LEGAL NOTICE:

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, Inc.
3855 SW 153rd Drive
Beaverton, OR 97003
info@nvmexpress.org

NVM Express Technical Proposal for New Feature

Technical Proposal ID	**4005a – Namespace Write Protect**
Change Date	**05/29/2018**
Builds on Specification	**NVM Express 1.3**

Technical Proposal Author(s)

Name	Company
Dave Landsman	Western Digital
Christoph Hellwig	Western Digital
Nadesan Narenthiran	Western Digital
John Carroll	Intel
Anthony Constantine	Intel
Michael Allison	Intel
Lee Prewitt	Microsoft
Kevin Marks	Dell

This technical proposal defines enhancements for write protection for namespaces.

Namespace Write Protection is a new (optional) Feature identifier that controls write protection on a per namespace basis. This Feature may be used to prevent modification of the specified namespace. The controller fails commands that attempt to modify the namespace while it is write protected.

Revision History

Revision Date	Change Description
09/03/2015	Initial proposal as drafted by John Carroll when we split Boot and WP
05/10/2017	Updates after sub-team meetings in Jun-2016 • Resolved some issues, such as Sanitize interactions, which were outside of the TP • Added two new bits/meta-modes o Disable capability to set Power On write protection o Disable capability to set Permanent Write Protect
07/18/2017	• Major architectural addition: Authenticated access, via RPMB, for Write Protect Until Reset, and Permanent Write Protect states • Numerous other updates for consistent naming, and simply fleshing out the proposal.
08/02/2017	• Added informative note in 8.TBD2 re: external WP features. • Added text in 8.TBD2 and 8.10 establishing a default for Namespace Write Protect State Control bits following an NVM Subsystem Reset.
08/03/2017	• Few updates from 8/3 TG review.
09/14/2017	• Many editorials • Changed feature to explicitly “not saveable”
09/27/2017	• Went back to “not savable / persistent”. • Declared “default” value as “no write protect”. • Other editorial fix ups.
9/28/2017	• Resolved remaining comments from 9/27 version. • Removed Set Features arc from Write Protect Until Reset state to Permanent Write Protect state; this was not intended to be allowed.
10/06/2017	• Resolved editorials from 10/05 WG and VTM (Harvey Neumann)
10/12/2017	• Substantive o Removed “default value” for the feature; the value of the Feature is determined by the write protection state of the namespace. Changed “persistent across power cycle and resets” in features table back to NO. The feature itself has no persistent values; the namespaces do. o Changed name of “WP Until Reset” state to be “WP Until Power Cycle”; this is more precise and less confusing. • Editorial o Added back theory of operation and other rearrangements o moved the “fail with status code Feature Not Changeable” down to the feature definition itself; o Separated “if controller supports write protection” from “if controller supports write protection and either Permanent WP or WP Until Power Cycle”, in order not have the embedded “if”. o updated all the other things discussed on previous review
10/29/2017	• Added language to describe difference between power cycles and Controller Level Resets regarding WP state persistence rules, Authentication Control, and WP Config feature (8.TBD2, 8.10, and 5.2.1.TBD). • Cleaned up/restructured Namespace Authentication Control language in 8.TBD2 and 8.TBD2.1. • Added language in 8.TBD2.1 to cover behavior of commands like Sanitize, which must not change a namespace which is write protected but do not use an NSID. • Added language in WP Config (5.22.1.TBD) which specifies command failure modes related to combination of Permanent and Until Power Cycle states, and their respective RPMB control bits. • Added text to change Sanitize Command Restrictions to explicitly disallow WP Config command when Sanitize in progress.
11/13/2017	• Distinguish media error vs. host initiated namespace write protect in various commands • Various editorial
12/04/2017	• Added statement in 8.TBD2.1 which clarifies the behavior of multi-controller subsystems in the presence of write protected namespaces; • Various editorial changes.
12/09/2017	• Accepted final changes and submitted for 30-day review.
01/10/2018	• Accepted few editorial comments from 30-day review.
1/23/2018	• Ratified

03/08/2018	• Added the requirement that all volatile write cache data and meta data is flush when transitioning to a write protected state.
03/15/2018	• Clarified requirement added 03/08/2018. Added Flush command as a command allowed while in the write protected state.
05/10/2018	• Moved Namespace Write Protection Config from figure 134 to figure 135 because the feature is namespace specific
05/11/2018	• After integration by technical writer, changed section 5.22 to 5.21, fixed figure titles, applied text to match NVMe 1.3 version that is not changed, updated TBD values for fields, and applied editorial changes.
5/29/2018	• Ratified

Description of Specification Changes

Add a new section 8.TBD2 as shown below:

8.TBD2 Namespace Write Protection (Optional)

Namespace Write Protection is an optional configurable controller capability that enables the host to control the write protection state of a namespace. Support for this capability is reported in the Namespace Write Protection Capabilities (NWPC) field in the Identify Controller data structure (refer to Figure 109, section 5.15).

