



LEGAL NOTICE:

© **Copyright 2008 to 2022 NVM Express®, Inc. ALL RIGHTS RESERVED.**

This 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 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: "**© 2008 to 2022 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	TP4055 Key per I/O
Change Date	2022-12-01
Builds on Specification	NVM Express NVM Command Set Specification 1.0a Zoned Namespace Specification 1.1a Key Value Command Set Specification 1.0a
Builds on Technical Proposal	TP 4095a Namespace Capability Reporting

Technical Proposal Author(s)

Name	Company
Michael Allison, Adrian Pearson, Thomas Bowen, Jonathan Hughes	Intel
David Black, Amnon Izhar, Judith Furlong	DellEMC
Fred Knight	NetApp
Michael Allison, Bill Martin, Eric Hibbard, Judy Brock	Samsung

This technical proposal addresses the following:

- Existing Self Encrypting Drive standards such as those described in the joint whitepaper between NVM Express and Trusted Computing Group (https://nvmexpress.org/wp-content/uploads/TCGandNVMe_Joint_White_Paper-TCG_Storage_Opal_and_NVMe_FINAL.pdf) perform encryption on user accessible data based on contiguous LBA ranges per namespace all the time, at PCIe interface speeds using keys generated/held in NVM by the storage device.
- This proposal makes use of encryption keys to be managed and downloaded securely by the host to the NVM subsystem, utilizing means outside the scope of this proposal. Encryption of user data is based on each I/O command indication which key to utilize for encryption of that command. This provides a finer granularity of data encryption that enables:
 - Easier support of European Union's General Data Protection Regulations' "Right to be forgotten".
 - Easier support of data erasure when data is spread over many disks (e.g., RAID/Erasure Coded) and mixed with other data needing to be preserved
 - Assigning an encryption key to a single sensitive file or host object
- The securely downloaded keys are maintained in volatile memory that are lost when powered off providing stronger confidentiality assurance against physical theft/loss of device.

Revision History

Revision Date	Change Description
2019/05/14	Initial version
2019/07/24	Updated after completing offline alignment discussion and subsequent review in NVMe Technical working group on June 13. Also added document header explaining the justification for the feature proposal and defined interactions with NVMe Sanitize, Format NVM, RPMB, Boot partitions, CMB, and PMR.
2019/8/13	Made the use of SDSPEC optional if SDTYPE is cleared to 0h. On commands that support DSM in these cases, Command Dword 13 have two different formats.
2019/8/26	Added a separate bits for supporting Key Per I/O and multiple directives. Updated to reference NVMe 1.4 and changed multiple locations to include TBD references instead of specific section numbers.
2019/10/2	Removed 2 nd directive (SDSPEC and SDTYPE fields) and replaced with using bits 15:00 in DWORD 13. Moved DSM to a directive when KPIO OACS bit is set.
2019/10/23	Added fields to allow the Host know whether transitioning to KPIO will affect all namespaces in the controller, as well as a notification of when the namespace is in KPIO mode.
2019/10/30	Added CETYPE (Command Extension Type) and CEV (Command Extension Value) fields to all affected commands. These new fields allow each affected command to specify whether the command is providing a Key Per I/O Tag value in the command. Also specified the error condition when the command does not specify a CETYPE or KPIOTAG when the namespace is set into KPIO mode.
2019/11/13	Made the CEV fields in commands refer to the figure defining the definition based on the CETYPE value. Other editorial edits
2019/11/18	Cleaned up the comments that were already addressed.
2019/12/2	Editorial updates from Paul Suhler.
2019/12/4	Phase 3 formatting by aligning tables and highlighting references.
2019/12/5	Removed DSM. Removed the questions 1 & 2 at the end as not part of the TP.
2019/12/12	Added a new status code for TCG usage to error out on a command that uses an invalid keytag as defined by TCG.
2020/1/2	Adding clarification statement for sanitize.
2020/1/30	Removed adding the introduction section as that can wait until TCG completes its specification. Modified nested conditionals in text to bullets for clarity. Changed a requirement that listed each command to just state the commands that support the new field to avoid having to update the text from each new read/write command added.
2020/1/30	Added references and editorial changes in the Technical WG meeting. Removed comments on presented material.
2020/2/13	Added the word rejected and added clarified read commands and write commands.
2020/4/10	Added Copy command information and TP 4068 PI Enhancement changes
2020/4/13	Fixing unchanged black text in Command Data Words to match NVMe 1.4.
2020/4/15	Copy command is changing so matching the changes. Fixed allocation of 2 byte field into 4 byte field.
2020/4/16	Ready for integration
2020/04/21	Integrated into NVM Express Base Specification
2020/04/23	Ready for ratification
2020/05/28	Added CETYPE and CEV to the Compare command.
2020/06/05	Accepted all changes.

