



LEGAL NOTICE:

© **Copyright 2007 to 2021 NVM Express™, Inc. ALL RIGHTS RESERVED.**

This NVM Express revision 2.0 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 2.0 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 to 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.

The NVM Express® design mark is a registered trademark of NVM Express, Inc.

NVM Express 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	TP 4099 MDTS Enhancement
Change Date	2021-9-21
Builds on Specification	NVM Express Base Specification 2.0 NVM Express NVM Command Set Specification 1.0
References Specification	

Technical Proposal Author(s)

Name	Company
Kapil Karkra, Mike Allison, Jonathan Hughes, Jackson Ellis	Intel

This proposal intends to:

- Define the new MEM (MDTS and non-MDTS Size Limits Exclude Metadata) bit in the Controller Attributes field to indicate that MDTS, VSL, WZSL, and WUSL values do not include interleaved metadata.
- Clarify the MDTS definition to state that it only includes interleaved metadata if metadata is interleaved and MEM bit is cleared to 0.

Revision History

Revision Date	Change Description
2021-03-10	Initial rough version
2021-08-01	Integrated into the NVMe Base Specification, revision 2.0, and the NVMe NVM Command Set Specification, revision 1.0.
2021-09-21	Changed "Specific" to "specific". Removed "NVMe" before the NVM Command Set Specification. Removed all references and cross references. Aligned to NVMe 2.0 specification after references were changed. Removed large chunks of tables not modified.

Description for NVM Express Base Specification 2.0 Changes Document

Feature Enhancement:

- Added MEM bit in CRATT field of the Identify Controller Data Structure
- Changed the definition to MDTS in a backward compatible manner
- **New requirement / incompatible change:**
 - An old host that issued a command within the increased size of the MDTS, and expected the controller to fail the command, will now succeed when MEM bit is set to 1.
- References:
 - Base Spec: Sections 5.17.2
 - NVM Command Set Spec: Section 4.1.5.4

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

Modify portions of NVM Express Base Specification 2.0 as shown below:

Modify section 5.17.2 as shown below:

5.17.2 Identify Data Structures

5.17.2.1 Identify Controller data structure (CNS 01h)

...

Figure 275: Identify – Identify Controller Data Structure, I/O Command Set Independent

Bytes	I/O ¹	Admin ¹	Disc ¹	Description
Controller Capabilities and Features				
...				
77	M	M	M	<p>Maximum Data Transfer Size (MDTS): This field indicates the maximum data transfer size for a command that transfers data between host-accessible memory (refer to section 1.5.26) and the controller. The host should not submit a command that exceeds this maximum data transfer size. If a command is submitted that exceeds this transfer size, then the command is aborted with a status of Invalid Field in Command. The value is in units of the minimum memory page size (CAP.MPSMIN) and is reported as a power of two (2^n). A value of 0h indicates that there is no maximum data transfer size.</p> <p>If the MEM bit is cleared to '0' in the CTRATT field, then this field includes the length of metadata, if metadata is interleaved with the user data.</p> <p>If the MEM bit is set to '1', then this field excludes the length of metadata.</p> <p>This field does not apply to commands that do not transfer data between host-accessible memory and the controller (e.g., the Verify command, the Write Uncorrectable command, and the Write Zeroes command); refer to the ONCS field for restrictions on these commands and other commands that transfer data.</p> <p>If SGL Bit Bucket descriptors are supported, their lengths shall be included in determining if a command exceeds the Maximum Data Transfer Size for destination data buffers. Their length in a source data buffer is not included for a Maximum Data Transfer Size calculation.</p>