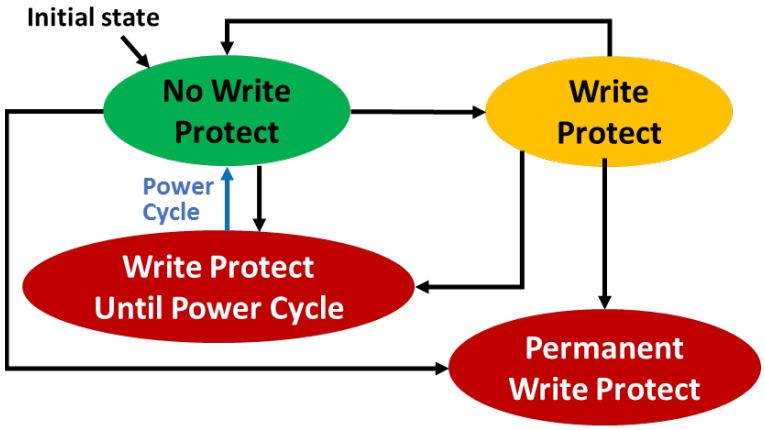
Figure TBD0 defines the write protection states that may be supported for a namespace. All states persist across power cycles and NVMe Controller Level Resets (refer to section 7.3) except Write Protect Until Power Cycle state, which is defined to transition to the No Write Protect state on the occurrence of a power cycle.

Figure TBD0 – Namespace Write Protection State Definitions

State	Definition	Persistent Across	
		Power Cycles	Controller Level Resets
No Write Protect	The namespace is not write protected.	Yes	Yes
Write Protect	The namespace is write protected.	Yes	Yes
Write Protect Until Power Cycle	The namespace is write protected until the next power cycle.	No	Yes
Permanent Write Protect	The namespace is permanently write protected.	Yes	Yes

Figure TBD1 defines the transition between write protection states. All state transitions are based on Set Features commands unless otherwise specified. The initial state of a namespace at the time of its creation is the No Write Protect state.

Figure TBD1 – Namespace Write Protection State Machine Model



The Write Protect Until Power Cycle and Permanent Write Protect states are subject to the Namespace Write Protection Authentication Control mechanism, which determines whether the controller processes or aborts Set Features commands which cause a transition into either of these two states (refer to section 8.10).

The results of using Namespace Write Protection in combination with an external write protection system (e.g., TCG Storage Interface Interactions Specification) are outside the scope of this specification.

8.TBD2.1 Namespace Write Protection – Theory of Operation

If Namespace Write Protection is supported by the controller, then the controller shall:

- Indicate the level of support for Namespace Write Protection capabilities in the Namespace Write Protection Capabilities (NWPC) field in the Identify Controller data structure; and
- Support the Namespace Write Protection Config Feature (refer to section 5.21.1.TBD).

If the Write Protect Until Power Cycle or the Permanent Write Protect states are supported by the controller, then the controller shall support the Namespace Write Protection Authentication Control field in the RPMB Device Configuration Block data structure (refer to section 8.10).

The controller shall not set the Critical Warning field, bit 3 (refer to Figure 93) if the read-only condition on the media is a result of the host using the Namespace Write Protection Config Feature, or due to any autonomous namespace write protection state transitions (e.g., power cycle). Host software may check the current namespace write protection state of a namespace using the Get Features command with the Namespace Write Protection Config Feature Identifier.

If any controller in the NVM subsystem supports Namespace Write Protection, then the write protection state of a namespace shall be enforced by any controller to which that namespace is attached.

8.TBD2.1.1 Namespace Write Protection – Command Interactions

Unless otherwise noted, the commands listed in Figure TBD2 are processed normally when specifying an NSID for a namespace that is write protected.