Revision Date	Change Description
2020/07/20	Integrated into the NVM Express Base Specification.
2020/07/31	Added the ZNS Append command
2020/08/10	Matched TP 4053a change to add Directives to Zone Append command.
2020/08/27	Editorial fix to CEV description.
2020/10/22	Allowing vendor specific way to manage the keys used by Key Per I/O.
2020/12/09	Removed the word “first” in Figure NEW-4.
2021/01/07	Editorial fixes per Technical WG review.
2021/02/08	Updated description to remove items for TCG to define.
2021/03/31	Aligned to NVMe 2.0 specifications.
2021/04/02	Per Technical WG direction, updated Key Per I/O to be only a capability for the I/O command sets that supports LBAs. Defined Key Tag to be scoped on a per namespace basis. Added a new field specifying the available Key Tag values allowed in an I/O command in an active namespace. Reserved a value of the CEV to be vendor specific.
2021/06/24	Updated to the released version of the NVMe 2.0 Library of Specifications.
2021/11/04	Aligned to latest specification and formats for NVMe 2.0 library of specification. Moving the definition of KPIO to the NVMe Base Specification and the scope is namespace and is independent of any I/O Command Set as a feature; however there are specific I/O Command Set exceptions. Added KPIO support for the Key Value Command Set.
2021/11/18	Minor edits during review in NVMe Technical WG on 11/18/2021.
2021/12/01	Updates the TCG document name to be “I/O” instead of “IO”. Moved CEV field usage in the Key Value Command Set Specification to use Command DW13 to align with NVM Command Set Specification changes. Added decision on CETYE in the Key Value Command Set Specification to use either CDW11 or CDW12 – question if we want to just take part of a whole CDW just to align to NVM Command Set Specification.
2021/12/9	Moved all of the KPIO fields in the Key Value Command Set Specification to CDW13. Added clarity to the KPIO bit that it is to be ignored if KPIO is not supported.
2021/12/22	Rolled the copyright date. Added section numbers for the Navigation pane. Editorial fixes from Paul Suhler.
2021/12/28	Added a comment. Separated some sections on page breaks.
2022/1/27	Moved the sanitize operation requirements in the Key Per I/O capability section to Sanitize Operation section with editorial changes. Updated sections so the navigation pane should the complete nesting of the sections.
2022/02/06	Integrated
2022/02/09	Removed the second KPIOS acronym. Fixed reference to NVMe Base Specification.
2022/02/15	Updated section 5.TBD of the NVM Command Set Specification change to be “i.e.,” rather than “i.e.”
2022/02/16	Moved section 1.8.1 in the NVMe Base Specification to section 1.9 as that title already existed. Corrected broken cross-references in black text.
2022/04/07	Correcting KPIODAAAG to be within MDTs.
2022/04/14	Added the reference in the KPIODAAAG field definition to section 5.TBD for the requirements of aborting commands that do not meet the KPIODAAAG alignment and granularity definition.
2022/07/26	Added a status code for an incorrect key.

Revision Date	Change Description
2022/08/17	Compared the 2/15/2022 version last integrated to the current revision to make sure the tracking of changes are completely identified.
2022/08/18	Updated the definition of the Invalid Key Tag and Incorrect Key status codes.
2022/08/25	Updated the Invalid Key Tag and Incorrect Key status codes.
2022-09-01	Removed comments and tracking for 30 day member review.
2022-09-05	CETPE defintion should not abort for a Fh vendor specific usage.
2022-10-06	Accepted all changes. Ready for Integration.
2022-11-27	Integrated
2022-11-29	Incorporated updates from Mike Allison
2022-12-01	Added a row of ... to account for bits defined by other TPs. Approved for ratification request.

Description for NVMe-Specification2.0a Changes:

Feature Enhancement:

Added the definition section of Key Per I/O

New requirement

- o Identify Controller data structure update to include indication of Key Per I/O support and scope.

References:

- o Technical Proposal 4055.

Description for NVMe-Command Set Specification 1.0a Changes:

Feature Enhancement:

New requirement

- o Identify Namespace data structure extended to include Key Per I/O Alignment granularity
- o Moved DSM to a directive to make space in Dword 13 for a 16 bit key tag field for Key Per I/O
- o Added CETYPE and CEV to I/O commands.

References:

- o Technical Proposal 4055.

Description for Zoned Namespace Command Set Specification 1.1a Changes:

Feature Enhancement:

New requirement

- o Added CETYPE and CEV to the Command Set specific I/O commands.

References:

- o Technical Proposal 4055.

Description for Key Value Command Set Specification 1.0a Changes:

Feature Enhancement:

New requirement

- o Added CETYPE and CEV to the Command Set specific I/O commands.

References:

- o Technical Proposal 4055.

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 2.0a as shown below:

Modify section 1.8 as shown below:

1 Introduction

...

1.8 References

...

TCG Storage Interactions Specification (SIIS). Available from <https://www.trustedcomputinggroup.org>.

Update section 1.9 as shown below:

1.9 References Under Development

TCG Storage Security Subsystem Class: Key Per I/O Specification

Modify a portion of section 3.3.3.2.1.1 as shown below:

3 NVM Express Architecture

...

