



LEGAL NOTICE:

© Copyright 2007 - 2021 NVM Express, Inc. ALL RIGHTS RESERVED.

This NVM Express NVM Express revision 1.4 technical proposal is proprietary to the NVM Express, Inc. (also referred to as “Company”) and/or its successors and assigns.

NOTICE TO USERS WHO ARE NVM EXPRESS, INC. MEMBERS: Members of NVM Express, Inc. have the right to use and implement this NVM Express revision 1.4 technical proposal subject, however, to the Member’s continued compliance with the Company’s Intellectual Property Policy and Bylaws and the Member’s Participation Agreement.

NOTICE TO NON-MEMBERS OF NVM EXPRESS, INC.: If you are not a Member of NVM Express, Inc. and you have obtained a copy of this document, you only have a right to review this document or make reference to or cite this document. Any such references or citations to this document must acknowledge NVM Express, Inc. copyright ownership of this document. The proper copyright citation or reference is as follows: “© 2007 - 2021 NVM Express, Inc. ALL RIGHTS RESERVED.” When making any such citations or references to this document you are not permitted to revise, alter, modify, make any derivatives of, or otherwise amend the referenced portion of this document in any way without the prior express written permission of NVM Express, Inc. Nothing contained in this document shall be deemed as granting you any kind of license to implement or use this document or the specification described therein, or any of its contents, either expressly or impliedly, or to any intellectual property owned or controlled by NVM Express, Inc., including, without limitation, any trademarks of NVM Express, Inc.

LEGAL DISCLAIMER:

THIS DOCUMENT AND THE INFORMATION CONTAINED HEREIN IS PROVIDED ON AN “AS IS” BASIS. TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, NVM EXPRESS, INC. (ALONG WITH THE CONTRIBUTORS TO THIS DOCUMENT) HEREBY DISCLAIM ALL REPRESENTATIONS, WARRANTIES AND/OR COVENANTS, EITHER EXPRESS OR IMPLIED, STATUTORY OR AT COMMON LAW, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE, VALIDITY, AND/OR NONINFRINGEMENT.

All product names, trademarks, registered trademarks, and/or servicemarks may be claimed as the property of their respective owners.

NVM Express Management Interface Workgroup
c/o VTM, Inc.
3855 SW 153rd Drive
Beaverton, OR 97003 USA
info@nvmexpress.org

NVM Express Technical Proposal for New Feature

Technical Proposal ID	4088 – Rotational Media
Change Date	2021-04-08
Builds on Specification	NVMe Base Specification, Revision 1.4b NVMe-MI TP6022 NVMe-MI 1.1 Maintenance TP4084 Time to Ready Enhancements TP4105a I/O Command Set Independent data structures

Technical Proposal Author(s)

Name	Company
Jim Hatfield David Allen	Seagate
Dave Landsman Christoph Hellwig Yoni Shternhell	Western Digital
Lee Prewitt Scott Lee	Microsoft

- Add support for rotational media (e.g., HDD)

Revision History

Revision Date	Change Description
2020-08-14	Initial version
2020-08-31	Updated based on recent workgroup discussions
2020-09-04	Updated based on today's workgroup discussion <ul style="list-style-type: none">- Suggest a field in Identify Namespace for the Rotational Media bit- Suggest a controller register for a Rotational Media Present bit that is visible prior to controller enable- Move 'number of reallocations' from the EG log to the new log

