SATA IO NCQ

November 14, 2006

Revision 4

e06150r4-SATA_IO_NCQ.fm

Technical Editor:

Curtis E. Stevens
20511 Lake Forest Dr.
Lake Forest, Ca. 92630
949-672-7933
Curtis.Stevens@wdc.com

Abstract: ...

Document Status

Revision History		
Rev	**Date**	**Description**
1	July 27, 2006	1) Initial Revision
2	October 22, 2006	1) Made edits based on meeting feedback. 2) Added definitions of Media, Cache and Non-volatile cache 3) Reworded command definition 4) Reworded FUA definition. There is still considerable disagreement on how FUA interracts with Non-volatile cache. 5) Agreed on the term Transport Specific Indicator in place of references to SDB FIS. This reference needs to be defined in the ATA8-AST doc. In theory, it should also be defined in ATA8-APT as well since this is being defined in a transport independent fashion. 6) Reworded error outputs and left it to the feature set to describe that the queue is aborted on errors.
3	November 3, 2006	1) Updated feature set to remove references to transport specific bits.
4	November 14, 2006	1) Changes based on meeting 14-Nov-2006

1 Introduction

This clause describes the background and the need for this proposal

e06130 provides guidelines for creating a proposal, but it can not really be used as a template. We are getting proposals that do not fully cover the changes required for the ATA/ATAPI Command Set (ACS). The new proposal template should be used to ensure that all the appropriate areas of ACS are updated when a new proposal is approved.

2 Scope

This clause describes the statement of purpose.

This purpose of this template is to facilitate the creation and integration of proposals into ATA8-ACS. This document contains the paragraph styles, fonts, table formatting and other items necessary for creating a complete proposal

3 Overview

This clause provides a high level description of the proposal.

The ACS document template contains all the necessary items for creating a proposal for an addition to ACS. The assumption for this proposal is that a new command or feature set is being added. All of the clauses listed here may not need to be changed. In that case, delete the subclause from the proposal.

4 Changes to ACS

4.1 Changes to clause 2

4.1.1 Changes to Approved ANSI References

Replace the references listed with the new references for this proposal.

Name	Reference
Protected Area Run Time Interface Extensions (PARTIES)	ANSI INCITS 346-2001
SCSI Primary Commands - 3 (SPC-3)	ANSI INCITS 408-2005 ISO/IEC 14776-453

4.1.2 Changes to ANSI References Under Development

Replaces the references listed below with the new references for this proposal.

Name	Project Number
AT Attachment-8 - Parallel Transport (ATA8-APT)	INCITS 1698D ISO/IEC 14776-881
AT Attachment-8 – ATA/ATAPI Architecture Model (ATA8-AAM)	INCITS 1700D ISO/IEC 14776-861

4.1.3 Changes to Other References

Replace the references list below with the new references for this proposal.

PC Card Standard, February 1995, PCMCIA (68-pin Connector)

For the PC Card Standard published by the Personal Computer Memory Card International Association, contact PCMCIA at 408-433-2273 or http://www.pc-card.org .

CompactFlash™ Association Specification, Revision 4.0

For the CompactFlash™ Association Specification published by the CompactFlash™ Association, contact the CompactFlash™ Association at http://www.compactflash.org.

4.2 Changes to clause 3

4.2.1 Changes to Definitions and abbreviations

4.2.2 **Media:** The material on which data is stored (e.g., a magnetic disk) that is accessible by application clients.

4.2.3 **Cache:** A temporary data storage area outside the area accessible by application clients that may contain a subset of the data stored on the media.

4.2.4 **non-volatile cache:** Cache that retains data through power cycles.

4.2.5 **Queued Command:** A NCQ command that has reported command acceptance but not command completion.

Editor's Note 1: Change the TCQ feature set to refer to TCQ command in place of queued command.

4.3 Changes to clause 4

4.3.1 Overview - Changes to the Feature Set Summary Table

Replace the names of the feature sets in the table below with the new feature set names.

Feature set	General Devices	Packet Devices
Native Command Queuing (NCQ) (see 4.3.4)	O	P
Your feature set (see TBD)	O	O
Key: M – Mandatory, O – Optional, P – Prohibited, N – Not defined		

4.3.2 Security Mode feature fet - Changes to the Security Mode Command Actions Table

Replace the command names with the new command names in this proposal.