Figure TBD2 – Commands Allowed when Specifying a Write Protected NSID

Admin Command Set	NVM Command Set
Device Self-Test	Compare
Directive Send[1]	Dataset Management[1]
Directive Receive	Read
Get Features	Reservation Register
Get Log Page	Reservation Report
Identify	Reservation Acquire
Namespace Attachment	Reservation Release
Security Receive[1]	Vendor Specific[1]
Security Send[1]	Flush[2]
Set Features[1]	
Vendor Specific[1]	

NOTES:
1. The controller shall fail commands if the specified action attempts to modify the medium of the specified namespace.
2. A Flush command shall complete successfully with no effect. All volatile write cache data and metadata associated with the specified namespace is written to non-volatile media as part of transitioning to the write protected state (refer to section 5.21.1.TBD).

Commands not listed in Table TBD2, and which meet the following conditions, shall be aborted with a status code of Namespace Is Write Protected (refer to Figure 31):

a) Commands that specify an NSID for a namespace that is write protected;

b) Commands that specify an NSID for a namespace that is not write protected and the execution of which would modify another namespace that is write protected (e.g., a Format NVM command); and
c) Commands which do not specify an NSID, and the execution of which would modify a namespace that is write protected (e.g., Sanitize).

Modify a portion of Figure 135: Set Features, NVM Command Set Specific – Feature Identifiers, as shown below:

Feature Identifier	O/M	Persistent Across Power Cycle and Reset[2]	Uses Memory Buffer for Attributes	Description
84h	O	No	No	Namespace Write Protection Config

Modify a portion of Figure 109 (Identify – Identify Controller Data Structure) as shown below:

531	M	~~Reserved~~ **Namespace Write Protection Capabilities (NWPC):** This field indicates the optional namespace write protection capabilities supported by the controller. Refer to section 8.TBD2. Bits 7:3 are reserved. Bit 2 if set to '1', then the controller supports the Permanent Write Protect state. If cleared to '0', then the controller does not support the Permanent Write Protect state. If this bit is set to '1', then the controller shall support the Namespace Write Protection Authentication field (refer to section 8.10). Bit 1 if set to '1', then the controller supports the Write Protect Until Power Cycle state. If cleared to '0', then the controller does not support Write Protect Until Power Cycle state. If this bit is set to '1', then the controller shall support the Namespace Write Protection Authentication field (refer to section 8.10). Bit 0 if set to '1', then the controller shall support the No Write Protect and Write Protect namespace write protection states and may support the Write Protect Until Reset and Permanent Write Protect namespace write protection states (refer to section 8.TBD2). If cleared to '0', then the controller does not support Namespace Write Protection and bits 2:1 shall be cleared to 00b.

Modify a portion of Figure 114 (Identify – Identify Namespace Data Structure, NVM Command Set Specific) as shown below:

99	O	**Namespace Attributes (NSATTR):** This field specifies attributes of the namespace. Bits 7:1 are reserved. Bit 0: If set to '1', then the namespace is currently write protected due to any condition (e.g., namespace write protection set for the namespace, media errors) and all write access to the namespace shall fail. If cleared to '0', then the namespace is not currently write protected.

Modify a portion of Figure 84 (Get Features – Feature Identifiers) as shown below:

Description	Section Defining Format of Attributes Returned
…	…
Reservation Persistence	Section 5.21.1.21
Namespace Write Protection Config	Section 5.21.1.TBD

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

Figure 31: Status Code – Generic Command Status Values

Value	Description
20h	**Namespace is Write Protected:** The command is prohibited while the namespace is write protected by the host.

Add a new section 5.21.1.TBD as shown below:

5.21.1.TBD Namespace Write Protection Config (Feature Identifier 84h)

This Feature is used by the host to configure the namespace write protection state. Refer to section 8.TBD2 for definition and behaviors of the namespace write protection states. The settings are specified in Command Dword 11.

This Feature is not savable (refer to section 7.8). There is no default value for this Feature; the value of the Feature after a power cycle or a Controller Level Reset is determined by the write protection state of the namespace prior to the power cycle or Controller Level Reset, except for the Write Protect Until Power Cycle write protection state (refer to section 8.TB2).

If a Get Features command is submitted for this Feature, the attributes specified in Figure 5.21.1.TBD1 are returned in Dword 0 of the completion queue entry for that command.