Revision Date	Change Description
	<ul style="list-style-type: none"> - Remove other changes to the EG log - Suggest a way for the host to test whether an EG contains rotational media - Initial spinup: model support/control after TP4084 Time to Ready <ul style="list-style-type: none"> o Provide a persistent default o Provide capabilities o Provide method to change the default o Mfg default=start spinup at first controller enable o Note: MI can process commands prior to controller enable
2020-09-29	<ol style="list-style-type: none"> 1. Removed the changes to controller registers 2. Removed the backup tutorial material 3. Removed absolute persistence of the new feature, instead relying on the existing mechanism of Saveable feature settings.
2020-10-06	<ol style="list-style-type: none"> 1. Changed new feature to more closely mimic SATA pin P11 staggered spinup (inhibit/permit spinup) 2. Removed option for spinup on first primary controller enable – default to spinup on power on (unless inhibited) 3. Added definition of 'spinup' 4. Removed 'spares' from the Rotational Media Information log page 5. Made rotational speed=FFFFh to be Reserved
2020-10-09	<ol style="list-style-type: none"> 1. Rebase on NVMe 1.4b 2. Specified that the new Identify bit is to be in the NSFEAT field 3. Added a new EGFEAT field to the Endurance Group Information log page, and a requirement to support the log page if there is an association with rotational media. 4. Added Figure 191 5. Rebased the Rotational Media Information log page on EGs, not namespaces 6. Renamed the new feature 'Spinup Control', and based it on EGs not namespaces 7. Added interaction with Sanitize
2020-11-03	<ol style="list-style-type: none"> 1. Added IO and Admin controller tables of logs and features supported 2. Clarified language about spinup and load counts 3. Clarified language about domain 'power on' events to refer to NVM Subsystem Resets instead 4. Added new section 8.new for a feature overview 5. Got rid of requirements specified in more than one place
2020-11-20	<ol style="list-style-type: none"> 1. Added model clause for rotational media 2. Endurance group support is now required 3. Added: allow reading of endurance group information log (with modifications) during sanitize 4. Added a note about the PCIe Slot Power Control feature
2021-01-04	<ol style="list-style-type: none"> 1. Changed copyright date to 2021 2. Replied to comments
2021-01-07	<ol style="list-style-type: none"> 1. Cleaned up approved comment resolutions 2. Added text about spin state and non-operational power states 3. Synched section, table and figure numbers with NVMe 1.4b 4. Put sections in order
2021-01-11	<ol style="list-style-type: none"> 1. Added modification on top of TP6022 NVMe-MI 1.1 Maintenance (new log and feature)

Revision Date	Change Description
2021-01-12	1. Add modification on top of TP4084 Time To Ready Enhancements (new log in a table)
2021-01-28	1. Clarified that 'nominal rotation speed' is in Power State 0 2. Changed filename to indicate Phase 3
2021-02-19	1. Moved the Identify changes to modification of TP4105a I/O command set independent data structure (and changed it from CNS=00h to CNS=08h) 2. Addressed member review comments received to date 3. Dealt with all instances of the phrase 'store data on rotational media' by creating new section 8.new.2, and referencing it as needed.
2021-02-22	1. Resolved comments from Fred Knight, Mike Allison and others 2. Substantive rewordings and rearrangements. Do a document compare with a previous version to see the differences, if you want.
2021-02-25	1. Accepted some editorial comments from Mike Allison 2. Changed the bits in the EGRFEAT field in figure 208 to a bit table
2021-04-01	1. Addressed editorial comments from the Member Review. 2. Now ready for integration
2021-04-06	1. Integrated into the NVMe Base Specification.
2021-04-07	1. Add a "the" to the definition of the RMEDIA bit. Updated the "Identify Namespace data structure" to "I/O Command Set Independent Identify Namespace data structure" in section 8.new.
2021-04-08	1. Removed comments, accepted all changes, and converted references/cross-references to text.

Markup Conventions:

Black:	Unchanged (however, hot links are removed)
Red Strikethrough:	Deleted
Blue:	New
Blue Highlighted:	TBD values, anchors, and links to be inserted in new text.
<Green Bracketed>:	Notes to editor

Summary of Specification Changes

1. Add some definitions
2. Add one bit to the Identify NSFEAT field
3. Define a new Set/Get feature: Spinup Control
4. Added a new EGFEAT field to the Endurance Group log page
5. Define a new log page: Rotational Media Information
6. Added interactions with Sanitize
7. Added support tables for IO and Admin controllers (logs and features)
8. Added feature overview section
9. Added interactions for management endpoints
10. Added interactions with TP4084 Time To Ready Enhancements
11. Added interactions with TP4105a Command Set Independent data structures

Description of Specification Changes to NVMe 1.4b

<add new definitions>

1.6.new3 rotational media

Media that stores data on rotating platters (refer to section [8.new](#)).

1.6.new4 spindown

The process of changing a spindle from an operational power state to a non-operational power state, for an Endurance Group that stores data on rotational media (refer to section [8.new](#)).

1.6.new5 spinup

The process of changing a spindle from a non-operational power state to an operational power state, for an Endurance Group that stores data on rotational media (refer to section [8.new](#)).

<define a new Feature for Get Features to the list of non-command set specific features >

<note 1: NVMe 1.MERGED has 2 new columns (Log Identifier and Scope) acknowledged here>

<note 2: this log is NOT I/O command set specific>

5.13 Get Features command

...

Description	Section Defining Format of Attributes Returned
...	
Spinup Control	5.21.1.new
...	

5.14 Get Log Page command

...