Command	Locked	Unlocked	Frozen
READ FPDMA QUEUED	Command aborted	Executable	Executable
WRITE FPDMA QUEUED	Command aborted	Executable	Executable

4.3.3 Long Logical Sector feature set - Block Size By Command Table

Replace the command names with the new command names in this proposal.

Command	Words Transferred
READ FPDMA QUEUED	IDENTIFY DEVICE data words (118:117)
WRITE FPDMA QUEUED	IDENTIFY DEVICE data words (118:117)

4.3.4 Native Command Queuing (NCQ) feature set

4.3.4.1 Overview

The NCQ feature set allows commands within the feature set to be accepted even though one or more previously accepted NCQ commands have not reported command completion. NCQ commands indicate command completion by returning a transport dependent indicator, see ATA8-AST for more information. The following commands are mandatory for devices implementing the NCQ feature set:

a) READ FPDMA QUEUED
b) WRITE FPDMA QUEUED

Devices which report support for the NCQ feature set shall also report support for the General Purpose Logging feature set and log addresses 00h and 10h.

Editor's Note 2: This has implications for DCO... If you disable GPL, NCQ shall be disabled as well.

All the commands in the NCQ feature set shall include an NCQ Tag. If the value of the NCQ Tag exceeds the value returned in IDENTIFY DEVICE data word 75xxx, then the device shall return command aborted for the new command and shall return command aborted for all the queued commands. If a NCQ command is issued with a NCQ Tag value that is identical to the NCQ Tag value for a queued command, then the device shall return command aborted for the new command and shall return command aborted for all the queued commands.

NOTE 1 — The NCQ Tag is used to identify return information (i.e., error status, data transfer and command completion).

If an error occurs while there are queued commands, the device shall abort all queued commands and shall complete any new commands (with the exception of READ LOG EXT or READ LOG DMA EXT requesting log address 10h) it receives with an error, until the device completes a READ LOG EXT or READ LOG DMA EXT command requesting log address 10h.

Editor's Note 3: READ LOG DMA EXT needs to be addressed by SATA IO

4.3.4.2 Command Phases

4.3.4.2.1 Command Acceptance

The device receives a NCQ command and returns command acceptance. Once the device reports command acceptance, it may then accept additional NCQ commands.

4.3.4.2.2 Data transmission

Data transfer should occur after acceptance of the command. The mechanism for data transfer is transport dependent.

4.3.4.3 Command completion

When the transfer of all requested data has occurred without error, the device returns a transport dependent indicator which informs the host that one or more command have been completed.

If an error occurs then the device shall return command aborted for the new command and shall return command aborted for all the queued commands. The condition of the data for any queued command that reports command aborted is indeterminate.

4.4 Changes to clause 6

4.4.1 Status Bits

4.4.1.1 Transport Dependent

All bits and fields that are labelled transport dependent are defined in the transport standards. For example, ATA/ATAPI-7 defines the status bits BSY, DRDY, and DRQ. These bits are documented in the transport standards. Although all of the commands in this standard use BSY=0, DRDY=1 and DRQ=0 to specify that the device is ready to accept a command and to specify that a command is complete, they are processed differently in the various transport standards.

4.4.2 Interrupt Reason bits

4.4.2.1 NCQ Tag

Count bits (7:3). If the device supports NCQ, this field contains the NCQ Tag value for the command. A NCQ Tag value may be any value that does not exceed IDENTIFY DEVICE data word 75xxx.

4.5 Changes to clause 7

4.5.1 Command Definition

4.5.1.1 READ FPDMA QUEUED - 60h, DMA Queued

4.5.1.2 Feature Set

This command is mandatory for devices implemeting the NCQ feature set (see feature set reference).

4.5.1.3 Description

This command requests that data to be transferred from the device to the host.

