LEGAL NOTICE:

© Copyright 2007 - 2016 NVM Express, Inc. ALL RIGHTS RESERVED.
This erratum to the NVM Express revision 1.2 specification 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 erratum to the NVM Express revision 1.2 specification 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 - 2016 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 Workgroup
c/o Virtual, Inc.
401 Edgewater Place, Suite 600
Wakefield, MA 01880
info@nvmexpress.org

Technical input submitted to the NVM Express™ Workgroup is subject to the terms of the NVM Express™ Participant's agreement. Copyright © 2014-16 NVMe™ Corporation.

NVM Express™ Technical Errata

Errata ID	**003**
Revision Date	**12/16/2016**
Affected Spec Ver.	**NVM Express™ 1.2.1**
Corrected Spec Ver.	

Errata Author(s)

Name	Company
Dave Landsman, Christoph Hellwig, Frank Chu	WD
Judy Brock	Samsung
Keith Busch	Intel
Fred Knight	NetApp

Errata Overview

The transfer direction bits use in vendor specific commands is clarified.

The initialization sequence as it relates to interrupts is clarified.

Clarified that Write Zeroes may deallocate the data.

Clarified Invalid Field in Command definition.

Clarifications on MDTS.

Clarification on Number of Dwords valid bits if Extended data for Get Log Page is not supported.

Clarified Controller detection of ACL limit exceeded is implementation specific.

Removed references to undefined term “accessible by”, replace with attached where applicable.

Clarified Private Namespace and Shared Namespace definitions.

Technical input submitted to the NVM Express™ Workgroup is subject to the terms of the NVM Express™ Participant’s agreement. Copyright © 2014-16 NVMe™ Corporation.

Revision History

Revision Date	Change Description
10/11/2016	Initial draft
11/14/2016	Clarified transfer direction bits for vendor specific commands. Added WD proposal for clarifying instantaneous power language. Clarified interrupts during initialization. Added Write Zeroes clarification for deallocating the data.
11/15/2016	command arbitration text clarification
11/16/2016	Updates based on 11/15 optional errata discussion.
11/29/16	Invalid Field in Command definition clarification
12/07/16	MDTS clarifications. Clarification on Number of Dwords valid bits if Extended data for Get Log Page is not supported. ACL clarifications. Remove "accessible" terminology, replace with "attached" where applicable. Clarification on scope of Delete all Namespaces. Clarified Private Namespace and Shared Namespace definitions. Remove Section 7.3.2 Controller Reset changes. Remove Section 9.5 CFS changes.
12/9/16	Included agreed upon change for "instantaneous" power state from WD. Wordsmithing of many fields and re-workings, while keeping same intent.
12/15/16	Fixed a section number for the Command Arbitration editorial change. Added "command" after both Write Uncorrectable and Write Zeroes for the MDTS clarification.
12/15/16 rev2	Added changes to remove the word "dataset" from Dword 13 of Read and Write commands.
12/16/16	Removed two namespace accessible changes that will be addressed in a TP. Reverted name of DSM field in Command Dword 13 for Read and Write.

Description of Specification Changes

Modify a portion of section 8.4 as shown below:

Associated with each power state is a Power State Descriptor in the Identify Controller data structure (refer to Figure 91). The descriptors for all implemented power states may be viewed as forming a table as shown in Figure 220 for a controller with seven implemented power states. Note that Figure 220 is illustrative and does not include all fields in the power state descriptor. The Maximum Power (MP) field indicates the ~~instantaneous~~ maximum power that may be consumed in that state. Refer to the appropriate form factor specification for power measurement methodologies for that form factor. The controller may employ autonomous power management techniques to reduce power consumption below this level, but under no circumstances is power allowed to exceed this level.

Modify Figure 41 (Opcodes for Admin Commands) as shown below:

Technical input submitted to the NVM Express™ Workgroup is subject to the terms of the NVM Express™ Participant's agreement. Copyright © 2014-16 NVMe™ Corporation.