<indicate that the new Rotational Media Information log page is related to Endurance Groups>

Figure 191: Get Log Page – Command Dword 11

Bits	Description
31:16	Log Specific Identifier: This field specifies an identifier that is required for a particular log page. The log pages that require a log specific identifier are indicated in the table below.
15:00	Number of Dwords (NUMDU): This field specifies the upper 16 bits of the number of dwords to return.

...

5.14.1 Log Specific Information

<define a new log page>

...

Figure 195: Get Log Page – Log Page Identifiers

Log Identifier	Scope	Log Page Name	Reference Section
...			
16h	Endurance Group	Rotational Media Information	5.14.1.new
...			

5.14.1.9 Endurance Group Information (Log Identifier 09h)

<add a new bit to the Endurance Group Information log page>

...

Figure 208: Get Log Page – Endurance Group Information Log (Log Identifier 09h)

Bytes	Description						
...	...						
01	<u>Endurance Group Features (EGFEAT): This field defines features of the Endurance Group.</u>						
	<table><tr><th><u>Bit</u></th><th><u>Description</u></th></tr><tr><td><u>7:1</u></td><td><u>Reserved</u></td></tr><tr><td><u>0</u></td><td><u>Endurance Group Rotational Media (EGRMEDIA): if set to '1' indicates that the Endurance Group stores data on rotational media (refer to section 8.new). If cleared to '0', then the Endurance Group does not store data on rotational media.</u></td></tr></table>	<u>Bit</u>	<u>Description</u>	<u>7:1</u>	<u>Reserved</u>	<u>0</u>	<u>Endurance Group Rotational Media (EGRMEDIA): if set to '1' indicates that the Endurance Group stores data on rotational media (refer to section 8.new). If cleared to '0', then the Endurance Group does not store data on rotational media.</u>
	<u>Bit</u>	<u>Description</u>					
<u>7:1</u>	<u>Reserved</u>						
<u>0</u>	<u>Endurance Group Rotational Media (EGRMEDIA): if set to '1' indicates that the Endurance Group stores data on rotational media (refer to section 8.new). If cleared to '0', then the Endurance Group does not store data on rotational media.</u>						
02:01	Reserved						
...							

<define a new log page>

5.14.1.new Rotational Media Information (Log Identifier 16h)

This log page provides rotational media information (refer to section 8.new) for Endurance Groups that store data on rotational media. The information provided is retained across power cycles and resets.

The Endurance Group Identifier is specified in the Log Specific Identifier field in Command Dword 11 of the Get Log Page command.

If the NVM subsystem does not contain any Endurance Groups that store data on rotational media, then the Rotational Media Information Log should not be supported.

Figure newlog: Get Log Page – Rotational Media Information Log

Bytes	Description
1:0	<u>Endurance Group Identifier: The Endurance Group Identifier specified by the Get Log Page command.</u>
3:2	<u>Number of Actuators: Contains the number of actuators in this Endurance Group.</u>

Figure newlog: Get Log Page – Rotational Media Information Log

Bytes	Description										
5:4	Nominal Rotational Speed (NRS):										
	<table><tr><th>Value</th><th>Description</th></tr><tr><td>0000h</td><td>Not reported</td></tr><tr><td>0001h</td><td>This value is prohibited to maintain backward compatibility with other standards. This value shall not be used.</td></tr><tr><td>FFFFh</td><td>Reserved</td></tr><tr><td>All others</td><td>Nominal rotational speed in revolutions per minute while the current Power State is 0 (refer to section 5.21.1.2).</td></tr></table>	Value	Description	0000h	Not reported	0001h	This value is prohibited to maintain backward compatibility with other standards. This value shall not be used.	FFFFh	Reserved	All others	Nominal rotational speed in revolutions per minute while the current Power State is 0 (refer to section 5.21.1.2).
	Value	Description									
	0000h	Not reported									
	0001h	This value is prohibited to maintain backward compatibility with other standards. This value shall not be used.									
	FFFFh	Reserved									
All others	Nominal rotational speed in revolutions per minute while the current Power State is 0 (refer to section 5.21.1.2).										
7:6	Reserved										
11:8	Spinup Count: Contains the total number of successful spinup events for this Endurance Group over the lifetime of the Endurance Group. If the Spinup Count is less than FFFFFFFFh, then the controller shall increment this count by one for each successful spinup event. A successful spinup event occurs when the controller power state transitions from a non-operational power state to an operational power state.										
15:12	Failed Spinup Count: Contains the total number of failed spinup events for this Endurance Group over the lifetime of the Endurance Group. If the Failed Spinup Count is less than FFFFFFFFh, then the controller shall increment this count by one for each failed spinup event. A failed spinup event occurs when the controller fails an attempt to transition from a non-operational power state to an operational power state.										
19:16	Load Count: Contains the total number successful actuator load events for this Endurance Group over the lifetime of the Endurance Group. If the Load Count is less than FFFFFFFFh, then the controller shall increment this count by one for each successful actuator load event. A successful actuator load event occurs if an actuator transitions from a non-operational state to an operational state.										
23:20	Failed Load Count: Contains the number of failed actuator load events for this Endurance Group over the lifetime of the Endurance Group. If the Failed Load Count is less than FFFFFFFFh, then the controller shall increment this count by one for each failed actuator load event. A failed actuator load event occurs if an actuator fails an attempt to transition from a non-operational state to an operational state.										
511:24	Reserved										

