the NCQ NON-DATA command.

The Count field of the SET FEATURES command is placed in the Count (15:8) field of the NCQ NON-DATA command.

See ACS-2 for the definition of LBA (27:0) of the SET FEATURES command. LBA (47:28) is reserved.

All other subcommand specific fields in the NCQ NON-DATA command are reserved.

1.1.1.12.2[Editor's note 13.6.7+2.6.3]Success Outputs

See 1.1.1.5[Editor's note 13.6.7.2]

1.1.1.12.3[Editor's note 13.6.7+2.6.4]Error Outputs

See 1.1.1.6 [Editor's note 13.6.7.3]