	I/O Command Set Specific					
1b	na	~~na~~ NOTE 4	80h – BFh	O		I/O Command Set specific
	Vendor Specific					
1b	na	~~na~~ NOTE 4	C0h – FFh	O		Vendor specific

NOTES:
1. O/M definition: O = Optional, M = Mandatory.
2. Opcodes not listed are reserved.
3. A subset of commands uses the Namespace Identifier field (CDW1.NSID). When not used, the field shall be cleared to 0h.
4. Indicates the data transfer direction of the command. All commands, including vendor specific commands, should follow this convention. 00b = no data transfer; 01b = host to controller; 10b = controller to host; 11b = bidirectional
5. For NVMe over PCIe implementations, the Keep Alive command is optional. For NVMe over Fabrics implementations, the associated NVMe Transport binding defines whether the Keep Alive command is optional or mandatory.

Modify Figure 146 (Opcodes for NVM Commands) as shown below:

	Vendor Specific				
1b	na	~~na~~ NOTE 5	80h – FFh	O	Vendor specific

NOTES:
1. O/M definition: O = Optional, M = Mandatory.
2. Opcodes not listed are reserved.
3. All NVM commands use the Namespace Identifier field (CDW1.NSID).
4. Mandatory if reservations are supported as indicated in the Identify Controller data structure.
5. Indicates the data transfer direction of the command. All commands, including vendor specific commands, should follow this convention. 00b = no data transfer; 01b = host to controller; 10b = controller to host; 11b = bidirectional

Modify a portion of section 7.6.1 (Initialization) as shown below:

The host should perform the following actions in sequence to initialize the controller to begin executing commands:

1. Set the PCI and PCI Express registers described in section 2 appropriately based on the system configuration. This includes configuration of power management features. A single interrupt (e.g. pin-based, single-MSI, or single MSI-X) should be used until the number of I/O Queues is determined. ~~Pin-based or single-message MSI interrupts should be used until the number of I/O Queues is determined.~~
2. The host waits for the controller to indicate that any previous reset is complete by waiting for CSTS.RDY to become '0.'

Modify a portion of section 6.16 (Write Zeroes) as shown below:

The Write Zeroes command is used to set a range of logical blocks to zero. After successful completion of this command, the value returned by subsequent reads of logical blocks in this range shall be zeroes until a write occurs to this LBA range. The metadata for this command shall be all zeroes and the protection information is updated based on CDW12.PRINFO. It is a valid implementation to deallocate the associated data if the logical blocks and metadata returns zeroes and the protection information is updated based on CDW12.PRINFO.

The fields used are Command Dword 10, Command Dword 11, Command Dword 12, Command Dword 14, and Command Dword 15 fields.

Modify a portion of section 4.11 (Command Arbitration) as shown below:

Technical input submitted to the NVM Express™ Workgroup is subject to the terms of the NVM Express™ Participant's agreement. Copyright © 2014-16 NVMe™ Corporation.

4.11 Command Arbitration

For NVMe over PCIe, a~~A~~ command is submitted when a Submission Queue Tail Doorbell write moves the Submission Queue Tail Pointer past the corresponding Submission Queue entry. For NVMe over Fabrics, refer to section 1.4.14 in the NVMe over Fabrics 1.0 specification for the definition of command submission. The controller transfers submitted commands to the controller's local memory for subsequent processing using a vendor specific algorithm.

Modify a portion of Figure 31 (Generic Command Status Values) as shown below:

Value	Description
00h	**Successful Completion:** The command completed successfully.
01h	**Invalid Command Opcode:** The associated command opcode field is not valid.
02h	**Invalid Field in Command:** ~~In the command parameters, a~~ A reserved coded value or an unsupported value in a defined field (other than the opcode field). The field may be in the command parameters as part of the Submission Queue Entry or in data structures pointed to by the command parameters.

Modify a portion of Figure 90 (Identify – Identify Controller Data Structure) as shown below:

77	M	**Maximum Data Transfer Size (MDTS):** This field indicates the maximum data transfer size between the host and the controller. The host should not submit a command that exceeds this transfer size. If a command is submitted that exceeds the 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 no restrictions on transfer size. The restriction includes metadata if it is interleaved with the logical block data. The restriction does not apply to commands that do not transfer data between the host and the controller (e.g., Write Uncorrectable command or Write Zeroes command). 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.

Modify a portion of section 5.10 (Get Log Page command) as shown below:

There are mandatory and optional Log Identifiers defined in **Error! Reference source not found.** and **Error! Reference source not found.**. If a Get Log Page command is processed that specifies a Log Identifier that is not supported, then the controller should abort the command with status Invalid Field in Command.

