

**Proposed
Draft**

**Serial ATA
International Organization**

**Version 2
March 16, 2020**

**Serial ATA Revision 3.4 Technical Proposal 087
Title: Digital: Fast Fail with Controls**

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.

Permission is granted to the Promoters, Contributors and Adopters of the Serial ATA International Organization to reproduce this document for the purposes of evolving the technical content for internal use only without further permission provided this notice is included. All other rights are reserved and may be covered by one or more Non Disclosure Agreements including the Serial ATA International Organization participant agreements. Any commercial or for-profit replication or republication is prohibited. Copyright © 2000 to 2020 Serial ATA International Organization. All rights reserved.

This Draft Specification is NOT the final version of the Specification and is subject to change without notice. A modified, final version of this Specification (“Final Specification”) when approved by the Promoters will be made available for download at this Web Site:
<http://www.sata-io.org>.

THIS DRAFT SPECIFICATION IS PROVIDED “AS IS” WITH NO WARRANTIES WHATSOEVER, INCLUDING ANY WARRANTY OF MERCHANTABILITY, NON-INFRINGEMENT, FITNESS FOR ANY PARTICULAR PURPOSE OR ANY WARRANTY OTHERWISE ARISING OUT OF ANY PROPOSAL, SPECIFICATION, OR SAMPLE. Except for the right to download for internal review, no license, express or implied, by estoppel or otherwise, to any intellectual property rights is granted or intended hereunder.

THE PROMOTERS DISCLAIM ALL LIABILITY, INCLUDING LIABILITY FOR INFRINGEMENT OF ANY PROPRIETARY RIGHTS, RELATING TO USE OF INFORMATION IN THIS DRAFT SPECIFICATION. THE PROMOTERS DO NOT WARRANT OR REPRESENT THAT SUCH USE WILL NOT INFRINGE SUCH RIGHTS.

THIS DOCUMENT IS AN INTERMEDIATE DRAFT FOR COMMENT ONLY AND IS SUBJECT TO CHANGE WITHOUT NOTICE.

* Other brands and names are the property of their respective owners.

Copyright © 2000 to 2020 Serial ATA International Organization. All rights reserved.

Author Information

Author Name	Company	Email address
Jim Hatfield for D220	Seagate Technology	james.c.hatfield@seagate.com
Ralph Weber for D225	Western Digital	ralph.weber@wdc.com

Workgroup Chair Information

Workgroup (Phy, Digital, etc...)	Chairperson Name	Email address
Digital	James C. Hatfield	james.c.hatfield@seagate.com

Document History

Version	Date	Comments
D220		
0	March 21, 2019	Initial version.
1	May 10, 2019	Renamed the feature, fields to be more aligned with T13 changes: command limits (not 'fast fail', not forever only for read)
2	October 17, 2019	Sync with T13/f18162r9, which T13 ratified on 10/16/2019 <ul style="list-style-type: none"> - See ACS-5 instead of putting behavior details here - Added WRITE FPDMA QUEUED
3	October 17, 2019	Better wording suggested by Ralph Weber
D225		
0	February 21, 2020	Add field usage controls, especially with respect to the interactions between the PRIO field and the COMMAND DURATION LIMITS INDEX field.
1	February 25, 2020	Make changes requested by February 24 th Digital Workgroup call. <ul style="list-style-type: none"> • Include interactions discussion in subclause 13.6.2.4 Priority. • Make non-zero PRIO field an error if the Command Duration Limits feature set is enabled.
2	March 16, 2020	Made the change requested by the March 16 th Digital Workgroup call, specifically... Made the PRIO field ignored if the Command Duration Limits feature set is enabled (as was the case in r0). Ignoring the PRIO field was preferred by all those present on the March 16 th Digital Workgroup call. A major concern was that making a non-zero PRIO field an error in such cases served to create an additional reason for aborting the entire NCQ Queue, which those present felt was carrying a worthy principle too far.

Introduction (Not part of any proposed text)

This proposal for addition to SATA Revision 3.4 defines a new command behavior for the READ FPDMA QUEUED command to standardize the method outlined by the Open Compute Platform (OCP) organization in a paper, Cloud Fast Fail Read Specification, outlining a method to separately track and control the amount of time that a device is actively servicing a read command versus the amount of time that the device is waiting to service that command.

This proposal adds to the Command Duration Limits feature as described in the T13 proposal: f18162 ACS-5 Command Duration Limits feature set

In addition to the changes approved in technical proposal D220, this proposal defines the interactions between the PRIO field and the COMMAND DURATION LIMITS INDEX field based on the supported/enabled state of the Command Duration Limits feature set (see ACS-5).

As discussed during the February T13 Plenary, consideration was given to defining the 11b value of the PRIO field as the indication that the COMMAND DURATION LIMITS INDEX field is not ignored. However, a careful review of this approach encountered several problems, all of which arise from a redundancy between the supported/enabled controls that are already defined in ACS-5. Having two mechanisms to control the same behavior:

- increases the complexity of defining the behaviors because the cases where the two control methods disagree must be specified with suitable error reporting included; and
- both methods must be tested by the device for each command processed, which represents an unacceptable overhead for READ/WRITE FPDMA queued commands.

1 Technical Specification Changes

1.1 <Title of section being changed>

[editor note: Existing text is black.

New text is marked as underlined in blue color or underlined in teal blue color.

Material to be deleted is ~~red with strikethrough markings~~.

<<.. indicates editorial text ..>>

Indicates that reference or code value may change when integrated

13.6.2.4 Priority

Host knowledge of I/O priority may be transmitted to the device as part of the command. There are two priority values for NCQ commands, normal and high. If the host marks an NCQ command as high priority, the host is requesting a better quality of service for that command than commands issued with normal priority.