4.5.1.4 Inputs

4.5.1.4.1 Overview

<table>
<tr><th>Word</th><th>Name</th><th>Description</th></tr>
<tr><td>00h</td><td>Feature</td><td>The number of logical sectors to be transferred. A value of 0000h indicates that 65,536 logical sectors are to be transferred.</td></tr>
<tr><td>01h</td><td>Count</td><td>Bit Description
15:8 Reserved
7:3 NCQ Tag - See (6.3.4xxxx)
2:0 Reserved</td></tr>
<tr><td>02h</td><td rowspan="3">LBA</td><td rowspan="3">(MSB)
Address of the first logical sector to be transferred.
(LSB)</td></tr>
<tr><td>03h</td></tr>
<tr><td>04h</td></tr>
<tr><td>05h</td><td>Device</td><td>Bit Description
15 FUA - See xxx(below)
14 Shall be set to one
13 Reserved
12 Shall be set to zero
11:8 Reserved</td></tr>
<tr><td></td><td>Command</td><td>7:0 60h</td></tr>
</table>

4.5.1.4.2 Forced Unit Access (FUA)

When the FUA bit is set to one the device shall retrive the data from the media regardless of whether the device holds the requested information in its cache. If the device holds a modified copy of the requested data as a result of having cached writes, the modified data shall be written to the media before being retrieved from the media as part of this operation. When the FUA bit is cleared to zero the data shall be retrieved either from the device's media or cache.

Editor's Note 4: The FUA definition may change due to clarifications on the relationship of Cache, NV Cache, and media.

4.5.1.5 Command Acceptance Outputs

See table yy

4.5.1.6 Normal Outputs

See table mm.

Editor's Note 5: How do we document the SDB FIS?

4.5.1.7 Error Outputs

This return indicates that the command was aborted due to LBA out of range, a duplicate tag number, an invalid tag number, or an ICRC error, see table zz for more information. Additionally, errors which occur during the processing of this command are reported by returning a transport dependent indicator with additional information available in log address 10h, see table nn for more information.

4.5.1.8 WRITE FPDMA QUEUED - 61h, DMA Queued

4.5.1.9 Feature Set

This command is mandatory for devices implemeting the NCQ feature set

4.5.1.10 Description

This command causes data to be transferred from the host to the device.

4.5.1.11 Inputs

4.5.1.11.1 Overview

Word	Name	Description
00h	Feature	The number of logical sectors to be transferred. A value of 0000h indicates that 65,536 logical sectors are to be transferred.
01h	Count	**Bit** **Description** 15:8 Reserved 7:3 NCQ Tag - See (6.3.4xxxx) 2:0 Reserved
02h	LBA	(MSB)
03h		Address of the first logical sector to be transferred.
04h		(LSB)
05h	Device	Bit **Description** 15 FUA - See below 14 Shall be set to one 13 Reserved 12 Shall be set to zero 11:8 Reserved
	Command	7:0 61h

4.5.1.11.2 Forced Unit Access (FUA)

When the FUA bit is set to one the device shall write the data to the media before indicating command completion. When the FUA bit is cleared to zero the device may return command completion before the data is written to the media.

4.5.1.12 Command Acceptance Outputs

See table yy

4.5.1.13 Normal Outputs

See table mm.

4.5.1.14 Error Outputs

This return indicates that the command was aborted due to LBA out of range, a duplicate tag number, an invalid tag number, or an ICRC error, see table zz for more information. Additionally, errors which occur during the processing of this command are reported by returning a transport dependent indicator with additional information available in log address 10h, see table nn for more information.

4.5.2 Changes to DCO Set and Restore

If the proposal adds a feature set or command, there may be be an addition to the Device Configuration Set data structure and the Device Configuration Identify data structure to allow the capability to be removed from the device. Fill in the description in the table and place a description of each bit or word underneath the table. This one table will indicate the changes for both.

Word	Content
8	Serial ATA Command set/feature set supported 15:1 Reserved 0 1 = Reporting for Native Command Queuing is allowed

DCO SET

When bit 0 cleared to zero, drive support for Native Command Queuing shall be disabled and Word 76 bit 8, Word 78 bit 1, Word 78 bit 2, Word 78 bit 4, Word 79 bit 1, Word 79 bit 2, and Word 79 bit 4 of IDENTIFY DEVICE shall all be cleared to zero.

DCO IDENTIFY

Word 8 bit 0 when set to one indicates that the device is allowed to report support for Native Command Queuing.

4.5.3 Changes to IDENTIFY DEVICE data

If the proposal adds a feature set or command, there should be an addition to the IDENTIFY DEVICE data to indicate both supported and enabled descriptions. Fill in the description in the table and place a description of each bit or word underneath the table.