Bytes	I/O ¹	Admin ¹	Disc ¹	Description								
99:96	M	M	R	Controller Attributes (CTRATT): This field indicates attributes of the controller.								
				<table><tr><th>Bits</th><th>Description</th></tr><tr><td>31:46 17</td><td>Reserved</td></tr><tr><td>16</td><td>MDTS and Size Limits Exclude Metadata (MEM): If set to '1', then:<ul style="list-style-type: none">The controller reported MDTS values do not include interleaved metadata.The controller reported VSL, WZSL, and WUSL values in the I/O Command Set specific Identify Controller data structure (refer to the NVM Command Set Specification) do not include interleaved metadata.If cleared to '0', then:<ul style="list-style-type: none">The controller reported MDTS values include interleaved metadata.The controller reported VSL, WZSL, and WUSL values in the I/O Command Set specific Identify Controller data structure include interleaved metadata.</td></tr><tr><td>15</td><td>Extended LBA Formats Supported (ELBAS): If set to '1' indicates that the controller supports the I/O command set specific extended protection information formats (refer to the Protection Information Formats section of the applicable I/O command set specification). If cleared to '0' indicates that the controller does not support the I/O command set specific extended protection information formats (refer to the Protection Information Formats section of the NVM Command Set Specification). Refer to the LBA Format Extension Enable (LBAFEE) field in the Host Behavior Support feature (refer to section 5.27.1.18) for details for host software to enable the controller to operate on namespaces using the protection information formats. NOTE: This bit field applies to all I/O Command Sets. The original name has been retained for historical continuity.</td></tr></table>	Bits	Description	31: 46 17	Reserved	16	MDTS and Size Limits Exclude Metadata (MEM): If set to '1', then: <ul style="list-style-type: none">The controller reported MDTS values do not include interleaved metadata.The controller reported VSL, WZSL, and WUSL values in the I/O Command Set specific Identify Controller data structure (refer to the NVM Command Set Specification) do not include interleaved metadata. If cleared to '0', then: <ul style="list-style-type: none">The controller reported MDTS values include interleaved metadata.The controller reported VSL, WZSL, and WUSL values in the I/O Command Set specific Identify Controller data structure include interleaved metadata.	15	Extended LBA Formats Supported (ELBAS): If set to '1' indicates that the controller supports the I/O command set specific extended protection information formats (refer to the Protection Information Formats section of the applicable I/O command set specification). If cleared to '0' indicates that the controller does not support the I/O command set specific extended protection information formats (refer to the Protection Information Formats section of the NVM Command Set Specification). Refer to the LBA Format Extension Enable (LBAFEE) field in the Host Behavior Support feature (refer to section 5.27.1.18) for details for host software to enable the controller to operate on namespaces using the protection information formats. NOTE: This bit field applies to all I/O Command Sets. The original name has been retained for historical continuity.
				Bits	Description							
				31: 46 17	Reserved							
16	MDTS and Size Limits Exclude Metadata (MEM): If set to '1', then: <ul style="list-style-type: none">The controller reported MDTS values do not include interleaved metadata.The controller reported VSL, WZSL, and WUSL values in the I/O Command Set specific Identify Controller data structure (refer to the NVM Command Set Specification) do not include interleaved metadata. If cleared to '0', then: <ul style="list-style-type: none">The controller reported MDTS values include interleaved metadata.The controller reported VSL, WZSL, and WUSL values in the I/O Command Set specific Identify Controller data structure include interleaved metadata.											
15	Extended LBA Formats Supported (ELBAS): If set to '1' indicates that the controller supports the I/O command set specific extended protection information formats (refer to the Protection Information Formats section of the applicable I/O command set specification). If cleared to '0' indicates that the controller does not support the I/O command set specific extended protection information formats (refer to the Protection Information Formats section of the NVM Command Set Specification). Refer to the LBA Format Extension Enable (LBAFEE) field in the Host Behavior Support feature (refer to section 5.27.1.18) for details for host software to enable the controller to operate on namespaces using the protection information formats. NOTE: This bit field applies to all I/O Command Sets. The original name has been retained for historical continuity.											
...												

Modify portions of NVM Command Set Specification 1.0 as shown below:

Modify section 4.1.5.4 as shown below:

4.1.5.4 I/O Command Set Specific Identify Controller data structure (CNS 06h)

Figure 102 defines the I/O Command Set Specific Identify Controller data structure for the NVM Command Set.

Figure 102: Identify – Identify Controller Data Structure, NVM Command Set (CSI 00h)

Bytes	O/M ¹	Description
00	O	<p>Verify Size Limit (VSL): If bit 7 in the Optional NVM Command Support (ONCS) field is set to '1' then:</p> <ul style="list-style-type: none"> a) a non-zero value in this field indicates the recommended maximum data size for a Verify command (refer to section 6.14); and b) a value of 0h in this field indicates that no recommended maximum data size for a Verify command is reported. <p>If bit 7 in the ONCS field is cleared to '0' then:</p> <ul style="list-style-type: none"> a) a non-zero value in this field indicates that the controller supports the Verify command with the maximum data size limit indicated by this field (refer to section 6.14); and b) a value of 0h in this field indicates that the controller does not support the Verify command. <p>The non-zero value is in units of the minimum memory page size (CAP.MPSMIN) and is reported as a power of two (2^n).</p> <p>If the MEM bit is cleared to '0' in the CTRATT field in the Identify Controller data structure, then this field includes the length of metadata, if metadata is interleaved with the logical block data.</p> <p>If the MEM bit is set to '1', then this field excludes the length of metadata.</p>
01	O	<p>Write Zeroes Size Limit (WZSL): If bit 3 in the Optional NVM Command Support (ONCS) field is set to '1' then:</p> <ul style="list-style-type: none"> a) a non-zero value in this field indicates the recommended maximum data size for a Write Zeroes command (refer to section 6.17); and b) a value of 0h in this field indicates that no recommended maximum data size for a Write Zeroes command is reported. <p>If bit 3 in the ONCS field is cleared to '0' then:</p> <ul style="list-style-type: none"> a) a non-zero value in this field indicates that the controller supports the Write Zeroes command with the maximum data size limit indicated by this field (refer to section 6.17); and b) a value of 0h in this field indicates that the controller does not support the Write Zeroes command. <p>The non-zero value is in units of the minimum memory page size (CAP.MPSMIN) and is reported as a power of two (2^n).</p> <p>If the MEM bit is cleared to '0' in the CTRATT field in the Identify Controller data structure, then this field includes the length of metadata, if metadata is interleaved with the logical block data.</p> <p>If the MEM bit is set to '1', then this field excludes the length of metadata.</p>