Figure 5.21.1.TBD1: Write Protection – Command Dword 11

<table>
<tr><th>Bit</th><th>Description</th></tr>
<tr><td>31:03</td><td>Reserved</td></tr>
<tr><td>02:00</td><td>Write Protection State: This field specifies the write protection state of the specified namespace.
<table>
<tr><th>Value</th><th>Definition</th></tr>
<tr><td>000b</td><td>No Write Protect</td></tr>
<tr><td>001b</td><td>Write Protect</td></tr>
<tr><td>010b</td><td>Write Protect Until Power Cycle</td></tr>
<tr><td>011b</td><td>Permanent Write Protect</td></tr>
<tr><td>100b to 111b</td><td>Reserved</td></tr>
</table>
</td></tr>
</table>

If a Set Features command attempts to change the namespace write protection state of a namespace that is in the Write Protect Until Power Cycle state or the Permanent Write Protect state, then the command shall fail with a status code of Feature Not Changeable.

If a Set Features command attempts to change the namespace write protection state of a namespace to the Write Protect Until Power Cycle state and bit 0 of the of the Write Protection Authentication Control field is cleared to ‘0’, then the command shall fail with a status code of Feature Not Changeable.

If a Set Features command attempts to change the namespace write protection state of a namespace to the Permanent Write Protect state and bit 1 of the of the Write Protection Authentication Control field is cleared to ‘0’, then the command shall fail with a status code of Feature Not Changeable.

If a Set Features command changes the namespace to a write protected state, then the controller shall commit all volatile write cache data and metadata associated with the specified namespace to non-volatile media as part of transitioning to the write protected state.

Modify a portion of section 8.10 (Replay Protected Memory Block (Optional)) as shown below:

Figure 269: RPMB Device Configuration Block Data Structure

Bytes	Component Name	Description
02	Namespace Write Protection Authentication Control	This field specifies whether the controller processes or aborts Set Features commands which enable certain namespace write protection states (refer to section 8.TBD2 and 5.21.1.TBD). If the controller does not support Namespace Write Protection, then this field shall be cleared to 0h. If the controller supports Namespace Write Protection, then bits 1:0 of this field shall be cleared to 00b after a power cycle or a Controller Level Reset. Bits 7:2 are reserved. Bit 1: If cleared to '0', indicates that the controller shall fail a Set Features command which attempts to set the namespace write protection state to Permanent Write Protect, as defined in section 8.TBD2. If set to '1', indicates that the controller shall process a Set Features command which attempts to set the namespace write protection state to Permanent Write Protect. Bit 0: If cleared to '0', indicates that the controller shall fail a Set Features command which attempts to set the namespace write protection state to Write Protect Until Power Cycle, as defined in section 8.TBD2. If set to '1', indicates that the controller shall process a Set Features command which sets the namespace write protection state to Write Protect Until Power Cycle.
511:023		Reserved

Modify a portion of section 8.15.1 (Command Restrictions) as shown below:

8.15.1 Sanitize Command Restrictions

While performing a sanitize operation and while a failed sanitize operation has occurred but successful recovery from that failure has not occurred, all enabled controllers and namespaces in the NVM subsystem are restricted to performing only a limited set of actions.

While a sanitize operation is in progress:

- All controllers in the NVM subsystem shall only process the Admin commands listed in Figure 287 subject to the additional restrictions stated in that figure;
- All I/O Commands shall be aborted with a status of Sanitize in Progress; and
- Any command or command option that is not explicitly permitted in Figure 287 shall be aborted with a status of Sanitize in Progress if fetched by any controller in the NVM subsystem.

…

Figure 287: Sanitize Operations – Admin Commands Allowed

<table>
<tr><th>Admin Command</th><th>Additional Restrictions</th></tr>
<tr><td>Abort</td><td></td></tr>
<tr><td>Asynchronous Event Request</td><td></td></tr>
<tr><td>Create I/O Completion Queue</td><td></td></tr>
<tr><td>Create I/O Submission Queue</td><td></td></tr>
<tr><td>Delete I/O Completion Queue</td><td></td></tr>
<tr><td>Delete I/O Submission Queue</td><td></td></tr>
<tr><td>Get Features</td><td></td></tr>
<tr><td>Get Log Page</td><td>The log pages allowed are listed below.