The controller indicates if it supports the Log Page Offset and extended Number of Dwords (32 bits rather than 12 bits) in the Log Page Attributes field of the Identify Controller data structure. If extended data is not supported, then bits 27:16 of the Number of Dwords Lower field specify the Number of Dwords to transfer.

Modify a portion of Figure 90 (*Identify – Identify Controller Data Structure*) as shown below:

Figure 1: Identify – Identify Controller Data Structure

Bytes	O/M	Description
Controller Capabilities and Features		
. . .	. . .	. . .

Technical input submitted to the NVM Express™ Workgroup is subject to the terms of the NVM Express™ Participant's agreement. Copyright © 2014-16 NVMe™ Corporation.

258	M	**Abort Command Limit (ACL):** This field is used to convey the maximum number of concurrently ~~outstanding~~ executing Abort commands supported by the controller (see section **Error! Reference source not found.**). This is a 0's based value. It is recommended that implementations support concurrent execution of a minimum of four Abort commands ~~outstanding simultaneously~~.

Modify a portion of section 5.1 (Abort command) as shown below:

The Abort command is used to abort a specific command previously submitted to the Admin Submission Queue or an I/O Submission Queue. ~~Host software may have multiple Abort commands outstanding, subject to the constraints of the Abort Command Limit indicated in the Identify Controller data structure in Figure 90.~~ An Abort command is a best effort command; the command to abort may have already completed, currently be in execution, or may be deeply queued. It is implementation specific if/when a controller chooses to complete the command when the command to abort is not found.

The Abort Command Limit field in Identify Controller indicates the controller limit on concurrent execution of Abort commands. A host should not allow the number of outstanding Abort commands to exceed this value. The controller may complete any excess Abort commands with Abort Command Limit Exceeded status.

Modify a portion of section 1.6.21 (private namespace) as shown below:

1.6.21 private namespace

A namespace that ~~is accessible by~~ may only be attached to one controller at a time.

Modify a portion of section 1.6.23 (shared namespace) as shown below:

1.6.23 shared namespace

A namespace that ~~is accessible by~~ may be attached to two or more controllers in an NVM subsystem concurrently. A host may determine whether a namespace is a private namespace or may be a shared namespace by the value of the Namespace Multi-path I/O and Namespace Sharing Capabilities (NMIC) field in the Identify Namespace data structure.

Modify a portion of section 5.10.1.6.1 (Reservation Notification (Log Identifier 80h)) as shown below:

5.10.1.6.1 Reservation Notification (Log Identifier 80h)

A Reservation Notification log page is created whenever an unmasked reservation notification occurs on any namespace that ~~may be~~ ~~accessed by~~ is attached to the controller.

Modify a portion of Figure 92 (*Identify – Identify Namespace Data Structure*) as shown below:

30	O	**Namespace Multi-path I/O and Namespace Sharing Capabilities (NMIC):** This field specifies multi-path I/O and namespace sharing capabilities of the namespace. Bits 7:1 are reserved Bit 0: If set to '1' then the namespace may be ~~accessible by~~ attached to two or more controllers in the NVM subsystem concurrently (i.e., may be a shared namespace). If cleared to '0' then the namespace is a private namespace and may only be ~~accessed by the controller that returned this namespace data structure~~ attached to one controller at a time.

Technical input submitted to the NVM Express™ Workgroup is subject to the terms of the NVM Express™ Participant's agreement. Copyright © 2014-16 NVMe™ Corporation.

Modify a portion of section 5.15.1.17 (Reservation Notification Mask (Feature Identifier 82h)) as shown below:

5.15.1.17 Reservation Notification Mask (Feature Identifier 82h), (Optional[2])

. . .

A Set Features command that uses a namespace ID other than FFFFFFFFh modifies the reservation notification mask for the corresponding namespace only. A Set Features command that uses a namespace ID of FFFFFFFFh modifies the reservation notification mask of all namespaces that are ~~accessible by~~ attached to the controller and that support reservations.

Modify a portion of section 5.15.1.18 (Reservation Persistence (Feature Identifier 83h)) as shown below:

5.15.1.13 Reservation Persistence (Feature Identifier 83h), (Optional[3])

. . .

The PTPL state is contained in the Reservation Persistence Feature that is namespace specific. A Set Features command that uses the namespace ID FFFFFFFFh modifies the PTPL state associated with all namespaces that are ~~accessible by~~ attached to the controller and that support PTPL (i.e., support reservations).