Word	O M	S P	F V	Description
76	O			Serial ATA Capabilities
				15:9 Reserved for Serial ATA
			F	8 1= Native Command Queuing feature set is supported
			F	7:0 Reserved for Serial ATA
78	O			Serial ATA features supported
			F	15:5 Reserved for Serial ATA
				4:0 Reserved for Serial ATA
79	O		F	Same as 76, except says enabled

Key:
O/M – Mandatory/optional requirement.
M – Support of the word is mandatory.
O – Support of the word is optional.
F/V – Fixed/variable content
F – The content of the field is fixed and does not change. The DCO command may change the value of a fixed field. For removable media devices, these values may change when media is removed or changed.
V – The contents of the field is variable and may change depending on the state of the device or the commands executed by the device.
X – The content of the field may be fixed or variable
S/P – Content applies to Serial or Parallel transport
S – Serial Transport
P – Parallel Transport
B – Both Serial and Parallel Transports
N – Belongs to a transport other than Serial or Parallel

Description 1

Description 2

Description 3

Description 4

Description 5

4.6 Changes to Normal and Error Outputs

4.6.1 Changes to Normal Outputs

All commands with normal outputs reference a table in this subclause. Some proposals may not need to define a new normal output. In this case, this subclause should be deleted. If a new normal output is required, please use the sample table included below[

Table 1: yy

Word	Name	Description
00h	Error	Shall be cleared to zero
01h	Count	N/A
02h-04h	LBA	N/A
05h	Device	**Bit** **Description** 15:8 N/A
	Status	7:6 Transport Dependent - See 6.1.9. 5 Device Fault – See 6.1.3 4 N/A 3 Transport Dependent - See 6.1.9. 2:1 N/A 0 Error - See 6.1.2

Table 2: mm

Word	Name	Description
00h		Transport dependent
01h	Error	15:8 Shall be cleared to 0
	Status	7 Shall be cleared to zero 6 Transport Dependent - See 6.1.9. 5 Device Fault – See 6.1.3 4 N/A 3 Shall be cleared to zero. 2:1 N/A 0 Error - See 6.1.2
02h-03h	SActive	31:0 Transport dependent completion indicator
04h-05h		Reserved

4.6.2 Changes to Error Outputs

All commands with error outputs reference a table in this subclause. Some proposals may not need to define a new error output. In this case, this subclause should be deleted. If a new error output is required, please use the sample table included below.

Table 3: zz

Word	Name	Description
00h	Error	**Bit** **Description** 15:8 Reserved 7 Interface CRC - See 6.2.6. [Editors note: add notation instructing editor to fill in cross reference] 6:3 N/A 2 Abort - See 6.2.1. 1:0 N/A
01h	Count	N/A
02h-04h	LBA	N/A
05h	Device	**Bit** **Description** 15:8 N/A
	Status	**Bit** **Description** 15:8 Reserved 7:6 Transport Dependent - See 6.1.9. 5 Device Fault - See 6.1.3 4 N/A 3 Transport Dependent - See 6.1.9. 2:1 N/A 0 Error - See 6.1.2

Table 4: nn

Word	Name	Description
00h		Transport dependent
01h	Error	**Bit** **Description** 7 Interface CRC - See 6.2.6 6 Uncorrectable Error - See 6.2.9 5 Obsolete 4 ID Not Found - See 6.2.4 3 Obsolete 2 Abort - See 6.2.1 1:0 Obsolete
	Status	7 Shall be cleared to zero 6 Transport Dependent - See 6.1.9. 5 Device Fault – See 6.1.3 4 N/A 3 Shall be cleared to zero. 2:1 N/A 0 Error - See 6.1.2
02h-03h	SActive	31:0 Transport dependent completion indicator
04h-05h		Reserved

Editor's Note 6: Fill in tables for Write error bits

4.7 Changes to Log Address Definitions

Log address 10h

Devices supporting the native queued capability shall support READ LOG EXT log page 10h. Page 10h is one page in length and is defined in table xxx

Table 5 — Queued Error Log

Byte	7	6	5	4	3	2	1	0
0	NQ	Reserved		NCQ Tag				
1	Reserved							
2	Status							
3	Error							
4	LBA (7:0)							
5	LBA (15:8)							
6	LBA (23:16)							
7	Device							
8	LBA (31:24)							
9	LBA (39:32)							
10	LBA (47:40)							
11	Reserved							
12	Count (7:0)							
13	Count (15:8)							
14-255	Reserved							
256-510	Vendor Specific							
511	Checksum							

4.7.1 NCQ Tag

If the NQ bit is cleared, the NCQ Tag field contains the NCQ Tag corresponding to the NCQ command that failed.