<table><tr><th>Log Pages</th><th>Additional Restrictions</th></tr><tr><td>Error Information</td><td>Return zeros in the LBA field.</td></tr><tr><td>SMART / Health Information</td><td></td></tr><tr><td>Changed Namespace List</td><td></td></tr><tr><td>Reservation Notification</td><td></td></tr><tr><td>Sanitize Status</td><td></td></tr></table></td></tr>
<tr><td>Identify</td><td></td></tr>
<tr><td>Keep Alive</td><td></td></tr>
<tr><td>Set Features</td><td>Namespace Write Protection Config Feature is not allowed.</td></tr>
<tr><td>Opcode 7Fh</td><td>The Fabric Commands allowed are listed below. Refer to the NVMe over Fabrics specification.

<table><tr><th>Fabrics Commands</th><th>Additional Restrictions</th></tr><tr><td>Property Set</td><td></td></tr><tr><td>Connect</td><td></td></tr><tr><td>Property Get</td><td></td></tr><tr><td>Authentication Send</td><td></td></tr><tr><td>Authentication Receive</td><td></td></tr><tr><td>Vendor Specific</td><td>Commands are allowed that do not affect or retrieve user data.</td></tr></table></td></tr>
<tr><td>Vendor Specific</td><td>Commands are allowed that do not affect or retrieve user data.</td></tr>
</table>

Modify a portion of Figure 48, Asynchronous Event Information – SMART/Health Status, as shown below:

Figure 48: Asynchronous Event Information – SMART / Health Status

Value	Description
0h	**NVM subsystem Reliability:** NVM subsystem reliability has been compromised. This may be due to significant media errors, an internal error, the media being placed in read only mode, or a volatile memory backup device failing. This status value shall not be used if the read-only condition on the media is due to a change in the write protection state of a namespace (refer to section 8.TBD2.1)
…	**…**
3h - FFh	Reserved

Modify a portion of Figure 93, Get Log Page - SMART/Health Information Log, as shown below:

Figure 93: Get Log Page – SMART / Health Information Log

Bytes	Description
0	**Critical Warning:** This field indicates critical warnings for the state of the controller. Each bit corresponds to a critical warning type; multiple bits may be set. If a bit is cleared to '0', then that critical warning does not apply. Critical warnings may result in an asynchronous event notification to the host. Bits in this field represent the current associated state and are not persistent.

Bit	Definition
…	…
3	If set to '1', then the media has been placed in read only mode. The controller shall not set this bit to '1' if the read-only condition on the media is a result of a change in the write protection state of a namespace (refer to section 8.TBD2.1)
…	…

Modify a portion of Figure 209, Dataset Management – Command Specific Status Values, as shown below:

Figure 1: Dataset Management – Command Specific Status Values

Value	Description
80h	**Conflicting Attributes:** The attributes specified in the command are conflicting.
82h	**Attempted Write to Read Only Range:** The controller may optionally report this status if a Deallocate is attempted for a read only range. The controller shall not set this status value if the read-only condition on the media is a result of a host directed change in the write protection state of a namespace (refer to section 8.TBD2).

Modify a portion of Figure 242, Write – Command Specific Status Values, as shown below:

Figure 242: Write – Command Specific Status Values

Value	Description
…	**…**
82h	**Attempted Write to Read Only Range:** The LBA range specified contains read-only blocks. The controller shall not set this status value if the read-only condition on the media is a result of a host directed change in the write protection state of a namespace (refer to section 8.TBD2).

Modify a portion of Figure 245, Write Uncorrectable – Command Specific Status Values, as shown below:

Figure 245: Write Uncorrectable - Command Specific Status Values

Value	Description
…	**…**
82h	**Attempted Write to Read Only Range:** The LBA range specified contains read-only blocks**.** The controller shall not set this status value if the read-only condition on the media is a result of a host directed change in the write protection state of a namespace (refer to section 8.TBD2).

Modify a portion of Figure 250, Write Zeroes – Command Specific Status Values, as shown below:

Figure 250: Write Zeroes – Command Specific Status Values

Value	Description
…	…
82h	**Attempted Write to Read Only Range:** The LBA range specified contains read-only blocks. The controller shall not set this status value if the read-only condition on the media is a result of a host directed change in the write protection state of a namespace (refer to section 8.TBD2).