



LEGAL NOTICE:

© **Copyright 2007 - 2018 NVM Express, Inc. ALL RIGHTS RESERVED.**

This NVM Express revision 1.3 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 1.3 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 - 2018 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 VTM Group
3855 SW 153rd Drive
Beaverton, OR 97003 USA
info@nvmexpress.org

NVM Express Technical Proposal for New Feature

Technical Proposal ID	4030 – Verify Command
Change Date	9/17/2018
Builds on Specification	NVM Express 1.3

Technical Proposal Author(s)

Name	Company
David Black	Dell EMC
Mike Allison	Intel
Peter Onufryk	Microsemi
Christoph Hellwig	WD

This technical proposal adds a Verify command to check the integrity of stored data (and metadata). The Verify command is roughly equivalent to a Read command that discards the data (and metadata) after reading. The important behavior of Verify is that errors are reported if the data (or metadata) cannot be read without expending the time and resources required to return the data (and metadata) to the host.

Revision History

Revision Date	Change Description
2018/02/02	Initial version
2018/02/05	Specify interaction of Verify with deallocated logical blocks. Additional minor edits.
2018/02/08	Add mention of Verify command to description of protection information field in Format NVM command.
2018/02/13	Verify contributes to Data Units Read in SMART/Health Information Log Make it clear that integrity checks include Protection Information. Specify FUA and PRINFO fields in more detail.
2018/02/15	Don't add requirement that Write Uncorrectable be supported if Verify is supported. Open Issue: how host discovers Verify command support (Identify Controller bit vs. Commands Supported and Effects log page)
2018/02/22	Use Identify Controller bit in ONCS to indicate Verify command support – that resolves the open issue. Other minor edits for phase 2 completion.
2018/02/23	Fix command completion wording based on ECNs for other commands.
2018/02/26	Add summary block (above) fix date at top of this page, and fix figure reference in 6.NEW.1.
2018/03/19	Add a sentence to make it clear that Verify command can return Unrecovered Read Error status.
2018/03/29	Explicitly state that integrity checks and reading (e.g. Verify and Read commands) don't have to return the same error. Additional minor edits.
2018/04/05	Define "Verify operation".
2018/04/12	Add sentence to require that LR (Limited Retry) have same behavior in Verify and Read commands. Editorial changes, e.g., put all instances of "e.g.," into parentheses.
2018/04/18	Remove comment indicating that ECN text needs to be written for FUA bit in other commands, as that text has now been written.
2018/04/19	Revise FUA text to state that cached data is flushed from the volatile cache.
2018/04/26	Clean version for 30-day member review
2018/06/12	Member review completed, no changes made.
2018/07/03	Editorial corrections from Harvey Newman and Mike Allison
2018/09/17	Ratified

Description of Specification Changes

Modify Figure 34 (Status Code – Command Specific Status Values, NVM Command Set) as shown below:

Figure 1: Status Code – Command Specific Status Values, NVM Command Set

Value	Description	Commands Affected
80h	Conflicting Attributes	Dataset Management, Read, Write
81h	Invalid Protection Information	Compare, Read, Verify , Write, Write Zeroes
82h	Attempted Write to Read Only Range	Dataset Management, Write, Write Uncorrectable, Write Zeroes
83h - BFh	Reserved	

Modify Figure 36 (Status Code – Media and Data Integrity Error Values, NVM Command Set) as shown below:

Value	Description
87h	Deallocated or Unwritten Logical Block: The command failed due to an attempt to read from or verify an LBA range containing a deallocated or unwritten logical block.

Modify Figure 93 (Get Log Page – SMART / Health Information Log) as shown below:

Bytes	Description
47:32	Data Units Read: Contains the number of 512 byte data units the host has read from the controller; this value does not include metadata. This value is reported in thousands (i.e., a value of 1 corresponds to 1000 units of 512 bytes read) and is rounded up. When the LBA size is a value other than 512 bytes, the controller shall convert the amount of data read to 512 byte units. For the NVM command set, logical blocks read as part of Compare, and Read, and Verify operations shall be included in this value.

Modify Figure 109 (Identify – Identify Controller data structure) as shown below:

77	M	<p>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., Verify command, Write Uncorrectable command, or Write Zeroes command).</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>
521:520	M	<p>Optional NVM Command Support (ONCS): This field indicates the optional NVM commands and features supported by the controller. Refer to section 6.</p> <p>Bits 15:NEW+17 are reserved.</p> <p>Bit NEW if set to '1', then the controller supports the Verify command. If cleared to '0', then the controller does not support the Verify command.</p> <p>Bit 6 if set to '1' then the controller supports the Timestamp feature. If cleared to '0', then the controller does not support the Timestamp feature. Refer to section 5.21.1.14.</p> <p>Bit 5 if set to '1' then the controller supports reservations. If cleared to '0' then the controller does not support reservations. If the controller supports reservations, then it shall support the following commands associated with reservations: Reservation Report, Reservation Register, Reservation Acquire, and Reservation Release. Refer to section 8.8 for additional requirements.</p> <p>Bit 4 if set to '1' then the controller supports the Save field set to a non-zero value in the Set Features command and the Select field set to a non-zero value in the Get Features command. If cleared to '0' then the controller does not support the Save field set to a non-zero value in the Set Features command and the Select field set to a non-zero value in the Get Features command.</p> <p>Bit 3 if set to '1' then the controller supports the Write Zeroes command. If cleared to '0' then the controller does not support the Write Zeroes command.</p> <p>Bit 2 if set to '1' then the controller supports the Dataset Management command. If cleared to '0' then the controller does not support the Dataset Management command.</p> <p>Bit 1 if set to '1' then the controller supports the Write Uncorrectable command. If cleared to '0' then the controller does not support the Write Uncorrectable command.</p> <p>Bit 0 if set to '1' then the controller supports the Compare command. If cleared to '0' then the controller does not support the Compare command.</p>

Modify Figure 176 (Format NVM – Command Dword 10) as shown below:

07:05	Protection Information (PI): This field specifies whether end-to-end data protection is enabled and the type of protection information. The values for this field have the following meanings:	
	Value	Definition
	000b	Protection information is not enabled
	001b	Protection information is enabled, Type 1
	010b	Protection information is enabled, Type 2
	011b	Protection information is enabled, Type 3
	100b – 111b	Reserved
When end-to-end data protected is enabled, the host shall specify the appropriate protection information in the Read, Verify , Write, or Compare commands.		

Modify a portion of section 6 (NVM Command Set) as shown below:

In the case of Compare, Read, **Verify**, Write, and Write Zeroes commands, the host may indicate whether a time limit should be applied to error recovery for the operation by setting the Limited Retry (LR) field in the command. The time limit is specified in the Error Recovery feature, specified in section 5.21.1.5. If the host does not specify a time limit should be applied, then the controller should apply all error recovery means to complete the operation.

Figure 2: Opcodes for NVM Commands

Opcode by Field			Combined Opcode ²	O/M ¹	Command ³
(07)	(06:02)	(01:00)			
Standard Command	Function	Data Transfer ⁵			
0b	000 00b	00b	00h	M	Flush
0b	000 00b	01b	01h	M	Write
0b	000 00b	10b	02h	M	Read
0b	000 01b	00b	04h	O	Write Uncorrectable
0b	000 01b	01b	05h	O	Compare
0b	000 10b	00b	08h	O	Write Zeroes
0b	000 10b	01b	09h	O	Dataset Management
0b	000 11b	00b	0Ch	O	Verify
0b	000 11b	01b	0Dh	O ⁴	Reservation Register
0b	000 11b	10b	0Eh	O ⁴	Reservation Report
0b	001 00b	01b	11h	O ⁴	Reservation Acquire
0b	001 01b	01b	15h	O ⁴	Reservation Release
Vendor Specific					
1b	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 options to the command shall transfer data as specified or transfer no data. All commands, including vendor specific commands, shall follow this convention: 00b = no data transfer; 01b = host to controller; 10b = controller to host; 11b = bidirectional.

Modify a portion of section 6.7.1.1 (Deallocate) as shown below:

6.7.1.1 Deallocate

A logical block that has been deallocated using the Dataset Management command is no longer deallocated when the logical block is written. Read operations **and Verify operations** do not affect the deallocation status of a logical block. The value read from a deallocated logical block shall be deterministic; specifically, the value returned by subsequent reads of that logical block shall be the same until a write occurs to that logical block.

The values read from a deallocated logical block and its metadata (excluding protection information) shall be all bytes set to 00h, all bytes set to FFh, or the last data written to the associated logical block and its metadata, except that access is prohibited to all data and metadata values written before the most recent successful sanitize operation, if any. The Deallocate Logical Block Features field in the Identify Namespace data structure may report the values read from a deallocated logical block and its metadata.

The values read from a deallocated or unwritten logical block's protection information field shall:

- have the Guard field value set to FFFFh or set to the CRC for the value read from the deallocated logical block and its metadata (excluding protection information) (e.g., set to 0000h if the value read is all bytes set to 00h); and
- have the Application Tag field value set to FFFFh and the Reference Tag field value set to FFFFFFFFh (indicating the protection information shall not be checked).

Host software may enable an error to be returned if a deallocated or unwritten logical block is read in the Error Recovery feature. If this error is supported for the namespace and enabled, then a **~~r~~Read, ~~V~~erify, or ~~e~~Compare command that includes ~~containing~~** a deallocated or unwritten logical block shall fail with the Unwritten or Deallocated Logical Block status code. Note: Legacy software may not handle an error for this case.