4.7.2 NQ

If set indicates that the error condition was a result of a non-NCQ command having been issued and that the NCQ Tag field is therefore not valid. If cleared indicates that the NCQ Tag field is valid and that the error condition applies to a NCQ command.

4.7.3 Return Fields

The Status, Error, LBA and Count fields indicate the error that caused the device to stop processing NCQ commands.

> NOTE 2 — The value returned in the Error field of the Queued Error Log may be different than the value returned in the Error field of table nn when the initial error condition is signaled. The Error field in table nn is used for the purpose of signaling a queued command error, while the value in the Error field of the Queued Error Log provides specific information about the error condition.

4.7.4 Checksum

The data structure checksum is the 2's complement of the sum of the first 511 bytes in the data structure. Each byte shall be added with 8-bit unsigned arithmetic and overflow shall be ignored. The sum of all 512 bytes of the data structure shall be zero.

4.8 Changes to Annex B - Command Set Summary

4.8.1 Changes to the Command Matrix Table

TBD

If the proposal adds commands, this subclause should be included with just the TBD. The reason is that the an actual command code is not assigned until the proposal is accepted.

4.8.2 Changes to the Command Codes Table

Fill in the proper command names and other information replacing the information in the table included below. The command code should remain TBD until a command code assignment is made.

Protocol	Command	General Feature Set	Packet Feature Set	Command Code
DMQ	READ FPDMA QUEUED	O	N	60
DMQ	WRITE FPDMA QUEUED	O	N	61

Key:

ND = Non-data command
PI = PIO data-in command
PO = PIO data-out command
DM = DMA command
DMQ = DMA QUEUED command
DR = DEVICE RESET command
DD = EXECUTE DEVICE DIAGNOSTIC command
P = PACKET command
VS = Vendor specific

M = Mandatory
O = Optional
N = Use prohibited
V = Vendor specific implementation
E = Retired
B = Obsolete
R = Reserved
F = If the device does not implement the CFA feature set, this command code is Vendor specific.

4.8.3 Changes to the Historical Command Assignments Table

Fill in the proper command names in the table included below. The ACS editor will fill in the historical information for the command codes that are assigned. The command code should remain TBD until a command code assignment is made.

Opcode	Command Name	ATA1	ATA2	ATA3	ATA4	ATA5	ATA6	ATA7	ATA8
60h	READ FPDMA QUEUED	R	R	R	R	R	R	S	C*
61h	WRITE FPDMA QUEUED	R	R	R	R	R	R	S	C*

Key:

C = a defined command.
E = a retired command.
O = Obsolete.
R = Reserved, undefined in current specifications.
V = Vendor specific commands.

A = Reserved for assignment by the CompactFlash Association
F = If the device does not implement the CFA feature set, this command code is Vendor specific.
M = Reserved for the Media Card Pass Through Command feature set.
S = Reserved for Serial ATA
*Indicates this definition is new to ATA8

4.8.4 Changes to the Historical Set Feature Code Assignments Table

If the proposal added a SET FEATURES Feature Code, then fill in the proper feature names in the table included below. The ACS editor will fill in the historical information for the command codes that are assigned. The command code should remain TBD until a command code assignment is made

Feature Code	Description	ATA1	ATA2	ATA3	ATA4	ATA5	ATA6	ATA7	ATA8
TBD	Your Feature Name	C	C	O	E	F	F	F	F
TBD	Your Feature Name	V	V	C	C	C	C	C	C

Key:

C = a defined command.
E = a retired command.
O = Obsolete.
R = Reserved, undefined in current specifications.
V = Vendor specific commands.
A = Reserved for assignment by the CompactFlash? Association.
F = If the device does not implement the CFA feature set, this command code is Vendor specific.
M = Reserved for the Media Card Pass Through Command feature set.
S = Reserved for Serial ATA.
T = Reserved for Technical Report T13/DT1696 (Time-Limited Commands).