3.3 NVM Queue Models

...

3.3.3 Queueing Data Structures

...

3.3.3.2 Common Completion Queue Entry

...

3.3.3.2.1 Status Field Definition

...

3.3.3.2.1.1 Generic Command Status Definition

...

Figure 94: Status Code - Generic Command Status Values

Value	Description	I/O Command Set Specific	I/O Command Set(s) ¹
...			
24h	<p>Admin Command Media Not Ready: The Admin command requires access to media and the media is not ready. The Do Not Retry bit indicates whether re-issuing the command at a later time may succeed. This status code shall only be returned:</p> <ul style="list-style-type: none"> a) for Admin commands; and b) if the controller is in Controller Ready Independent of Media mode (CC.CRIME is bit is set to '1'). <p>This status code shall not be returned with the Do Not Retry bit cleared to '0' after the amount of time indicated by the Controller Ready With Media Timeout field (CRTO.CRWMT) after the controller is enabled (i.e., CC.EN transitions from '0' to '1').</p> <p>Refer to Figure 103 for the list of Admin commands permitted to return this status code.</p>	No	
25h	<p>Invalid Key Tag: The command was aborted due to an invalid KEYTAG field value (refer to Figure NEW-6) as:</p> <ul style="list-style-type: none"> a) the value of the specified KEYTAG field is greater than the Maximum Key Tag (MAXKT) field in the I/O Command Set Independent Identify Namespace data structure (refer to Figure 280); or b) defined by the appropriate security specification (e.g., TCG Storage Interface Interactions specification). 	No	
...			
28h	<p>Incorrect Key: The command was aborted due to the key associated with the KEYTAG field being incorrect.</p> <p>The specific conditions under which a key is considered incorrect are defined by the appropriate security specification (e.g., TCG Storage Interface Interactions specification).</p>	No	
29h25h to 7Fh	Reserved		
80h	LBA Out of Range: See the applicable I/O Command Set specification for the description.	Yes	NVM, ZNS

Modify a portion of section 5.17.2.1 as shown below:

5 Admin Command Set

...

5.17 Identify command

...

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								
...												
357:356	O	O	O	Domain Identifier: This field indicates the identifier of the domain (refer to section 3.2.4) that contains this controller. If the MDS bit is set to '1', then this field shall be set to a non-zero value. If the NVM subsystem does not support multiple domains (i.e., the NVM subsystem consists of a single domain), then this field shall be cleared to 0h.								
358	O	P	P	Key Per I/O Capabilities: This field indicates the attributes for the Key Per I/O capability (refer to section 8.TBD).								
				<table><tr><th>Bits</th><th>Description</th></tr><tr><td>7:2</td><td>Reserved</td></tr><tr><td>1</td><td>Key Per I/O Scope (KPIOSC): If set to '1', then the Key Per I/O capability applies to all namespaces in the NVM subsystem when the Key Per I/O capability is enabled. If cleared to '0', then the Key Per I/O capability does not apply to all namespaces in the NVM subsystem and is allowed to be independently enabled and disabled uniquely on each namespace within the NVM subsystem. This bit should be ignored by the host if the KPIOS bit is cleared to '0'.</td></tr><tr><td>0</td><td>Key Per I/O Supported (KPIOS): If set to '1', then the controller supports the Key Per I/O capability. If cleared to '0', then the controller does not support the Key Per I/O capability.</td></tr></table>	Bits	Description	7:2	Reserved	1	Key Per I/O Scope (KPIOSC): If set to '1', then the Key Per I/O capability applies to all namespaces in the NVM subsystem when the Key Per I/O capability is enabled. If cleared to '0', then the Key Per I/O capability does not apply to all namespaces in the NVM subsystem and is allowed to be independently enabled and disabled uniquely on each namespace within the NVM subsystem. This bit should be ignored by the host if the KPIOS bit is cleared to '0'.	0	Key Per I/O Supported (KPIOS): If set to '1', then the controller supports the Key Per I/O capability. If cleared to '0', then the controller does not support the Key Per I/O capability.
				Bits	Description							
				7:2	Reserved							
1	Key Per I/O Scope (KPIOSC): If set to '1', then the Key Per I/O capability applies to all namespaces in the NVM subsystem when the Key Per I/O capability is enabled. If cleared to '0', then the Key Per I/O capability does not apply to all namespaces in the NVM subsystem and is allowed to be independently enabled and disabled uniquely on each namespace within the NVM subsystem. This bit should be ignored by the host if the KPIOS bit is cleared to '0'.											
0	Key Per I/O Supported (KPIOS): If set to '1', then the controller supports the Key Per I/O capability. If cleared to '0', then the controller does not support the Key Per I/O capability.											
367:358				Reserved								
359												
...												

Modify a portion of section 5.17.2.8 as shown below:

5.17.2.8 I/O Command Set Independent Identify Namespace data structure (CNS 08h)

...