Add a new section 6.NEW after section 6.13, renumbering as needed, as shown below:

(All text and tables in this new section [6.NEW and 6.NEW.1] are new, and hence are not marked in red.)

6.NEW Verify command

The Verify command verifies integrity of stored information by reading data and metadata, if applicable, for the LBAs indicated without transferring any data or metadata to the host. A Verify operation consists of the controller actions (e.g., reading) that verify integrity of stored information during execution of a Verify command. The command may specify protection information to be checked as part of the Verify operation.

Verify operations may be implemented via integrity checks of stored data and metadata. Metadata integrity checks shall include protection information if the Verify command specifies checking of protection information and the namespace is formatted with protection information.

If reading the data and metadata, if applicable, would result in an error being returned, then an error shall be returned as a result of the Verify operation on that data and metadata, if applicable. In this situation, the error that results from integrity checks may differ from the error that would result from reading (e.g., there is no requirement that the Verify and Read commands return the same error). Setting the Limited Retry (LR) bit to '1' shall have the same effect in both the Read and Verify commands.

All data that is read or has its integrity checked by a Verify operation shall be included in the value of the Data Units Read field in the SMART/Health Information log page, refer to 5.14.1.2.

The command uses Command Dword 10, Command Dword 11, Command Dword 12, Command Dword 14, and Command Dword 15 fields.

Figure NEW-1: Verify – Command Dword 10 and Command Dword 11

Bit	Description
63:00	Starting LBA (SLBA): This field indicates the 64-bit address of the first logical block of data to be verified as part of the operation. Command Dword 10 contains bits 31:00; Command Dword 11 contains bits 63: 32.

Figure NEW-2: Verify – Command Dword 12

Bit	Description
31	Limited Retry (LR): If set to '1', then the controller should apply limited retry efforts. If cleared to '0', then the controller should apply all available error recovery means before completing the command with failure.
30	Force Unit Access (FUA): If set to '1', then the controller shall flush any data and metadata specified by the Verify command from any volatile cache before performing the Verify operation and shall perform the Verify operation on data and metadata that have been committed to non-volatile media. There is no implied ordering with other commands. If cleared to '0', then this bit has no effect.
29:26	Protection Information Field (PRINFO): Specifies the protection information action and check field, as defined in Figure 196. The Protection Information Check (PRCHK) field in the PRINFO field specifies the protection information to be checked by the Verify operation. The Protection Information Action (PRACT) bit in the PRINFO field is cleared to '0' by the host. If the PRACT bit is not cleared to '0', then the controller shall abort the command with a status of Invalid Field in Command.
25:16	Reserved
15:00	Number of Logical Blocks (NLB): This field indicates the number of logical blocks to be verified. This is a 0's based value.

Figure NEW-3: Verify – Command Dword 14

Bit	Description
31:00	Expected Initial Logical Block Reference Tag (EILBRT): This field specifies the Initial Logical Block Reference Tag expected value. This field is only used if the namespace is formatted to use end-to-end protection information. Refer to section 8.3.

Figure 3: NEW-4: Verify – Command Dword 15

Bit	Description
31:16	Expected Logical Block Application Tag Mask (ELBATM): This field specifies the Application Tag Mask expected value. This field is only used if the namespace is formatted to use end-to-end protection information. Refer to section 8.3.
15:00	Expected Logical Block Application Tag (ELBAT): This field specifies the Application Tag expected value. This field is only used if the namespace is formatted to use end-to-end protection information. Refer to section 8.3.

6.NEW.1 Command Completion

Upon completion of the Verify command, the controller posts a completion queue entry (CQE) to the associated I/O Completion Queue. The status code types and values that may be used in a CQE for the Verify command include the status code type and status code values for all Media and Data Integrity Errors for the NVM Command Set that are applicable to the Read command (e.g., Unrecovered Read Error). Refer to Figure 30 and to Figure 36.

Verify command specific status values are defined in Figure NEW-5.

Figure 4: Verify – Command Specific Status Values

Value	Description
81h	Invalid Protection Information: The Protection Information Field (PRINFO) (refer to Figure NEW-2) settings specified in the command are invalid for the Protection Information with which the namespace was formatted (refer to the PI field in Figure 176 and the DPS field in Figure 114) or the EILBRT field is invalid (refer to section 8.3.1.5).

Modify Figure 268 (Command Behavior in the Presence of a Reservation) as shown below:

Figure 5: Command Behavior in the Presence of a Reservation

NVM Command	Write Exclusive Reservation		Exclusive Access Reservation		Write Exclusive Registrants Only or Write Exclusive All Registrants Reservation		Exclusive Access Registrants Only or Exclusive Access All Registrants Reservation	
	Non-Registrant	Registrant	Non-Registrant	Registrant	Non-Registrant	Registrant	Non-Registrant	Registrant
NVM Read Command Group: Read Compare Verify Security Receive (Admin)	A	A	C	C	A	A	C	A