Technical input submitted to the NVM Express™ Workgroup is subject to the terms of the NVM Express™ Participant's agreement. Copyright © 2014-16 NVMe™ Corporation.

Modify Figure 197 (Write – Command Dword 13) as shown below:

<table>
<tr><th>Bit</th><th>Description</th></tr>
<tr><td>31:08</td><td>Reserved</td></tr>
<tr><td>07:00</td><td>
Dataset Management (DSM): This field indicates attributes for the ~~dataset that the~~ LBA(s) being ~~read~~ written ~~are associated with~~.
<table>
<tr><th>Bits</th><th>Attribute</th><th>Definition</th></tr>
<tr><td>07</td><td>Incompressible</td><td>If set to ‘1’, then data is not compressible for the logical blocks indicated. If cleared to ‘0’, then no information on compression is provided.</td></tr>
<tr><td>06</td><td>Sequential Request</td><td>If set to ‘1’, then this command is part of a sequential write that includes multiple Write commands. If cleared to ‘0’, then no information on sequential access is provided.</td></tr>
<tr><td>05:04</td><td>Access Latency</td><td>
<table>
<tr><th>Value</th><th>Definition</th></tr>
<tr><td>00b</td><td>None. No latency information provided.</td></tr>
<tr><td>01b</td><td>Idle. Longer latency acceptable.</td></tr>
<tr><td>10b</td><td>Normal. Typical latency.</td></tr>
<tr><td>11b</td><td>Low. Smallest possible latency.</td></tr>
</table>
</td></tr>
<tr><td>03:00</td><td>Access Frequency</td><td>
<table>
<tr><th>Value</th><th>Definition</th></tr>
<tr><td>0000b</td><td>No frequency information provided.</td></tr>
<tr><td>0001b</td><td>Typical number of reads and writes expected for this LBA range.</td></tr>
<tr><td>0010b</td><td>Infrequent writes and infrequent reads to the LBA range indicated.</td></tr>
<tr><td>0011b</td><td>Infrequent writes and frequent reads to the LBA range indicated.</td></tr>
<tr><td>0100b</td><td>Frequent writes and infrequent reads to the LBA range indicated.</td></tr>
<tr><td>0101b</td><td>Frequent writes and frequent reads to the LBA range indicated.</td></tr>
<tr><td>0110b</td><td>One time write. E.g. command is due to virus scan, backup, file copy, or archive.</td></tr>
<tr><td>0111b – 1111b</td><td>Reserved</td></tr>
</table>
</td></tr>
</table>
</td></tr>
</table>

Technical input submitted to the NVM Express™ Workgroup is subject to the terms of the NVM Express™ Participant’s agreement. Copyright © 2014-16 NVMe™ Corporation.

Modify Figure 172 (Read – Command Dword 13) as shown below:

Bit	Description
31:08	Reserved
07:00	**Dataset Management (DSM):** This field indicates attributes for the ~~dataset that the~~ LBA(s) being read ~~are associated with~~.

Bits	Attribute	Definition
07	Incompressible	If set to '1', then data is not compressible for the logical blocks indicated. If cleared to '0', then no information on compression is provided.
06	Sequential Request	If set to '1', then this command is part of a sequential read that includes multiple Read commands. If cleared to '0', then no information on sequential access is provided.
05:04	Access Latency	
03:00	Access Frequency	

Value	Definition
00b	None. No latency information provided.
01b	Idle. Longer latency acceptable.
10b	Normal. Typical latency.
11b	Low. Smallest possible latency.

Value	Definition
0000b	No frequency information provided.
0001b	Typical number of reads and writes expected for this LBA range.
0010b	Infrequent writes and infrequent reads to the LBA range indicated.
0011b	Infrequent writes and frequent reads to the LBA range indicated.
0100b	Frequent writes and infrequent reads to the LBA range indicated.
0101b	Frequent writes and frequent reads to the LBA range indicated.
0110b	One time read. E.g. command is due to virus scan, backup, file copy, or archive.
0111b	Speculative read. The command is part of a prefetch operation.
1000b	The LBA range is going to be overwritten in the near future.
1001b – 1111b	Reserved

Technical input submitted to the NVM Express™ Workgroup is subject to the terms of the NVM Express™ Participant's agreement. Copyright © 2014-16 NVMe™ Corporation.