Figure 280: Identify - I/O Command Set Independent Identify Namespace Data Structure

Bytes	O/M 1	Description	Capability Field <<TP4095a >>				
...							
14	M	Namespace Status (NSTAT): This field indicates the status of the namespace with the specified NSID. <table><tr><th>Bits</th><th>Description</th></tr><tr><td>7:1</td><td>Reserved</td></tr></table>	Bits	Description	7:1	Reserved	
Bits	Description						
7:1	Reserved						

**Figure 280: Identify - I/O Command Set Independent Identify
Namespace Data Structure**

Bytes	O/M 1	Description		Capability Field <<TP4095a >>								
		0	Namespace Ready (NRDY): A value of '1' indicates that the namespace is ready (refer to section 3.5.3). A value of '0' indicates that the namespace is not ready.									
15	O	Key Per I/O Status: This field indicates the namespace Key Per I/O capability status. <table><tr><th>Bit s</th><th>Description</th></tr><tr><td>7:2</td><td>Reserved</td></tr><tr><td>1</td><td>Key Per I/O Supported in Namespace (KPIOSNS): If set to '1', then the Key Per I/O capability is supported by the namespace. If cleared to '0', then the Key Per I/O capability is not supported by the namespace.</td></tr><tr><td>0</td><td>Key Per I/O Enabled in Namespace (KPIOENS): If set to '1', then the Key Per I/O capability is enabled on the namespace. The mechanism to enable the Key Per I/O capability on the namespace is outside the scope of this specification (refer to section 8.TBD). If cleared to '0', then the Key Per I/O capability is disabled on the namespace. If the KPIOSNS bit is cleared to '0', then this bit shall be cleared to '0'.</td></tr></table>		Bit s	Description	7:2	Reserved	1	Key Per I/O Supported in Namespace (KPIOSNS): If set to '1', then the Key Per I/O capability is supported by the namespace. If cleared to '0', then the Key Per I/O capability is not supported by the namespace.	0	Key Per I/O Enabled in Namespace (KPIOENS): If set to '1', then the Key Per I/O capability is enabled on the namespace. The mechanism to enable the Key Per I/O capability on the namespace is outside the scope of this specification (refer to section 8.TBD). If cleared to '0', then the Key Per I/O capability is disabled on the namespace. If the KPIOSNS bit is cleared to '0', then this bit shall be cleared to '0'.	Yes
Bit s	Description											
7:2	Reserved											
1	Key Per I/O Supported in Namespace (KPIOSNS): If set to '1', then the Key Per I/O capability is supported by the namespace. If cleared to '0', then the Key Per I/O capability is not supported by the namespace.											
0	Key Per I/O Enabled in Namespace (KPIOENS): If set to '1', then the Key Per I/O capability is enabled on the namespace. The mechanism to enable the Key Per I/O capability on the namespace is outside the scope of this specification (refer to section 8.TBD). If cleared to '0', then the Key Per I/O capability is disabled on the namespace. If the KPIOSNS bit is cleared to '0', then this bit shall be cleared to '0'.											
17:16 <2 byte aligned>	O	Maximum Key Tag (MAXKT): This field indicates the maximum Key Tag field value allocated to this namespace in an I/O command that supports a Key Tag. If the KPIOENS bit is set to '1', then the valid Key Tag field values for this namespace are 0h to the value of this field. If the KPIOENS bit is cleared to '0', then this field is reserved.		No								
4095:1815		Reserved										
NOTES:												
1. O/M definition: O = Optional, M = Mandatory.												

Modify portions of section 8.21 as shown below:

8 Extended Capabilities

...

8.21 Sanitize Operations

...

This additional media modification shall complete before the NVM subsystem:

- a) reports sanitize completion by Asynchronous Event (refer to section 5.2); and
- b) reports sanitize completion in the Sanitize Status log (refer to section 5.16.1.25).

The Overwrite sanitize operation is media specific and may not be appropriate for all media types. For example, if the media is NAND, multiple pass overwrite operations may have an adverse effect on media endurance.

A sanitize operation alters all user data in the NVM subsystem such that recovery of any previous user data from any cache, the non-volatile media, or any Controller Memory Buffer is not possible. It is implementation specific whether Submission Queues and Completion Queues within a Controller Memory Buffer are altered by a sanitize operation; all other data stored in all Controller Memory Buffers is altered by a sanitize operation. If a portion of the user data was not altered and the sanitize operation completed successfully, then the NVM subsystem shall ensure permanent inaccessibility of that portion of the user data for any future use within the NVM subsystem (e.g., retrieval from NVM media, caches, or any Controller Memory Buffer) and permanent inaccessibility of that portion of the user data via any interface to the NVM subsystem, including management interfaces as defined by the NVMe Management Interface Specification.