The classes are forms of soft priority. The device may choose to complete a normal priority command before an outstanding high priority command, although preference should be given to the high priority commands.

EXAMPLE – One example where a normal priority command may be completed before a high priority command is if the normal priority command is a cache hit, whereas the high priority command requires access of the device media.

The priority class is specified in the Priority (PRIO) field for READ FPDMA QUEUED commands, WRITE FPDMA QUEUED commands, RECEIVE FPDMA QUEUED commands, SEND FPDMA QUEUED commands, and DURABLE/ORDERED WRITE NOTIFICATION subcommands. This bit may specify either the normal priority or high priority value. If a command is marked by the host as high priority, the device should attempt to provide better quality of service for the command. It is not required that devices process all high priority requests before satisfying normal priority requests.

[For the READ FPDMA QUEUED command \(see 13.6.4\) and the WRITE FPDMA QUEUED command \(see 13.6.5\), processing of the PRIO field may be affected by the Command Duration Limits feature set \(see ACS-5\).](#)

The device should complete high priority requests in a more timely fashion than normal and isochronous requests. The device should complete isochronous request prior to its associated deadline.

The device should complete isochronous request prior to its associated deadline (see Table 106).

Table 106 – Priority

PRIO field (1:0)	Description
00b	Normal Priority
01b	Isochronous – deadline dependent priority
10b	High priority
11b	Reserved

13.6.4 READ FPDMA QUEUED command

13.6.4.1 READ FPDMA QUEUED command definition

Queued native read commands use this command. The command supports LBA mode only and uses 48 bit addressing only. The format of the command is defined in Figure 337.

Field	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(4:0)					Reserved	RARC	
COUNT(15:8)	PRIO(1:0)		GROUP ID (5:0)					
LBA(7:0)	LBA(7:0)							
LBA(15:8)	LBA(15:8)							
LBA(23:16)	LBA(23:16)							
LBA(31:24)	LBA(31:24)							
LBA(39:32)	LBA(39:32)							
LBA(47:40)	LBA(47:40)							
ICC(7:0)	ICC(7:0)							
AUXILIARY(7:0)	Reserved					Reserved COMMAND DURATION LIMITS INDEX		
AUXILIARY(15:8)	Reserved							
AUXILIARY(23:16)	HYBRID INFORMATION(7:0)							
AUXILIARY(31:24)	Reserved							
DEVICE(7:0)	FUA	1	Res	0	Reserved			
COMMAND(7:0)	60h							

Figure 337 – READ FPDMA QUEUED command definition

Field Definitions

....

PRIO [If the Command Duration Limits feature set \(see ACS-5\) is not supported or not enabled, then the](#) ~~The~~ PRIO field value is assigned by the host based on the priority of the command issued. The device should complete high priority requests in a more timely fashion than normal and isochronous requests. The device should complete isochronous request prior to its associated deadline (see Table 106).
[If the Command Duration Limits feature set is supported and enabled, then the device shall ignore the PRIO field.](#)

....

[COMMAND DURATION LIMITS INDEX](#)

[If the Command Duration Limits feature set is supported and enabled, then usage of the COMMAND DURATION LIMITS INDEX field is defined by ACS-5.](#) ~~usage defined by the Command Duration Limits feature set (see ACS-5)~~
[If the Command Duration Limits feature set \(see ACS-5\) is not supported or not enabled, then the device shall ignore the COMMAND DURATION LIMITS INDEX field.](#)

.....

<<Ed. Note: there are no other changes to READ FPDMA QUEUED>>

13.6.5 WRITE FPDMA QUEUED command

13.6.5.1 WRITE FPDMA QUEUED command definition

Queued native write commands use this command. The command supports LBA mode only and uses 48 bit addressing only. The format of the command is defined in Figure 342.

Field	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(4:0)					Reserved		
COUNT(15:8)	PRIO(1:0)		GROUP ID (5:0)					
LBA(7:0)	LBA(7:0)							
LBA(15:8)	LBA(15:8)							
LBA(23:16)	LBA(23:16)							
LBA(31:24)	LBA(31:24)							
LBA(39:32)	LBA(39:32)							
LBA(47:40)	LBA(47:40)							
ICC(7:0)	ICC(7:0)							
AUXILIARY(7:0)	Reserved					Reserved COMMAND DURATION LIMITS INDEX		
AUXILIARY(15:8)	Reserved							
AUXILIARY(23:16)	HYBRID INFORMATION(7:0)							
AUXILIARY(31:24)	Reserved							
DEVICE(7:0)	FUA	1	Res	0	Reserved			
COMMAND(7:0)	61h							

Figure 342 – WRITE FPDMA QUEUED command definition

Field Definitions

....

PRIO [If the Command Duration Limits feature set \(see ACS-5\) is not supported or not enabled, then the PRIO field value is assigned by the host based on the priority of the command issued. The device should complete high priority requests in a more timely fashion than normal and isochronous requests. The device should complete isochronous request prior to its associated deadline \(see Table 106\).](#)
[If the Command Duration Limits feature set is supported and enabled, then the device shall ignore the PRIO field.](#)

....

[COMMAND DURATION LIMITS INDEX](#)

[If the Command Duration Limits feature set is supported and enabled, then usage of the COMMAND DURATION LIMITS INDEX field is defined by ACS-5. ~~usage defined by the Command Duration Limits feature set \(see ACS-5\)~~](#)
[If the Command Duration Limits feature set \(see ACS-5\) is not supported or not enabled, then the device shall ignore the COMMAND DURATION LIMITS INDEX field.](#)

....

<<Ed. Note: there are no other changes to WRITE FPDMA QUEUED>>