02	O	<p>Write Uncorrectable Size Limit (WUSL): If bit 1 in the Optional NVM Command Support (ONCS) field is set to '1' then:</p> <ul style="list-style-type: none"> a) a non-zero value in this field indicates the recommended maximum data size for a Write Uncorrectable command (refer to section 6.16); and b) a value of 0h in this field indicates that no recommended maximum data size for a Write Uncorrectable command is reported. <p>If bit 1 in the ONCS field is cleared to '0' then:</p> <ul style="list-style-type: none"> a) a non-zero value in this field indicates that the controller supports the Write Uncorrectable command with the maximum data size limit indicated by this field (refer to section 6.16); and b) a value of 0h in this field indicates that the controller does not support the Write Uncorrectable command. <p>The non-zero value is in units of the minimum memory page size (CAP.MPSMIN) and is reported as a power of two (2^n).</p> <p>If the MEM bit is cleared to '0' in the CTRATT field in the Identify Controller data structure, then this field includes the length of metadata, if metadata is interleaved with the logical block data.</p> <p>If the MEM bit is set to '1', then this field excludes the length of metadata.</p>
03	O	<p>Dataset Management Ranges Limit (DMRL): If bit 2 in the Optional NVM Command Support (ONCS) field is set to '1' then:</p> <ul style="list-style-type: none"> a) a non-zero value in this field indicates the recommended maximum number of logical block ranges for a Dataset Management command (refer to section 6.7.1); and b) a value of 0h in this field indicates that no recommended maximum number of logical block ranges for a Dataset Management command is reported. <p>If bit 2 in the ONCS field is cleared to '0', then:</p> <ul style="list-style-type: none"> a) a non-zero value in this field indicates that the controller supports the Dataset Management command with the maximum number of logical block ranges limit indicated by this field (refer to section 6.7.1); and <p>a value of 0h in this field indicates that the controller does not support the Dataset Management command.</p>
07:04	O	<p>Dataset Management Range Size Limit (DMRSL): If bit 2 in the Optional NVM Command Support (ONCS) field is set to '1' then:</p> <ul style="list-style-type: none"> a) a non-zero value in this field indicates the recommended maximum number of logical blocks in a single range for a Dataset Management command (refer to section 6.7.1); and b) a value of 0h in this field indicates that no recommended maximum number of logical blocks in a single range for a Dataset Management command is reported. <p>If bit 2 in the ONCS field is cleared to '0', then:</p> <ul style="list-style-type: none"> a) a non-zero value in this field indicates that the controller supports the Dataset Management command with the maximum number of logical blocks in a single range limit indicated by this field (refer to section 6.7.1); and <p>a value of 0h in this field indicates that the controller does not support the Dataset Management command.</p>

15:08	O	<p>Dataset Management Size Limit (DMSL): If bit 2 in the Optional NVM Command Support (ONCS) field is set to '1' then:</p> <ul style="list-style-type: none"> a) a non-zero value in this field indicates the recommended maximum total number of logical blocks for a Dataset Management command (refer to section 6.7.1). b) a value of 0h in this field indicates that no recommended maximum total number of logical blocks for a Dataset Management command is reported. <p>If bit 2 in the ONCS field is cleared to '0', then:</p> <ul style="list-style-type: none"> a) a non-zero value in this field indicates that the controller supports the Dataset Management command with the maximum total number of logical blocks limit indicated by this field (refer to section 6.7.1); and <p>a value of 0h in this field indicates that the controller does not support the Dataset Management command.</p>
4095:16	M	Reserved
<p>NOTES:</p> <p>1. O/M definition: O = Optional, M = Mandatory.</p>		