If the NVM subsystem supports the Key Per I/O capability (refer to section 8.TBD), then a sanitize operation shall alter all user data such that recovery of any previous user data using the KEYTAG values specified when that previous user data was written (i.e., original KEYTAG values) is infeasible for a given level of effort (refer to ISO/IEC 27040).

...

Add the following section:

8.TBD Key Per I/O

The Key Per I/O capability provides a mechanism to use encryption keys that have been injected into an NVM subsystem by a host. The mechanism to perform activation of the Key Per I/O capability, encryption key injection, management, and association to encryption key tag is outside the scope of this specification. One mechanism is defined in the TCG Storage Security Subsystem Class: Key Per I/O Specification using the NVMe Admin commands Security Send and Security Receive. If the TCG mechanism is used, any additional modifications to the NVM subsystem as a result of activation of the Key Per I/O Security Provider are defined in the TCG Storage Interface Interactions Specification (SIIS).

Encryption keys injected into the NVM subsystem may be referenced by I/O commands through the use of the encryption key tag (refer to the KEYTAG field in Figure NEW-6) associated with the encryption keys. If an I/O command CETYPE field is set to the KPIOTAG value, then the CEV field (refer to Figure NEW-6) of that I/O command specifies the encryption key tag associated with the encryption key to be used to encrypt or decrypt the data in that I/O command. The association of an encryption key tag to a specific encryption key is outside the scope of this specification. One association mechanism is defined in the TCG Storage Security Subsystem Class: Key Per I/O Specification.

The Key Per I/O Scope bit (refer to Figure 275) indicates if the Key Per I/O capability:

- applies to all read and write commands in all namespaces within the NVM subsystem; or
- independently applies to read and write commands in each namespace within the NVM subsystem.

The Key Per I/O capability does not have any effect on host accesses to RPMBs and Boot Partitions as these features are not addressed through I/O commands that specify a namespace.

A controller that supports the Key Per I/O capability shall set the KPIOS bit to '1' in the Identify Controller data structure (refer to Figure 275).

A namespace that supports the Key Per I/O capability shall set the KPIOSNS bit to '1' in the I/O Command Set Independent Identify Namespace data structure (refer to Figure 280).

The Key Per I/O capability uses the Command Extension Type (CETYPE) and Command Extension Value (CEV) fields in all read and write commands. Definition of the CETYPE fields are shown in [Figure NEW-6](#).

Figure NEW-6: CETYPE Definition

Value	Description	CEV Field Definition
0h	Reserved.	
1h	Key Per I/O Tag (KPIOTAG): This command is using the Key Per I/O capability.	<p>Key Tag (KEYTAG): Specifies a namespace-specific 16-bit encryption key tag that identifies the encryption key used to encrypt or decrypt the data of the command.</p> <p>The same Key Tag value on different namespaces may or may not identify the same encryption key.</p> <p>Refer to the Maximum Key Tag field in the I/O Command Set Independent Identify Namespace data structure (refer to Figure 280) for the supported values.</p>
2h to Eh	Reserved	
Fh	Vendor Specific	

If:

the Key Per I/O capability is enabled in a namespace (i.e., the KPIOENS bit set to '1');
an I/O command supports the CETYPE field; and
the CETYPE field in that I/O command is set to a value that is reserved,

then the controller shall abort that command with a status code of Invalid Field in Command.

...

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

Modify a portions of section 3.2.1 as shown below:

3 I/O Commands for the NVM Command Set

...

3.2 NVM Command Set Commands

...

3.2.1 Compare command

...

The command uses Command Dword 2, Command Dword 3, Command Dword 10, Command Dword 11, Command Dword 12, [Command Dword 13](#), Command Dword 14, and Command Dword 15 fields. If the command uses PRPs for the data transfer, then the Metadata Pointer, PRP Entry 1, and PRP Entry 2 fields are used. If the command uses SGLs for the data transfer, then the Metadata SGL Segment Pointer and SGL Entry 1 fields are used. All other command specific fields are reserved.

...

Figure 23: Compare - Command Dword 12

Bits	Description
...	
25:16	Reserved
19:16	Command Extension Type (CETYPE) : Specifies the Command Extension Type that applies to the command (refer to the Key Per I/O section in the NVM Express Base Specification).

...

The definition of [Command Dword 13](#) is based on the CETYPE value. If the CETYPE value is cleared to 0h, then [Command Dword 13](#) is reserved. If the CETYPE value is non-zero, then [Command Dword 13](#) is defined in [Figure 23_TBD](#).

...

Figure 23_TBD: Compare - Command Dword 13 if CETYPE is non-zero

Bits	Description
31:16	Reserved
15:00	Command Extension Value (CEV) : The definition of this field is dependent on the value of the CETYPE field. Refer to the Key Per I/O section in the NVM Express Base Specification.

...

Modify portions of section 3.2.2 as shown below:

3.2.2 Copy command

...

Figure 30: Copy - Command Dword 12

Bits	Description
...	
19:16	Command Extension Type (CETYPE): Specifies the Command Extension Type that applies to the command (refer to the Key Per I/O section in the NVM Express Base Specification). This field is used for the write portion of the copy operation.
...	