5.21.1 Feature Specific Information

...

<define a new Feature Identifier for Set Features>

Figure 275: Set Features – Feature Identifiers

Feature Identifier	Current Setting Persists Across Power Cycle and Reset ²	Uses Memory Buffer for Attributes	Feature Name
...			
1Ah	Yes	No	Spinup Control
...			

<define new Spinup Control feature>

5.21.1.new Spinup Control (Feature Identifier 1Ah)

This Feature allows the host to configure the method for initial spinup for Endurance Groups that stores data on rotational media (refer to section [8.new](#)). The NVM Subsystem is the scope for this feature.

If the NVM subsystem does not contain any Endurance Groups that store data on rotational media, then the controller shall abort the Set Features command and the Get Features command for this feature with status code of Invalid Field In Command.

The method is specified in Command Dword 11 (refer to [figure newfeatureA](#)).

Figure newfeatureA: Spinup Control – Command Dword 11

Bits	Description
31:01	Reserved
0	If set to '1', then the Spinup Control feature is enabled. If cleared to '0', then the Spinup Control feature is disabled. The setting is persistent.

If a Get Features command is submitted for this Feature, the attributes described in (refer to [figure newfeatureB](#)) are returned in Dword 0 of the completion queue entry for that command.

Figure newfeatureB: Completion Queue Entry Dword 0

Bits	Description
31:01	Reserved
0	If set to '1', then the Spinup Control feature is enabled. If cleared to '0', then the Spinup Control feature is disabled.

7.1.1 I/O Controller

...

<modify Figure 424 IO controller – Log Page Support>

Figure 424: I/O Controller – Log Page Support

Log Page Name	Command Support Requirements ¹
...	
Rotational Media Information	O
...	

...

<modify Figure 426 IO controller – Feature Support>

Figure 426: I/O Controller – Feature Support

Feature Name	Feature Support Requirements ¹	Logged in Persistent Event Log ¹
...		
Rotational Media	O	O
...		

...

7.1.2 Administrative Controller

...

<modify Figure 432 Administrative - Controller Log Page Support>

Figure 432: Administrative [Controller](#) – ~~Controller~~ Log Page Support

Log Page Name	Command Support Requirements ¹
...	
Rotational Media Information	P
...	

...

<modify Figure 434 Administrative controller – Feature Support>

Figure 434: Administrative Controller – Feature Support

Feature Name	Feature Support Requirements ¹	Logged in Persistent Event Log ¹
...		
Rotational Media	P	P
...		

...

8.15.1 Sanitize Command Restrictions

...

<indicate that the new log is readable (with restrictions) during sanitize >

Figure Figure 486: Sanitize Operations – Admin Commands Allowed

Admin Command	Additional Restrictions	
...		
Get Log Page	The log pages allowed are listed below.	
	Log Pages	Additional Restrictions
	Error Information	Return zeroes in the LBA field.
	SMART / Health Information	
	Changed Namespace List	
	Reservation Notification	
	Sanitize Status	
	Asymmetric Namespace Access	
	Endurance Group Information	Return zeroes for all fields except: Critical Warning field and EGFEAT field.
	Rotational Media Information	Return zeroes in the Spinup Count, Failed Spinup Count, Load Count, and Failed Load Count fields.
...		

<add new overview material in clause 8>

8.new Rotational Media (Optional)

Rotational media has different operational, endurance, and performance characteristics than non-rotational media (e.g., NAND). Rotational media utilizes electromechanical methods for accessing data.

Rotational media contains one or more spinning platters containing the media, and one or more actuators that provide physical access to the data on that media (e.g., a hard disk drive or a CD-ROM).