Figure 31: Copy - Command Dword 13

Bits	Description
....	
15:00	Reserved Command Extension Value (CEV): The definition of this field is dependent on the value of the CETYPE field. Refer to the Key Per I/O section in the NVM Express Base Specification. This field is used for the write portion of the copy operation.

...

Figure 34: Copy - Source Range Entries Descriptor Format 0h

Range	Bytes	Description									
Source Range 0	07:00	Reserved									
	15:08	Starting LBA									
	19:16	Read Parameters as follows:									
		<table><tr><th>Bits</th><th>Description</th></tr><tr><td>31:20</td><td>Reserved</td></tr><tr><td>19:16</td><td>Command Extension Type (CETYPE): Specifies the Command Extension Type that applies to the command (refer to the Key Per I/O section in the NVM Express Base Specification). This field is used for the read portion of the copy operation for the LBAs specified in this Source Range entry.</td></tr><tr><td>15:00</td><td>Number of Logical Blocks (NLB): This field indicates the number of logical blocks to be copied. This is a 0's based value.</td></tr></table>		Bits	Description	31:20	Reserved	19:16	Command Extension Type (CETYPE): Specifies the Command Extension Type that applies to the command (refer to the Key Per I/O section in the NVM Express Base Specification). This field is used for the read portion of the copy operation for the LBAs specified in this Source Range entry.	15:00	Number of Logical Blocks (NLB): This field indicates the number of logical blocks to be copied. This is a 0's based value.
		Bits	Description								
		31:20	Reserved								
	19:16	Command Extension Type (CETYPE): Specifies the Command Extension Type that applies to the command (refer to the Key Per I/O section in the NVM Express Base Specification). This field is used for the read portion of the copy operation for the LBAs specified in this Source Range entry.									
	15:00	Number of Logical Blocks (NLB): This field indicates the number of logical blocks to be copied. This is a 0's based value.									
21:20	Command Extension Value (CEV): The definition of this field is dependent on the value of the CETYPE field. Refer to the Key Per I/O section in the NVM Express Base Specification. This field is used for the read portion of the copy operation for the LBAs specified in this Source Range entry.										
23:22	Reserved										
....											
Source Range 1	39:32	Reserved									
	47:40	Starting LBA									
	51:48	Read Parameters									
	53:52	CEV									
	55:54	Reserved									
	59:56	EILBRT									
	61:60	ELBAT									
63:62	ELBATM										
...											
Source Range 127	4071:4064	Reserved									
	4079:4072	Starting LBA									
	4083:4080	Read Parameters									
	4085:4084	CEV									
	4087:4086	Reserved									
	4091:4088	EILBRT									
	4093:4092	ELBAT									
	4095:4094	ELBATM									

...

Figure 35: Copy - Source Range Entries Descriptor Format 1h

Range	Bytes	Description	
Source Range 0	07:00	Reserved	
	15:08	Starting LBA	
	19:16	Read Parameters as follows:	
		Bits	Description

Figure 35: Copy - Source Range Entries Descriptor Format 1h

Range	Bytes	Description
	31:20	Reserved
	19:16	Command Extension Type (CETYPE): Specifies the Command Extension Type that applies to the command (refer to the Key Per I/O section in the NVM Express Base Specification). This field is used for the read portion of the copy operation for the LBAs specified in this Source Range entry.
	15:00	Number of Logical Blocks (NLB): This field indicates the number of logical blocks to be copied. This is a 0's based value.
	21:20	Command Extension Value (CEV): The definition of this field is dependent on the value of the CETYPE field. Refer to the Key Per I/O section in the NVM Express Base Specification. This field is used for the read portion of the copy operation for the LBAs specified in this Source Range entry.
	25:22	Reserved
	...	
Source Range 1	47:40	Reserved
	55:48	Starting LBA
	59:56	Read Parameters
	61:60	CEV
	65:62	Reserved
	75:66	The variable sized ELBST and EILBRT
	77:76	ELBAT
	79:78	ELBATM
	...	
Source Range 101	4047:4040	Reserved
	4055:4048	Starting LBA
	4059:4056	Read Parameters
	4061:4060	CEV
	4065:4062	Reserved
	4075:4066	The variable sized ELBST and EILBRT
	4077:4076	ELBAT
	4079:4078	ELBATM
	4095:4080	Reserved

Modify a portions of section 3.2.4 as shown below:

3.2.4 Read command

...

Figure 48: Read - Command Dword 12

Bits	Description
...	
25:1620	Reserved

Figure 48: Read - Command Dword 12

Bits	Description0
19:16	Command Extension Type (CETYPE): Specifies the Command Extension Type that applies to the command (refer to the Key Per I/O section in the NVM Express Base Specification).
...	

The definition of Command Dword 13 is based on the CETYPE value. If the CETYPE value is cleared to 0h, then Command Dword 13 is defined in Figure 49. If the CETYPE value is non-zero, then Command Dword 13 is defined in Figure 49_TBD.

Figure 49: Read - Command Dword 13 if CETYPE is cleared to 0h

Bits	Description
31:08	Reserved
...	

Figure 49_TBD: Read - Command Dword 13 if CETYPE is non-zero

Bits	Description
31:16	Reserved
15:00	Command Extension Value (CEV): The definition of this field is dependent on the value of the CETYPE field. Refer to the Key Per I/O section in the NVM Express Base Specification.

Modify portions of section 3.2.5 as shown below:

3.2.5 Verify command

...

The command uses Command Dword 10, Command Dword 11, Command Dword 12, **Command Dword 13**, Command Dword 14, and Command Dword 15 fields.

...

Figure 55: Verify - Command Dword 12

Bits	Description
....	
25:20	Reserved
19:16	Command Extension Type (CETYPE): Specifies the Command Extension Type that applies to the command (refer to the Key Per I/O section in the NVM Express Base Specification).
...	

The definition of Command Dword 13 is based on the CETYPE value. If the CETYPE value is cleared to 0h, then Command Dword 13 is reserved. If the CETYPE value is non-zero, then Command Dword 13 is defined in Figure NEW-3.

Figure NEW-3: Verify - Command Dword 13 if CETYPE is non-zero

Bits	Description
31:16	Reserved
15:00	Command Extension Value (CEV): The definition of this field is dependent on the value of the CETYPE field. Refer to the Key Per I/O section in the NVM Express Base Specification.

Modify portions of section 3.2.6 as shown below:

3.2.6 Write command

...

Figure 63: Write - Command Dword 12

Bits	Description
...	
19:16	Command Extension Type (CETYPE): Specifies the Command Extension Type that applies to the command (refer to the Key Per I/O section in the NVM Express Base Specification).
...	

...

The definition of Command Dword 13 is based on the CETYPE value. If the CETYPE value is cleared to 0h, then Command Dword 13 is defined in Figure 64. If the CETYPE value is non-zero, then Command Dword 13 is defined in Figure 64_TBD.

Figure 64: Write - Command Dword 13 if CETYPE is cleared to 0h

Bits	Description
31:16	Directive Specific (DSPEC): Specifies the Directive Specific value associated with the Directive Type field (refer to section 9.1).
15:08	Reserved
07:00	Dataset Management (DSM): This field indicates attributes for the LBA(s) being written. ...

Figure 64_TBD: Write - Command Dword 13 if CETYPE is non-zero

Bits	Description
31:16	Directive Specific (DSPEC): Specifies the Directive Specific value associated with the Directive Type field (refer to the Key Per I/O section in the NVM Express Base Specification).
15:00	Command Extension Value (CEV): The definition of this field is dependent on the value of the CETYPE field. Refer to the Key Per I/O section in the NVM Express Base Specification.

Modify the text in section 3.2.8 as shown below:

3.2.8 Write Zeroes command

...

The fields used are Command Dword 10, Command Dword 11, Command Dword 12, Command Dword 13, Command Dword 14, and Command Dword 15 fields.

Figure 73: Write Zeroes - Command Dword 12

Bits	Description
...	
24:20	Reserved
19:16	Command Extension Type (CETYPE) : Specifies the Command Extension Type that applies to the command (refer to the Key Per I/O section in the NVM Express Base Specification).
...	

The definition of Command Dword 13 is based on the CETYPE value. If the CETYPE value is cleared to 0h, then Command Dword 13 is reserved. If the CETYPE value is non-zero, then Command Dword 13 is defined in Figure **NEW-5**.

Figure NEW-5: Write Zeros - Command Dword 13 if CETYPE is non-zero

Bits	Description
31:16	Reserved.
15:00	Command Extension Value (CEV) : The definition of this field is dependent on the value of the CETYPE field. Refer to the Key Per I/O section in the NVM Express Base Specification.

Modify portions of Figure 97 in section 4.1.5.1as shown below:

4 Admin Commands for the NVM Command Set

...

4.1 Admin Command behavior for the NVM Command Set

...

4.1.5 Identify Command

...

4.1.5.1 NVM Command Set Identify Namespace data structure (CNS 00h)

...