A controller that supports namespaces that store user data on rotational media shall:

- set the Rotational Media bit to '1' in the NSFEAT field of the I/O Command Set Independent Identify Namespace data structure (refer to **Figure TBDa**) for any namespace that stores data on rotational media;
- support the Rotational Media Information log page (refer to section **5.14.1.new**);
- support the Spinup Control feature (refer to section **5.21.1.new**);
- support Endurance Groups (refer to section **8.17**); and
- set the EG Rotational Media bit to '1' in the EGFEAT field in the Endurance Group Information log page for each Endurance Group that stores data on rotational media.

If a namespace that stores data on rotational media is attached to a controller, and the spindle used by that namespace is not spinning, then that controller shall be in a non-operational power state (i.e., NOPS is set to '1', refer to **Figure 252**). <editor note: "Identify – Power State Descriptor Data Structure" >

If:

- a) a domain contains an Endurance Group that stores data on rotational media;
- b) that domain processes an NVM Subsystem Reset; and
- c) the Spinup Control feature (refer to section 5.21.1.new) is:
 - a. disabled, then initial spinup for all such Endurance Groups in that domain shall be initiated; and
 - b. enabled, then initial spinup for all such Endurance Groups in that domain shall be inhibited during processing of the NVM Subsystem Reset until the controller processes a Set Features (Power Management) command that specifies an operational power state.

If the PCIe transport is used for a controller, then the PCIe Slot Power Control feature may affect the power states supported.

Description of Specification Changes to NVMe-MI

On top of TPAR 6022 NVMe-MI 1.1 Maintenance

<add a log to table TBD1>

Figure TBD1: Management Endpoint – Log Page Support

<u>Log Page Name</u>	<u>SMBus/I2C Log Page Support Requirements¹</u>		<u>PCIe VDM Log Page Support Requirements¹</u>	
	<u>NVMe Storage Device</u>	<u>NVMe Enclosure</u>	<u>NVMe Storage Device</u>	<u>NVMe Enclosure</u>

<u>Rotational Media Information</u>	<u>O</u>	<u>O</u>	<u>O</u>	<u>O</u>

<add a feature to table TBD3>

Figure TBD3: Management Endpoint – Feature Support

<u>Feature Name</u>	<u>SMBus/I2C Log Page Support Requirements¹</u>		<u>PCIe VDM Log Page Support Requirements¹</u>	
	<u>NVMe Storage Device</u>	<u>NVMe Enclosure</u>	<u>NVMe Storage Device</u>	<u>NVMe Enclosure</u>

<u>Spinup Control</u>	<u>O</u>	<u>O</u>	<u>O</u>	<u>O</u>

Modifications to TP4084 Time to Ready Enhancements >

<add one log to table TBD11>

7.6.1.1 Initialization: Controller Ready Modes During Initialization

...

**Figure TBD11: Admin Commands Permitted to Return a Status Code of
Admin Command Media Not Ready**

Admin Command	Additional Restrictions
...	
Get Log Page	<p>Get Log Page is only permitted to return a status code of Admin Command Media Not Ready for the following log pages:</p> <ul style="list-style-type: none">• Device Self-test• Firmware Slot Information• Telemetry Controller-Initiated• Telemetry Host-Initiated• Predictable Latency Per NVM Set• Predictable Latency Event Aggregate• Persistent Event Log• LBA Status Information• Endurance Group Event Aggregate• Media Unit Status• Supported Capacity Configuration List• Boot Partition• Reservation Notification• Rotational Media Information• Vendor Specific
...	

Modifications to TP4105a Command set independent data structure

<modify Identify CNS=08h to add one new bit to the NSFEAT field>

Figure **TBDa**: Identify – I/O Command Set Independent Identify Namespace Data Structure

Bytes	O/M ¹	Description
00	M	Common Namespace Features (NSFEAT): This field defines features of the namespace. Bits 7:4 are reserved. <u>Bit 4 (RMEDIA) if set to '1' indicates that the namespace stores data on rotational media (refer to section 8.new). If cleared to '0', indicates that the namespace does not store data on rotational media.</u> Bit 3 (UIDREUSE) if set to '1' indicates that the value in the NGUID field for this namespace, if non-zero, is never reused by the controller and that the value in the EUI64 field for this namespace, if non-zero, is never reused by the controller. If cleared to '0', then the NGUID value may be reused and the EUI64 value may be reused by the controller for a new namespace created after this namespace is deleted. This bit shall be cleared to '0' if both NGUID and EUI64 fields are cleared to 0h. Refer to section 7.11. Bit 2:0 are reserved
...		