Figure 97: Identify - Identify Namespace Data Structure, NVM Command Set

Bytes	O/M 1	Description	Capability Field <<TP 4095>>
...			
80	O	Maximum Source Range Count (MSRC) : This field indicates the maximum number of Source Range entries that may be used to specify source data in a Copy command. This is a 0's based value.	
...			
87:84 <all ready integrated value>	O	Key Per I/O Data Access Alignment and Granularity (KPIODAAG) : This field indicates the alignment and granularity in logical blocks that is required for commands that support a KPIOTAG value in the CETYPE field (refer to the Key Per I/O section in the NVM Express Base Specification).	No

Figure 97: Identify - Identify Namespace Data Structure, NVM Command Set

Bytes	O/M 1	Description	Capability Field <<TP 4095>>
		<p>This is a 0's based value.</p> <p>Refer to section 5.TBD on the behavior of commands not meeting the alignment or granularity defined by this field.</p> <p>The value of this field may change if the namespace is reformatted.</p> <p>If the KPIOSNS bit is cleared to '0' in the I/O Command Set Independent Identify Namespace data structure (refer to the NVM Express Base Specification), then this field is reserved.</p>	
91: 88 81		Reserved	
95:92	0	ANA Group Identifier (ANAGRPID): This field is as defined in the I/O Command Set Independent Identify Namespace data structure (refer to the I/O Command Set Independent Identify Namespace data structure section in the NVMe Base Specification).	
98:96		Reserved	
...			

Add the following section:

5 Extended Capabilities

...

5.TBD Key Per I/O

The Key Per I/O capability operates as defined by the NVM Express Base Specification with the exceptions specified by this section.

Any I/O command specifying the key tag (i.e., CETYPE value of KPIOTAG) that does not have:

- the SLBA field in that command aligned on a logical block boundary specified by the KPIODAAG field in the Identify Namespace data structure (refer to Figure 97); and
- the range of LBAs accessed by that command specify a granularity specified by the KPIODAAG field,

then the controller shall abort that command with a status code of Invalid Field in Command.

Modify portions of Zoned Namespace Command Set Specification 1.1a as shown below:

Update the Zone Append command to use the new key per I/O fields in section 3.4.1 as shown below:

3 I/O Commands for the Zoned Namespace Command Set

...

3.4 Zoned Namespace Command Set I/O Commands

...

3.4.1 Zone Append command

...

Figure 26: Zone Append - Command Dword 12

Bits	Description
...	
24:20 16	Reserved
19:16	Command Extension Type (CETYPE) : Specifies the Command Extension Type that applies to the command (refer to the Key Per I/O section in the NVM Express Base Specification).
...	

The definition of Command Dword 13 is based on the CETYPE value. If the CETYPE value is cleared to 0h, then Command Dword 13 is defined in Figure 27. If the CETYPE value is non-zero, then Command Dword 13 is defined in Figure **NEW-APPEND**.

Figure 27: Zone Append - Command Dword 13 if CETYPE is cleared to 0h

Bits	Description
31:16	Directive Specific (DSPEC) : Specifies the Directive Specific value associated with the Directive Type field (refer to the NVMe Base Specification).
15:00	Reserved

Figure NEW-APPEND: Zone Append - Command Dword 13 if CETYPE is non-zero

Bits	Description
31:16	Directive Specific (DSPEC) : Specifies the Directive Specific value associated with the Directive Type field (refer to the Key Per I/O section in the NVM Express Base Specification).
15:00	Command Extension Value (CEV) : The definition of this field is dependent on the value of the CETYPE field. Refer to the Key Per I/O section in the NVM Express Base Specification.

Modify portions of Key Value Command Set Specification 1.0a as shown below:

Update the Retrieve command to use the new Key Per I/O fields:

3 I/O Commands for the Key Value Command Set

...

3.2 Key Value Command Set Commands

...

3.2.3 Retrieve command

...

The command uses Command Dword 2, Command Dword 3, Command Dword 10, Command Dword 11, [Command Dword 13](#), Command Dword 14, and Command Dword 15 fields. All other command specific fields are reserved. If the command uses PRPs for the data transfer, then the PRP Entry 1, and PRP Entry 2 fields are used. If the command uses SGLs for the data transfer, then the SGL Entry 1 field is used.

...

FigureTBDA: Retrieve - Command Dword 13

Bit	Description
31:20	Reserved
19:16	Command Extension Type (CETYPE): Specifies the Command Extension Type that applies to the command (refer to the Key Per I/O section in the NVM Express Base Specification).
15:00	Command Extension Value (CEV): The definition of this field is dependent on the value of the CETYPE field. Refer to the Key Per I/O section in the NVM Express Base Specification.

Update the Store command to use the new Key Per I/O fields:

3.2.5 Store command

The Store command stores a value to the NVM KV controller for the KV key specified.

The command uses Command Dword 2, Command Dword 3, Command Dword 10, Command Dword 11, [Command Dword 13](#), Command Dword 14, and Command Dword 15 fields. If the command uses PRPs for the data transfer, then the PRP Entry 1, and PRP Entry 2 fields are used. If the command uses SGLs for the data transfer, then the SGL Entry 1 field is used.

...

FigureTBDB: Store - Command Dword 13

Bit	Description
31:20	Reserved
19:16	Command Extension Type (CETYPE) : Specifies the Command Extension Type that applies to the command (refer to the Key Per I/O section in the NVM Express Base Specification).
15:00	Command Extension Value (CEV) : The definition of this field is dependent on the value of the CETYPE field. Refer to the Key Per I/O section in the NVM Express Base Specification.