NVM Express®

Key Value Command Set Specification

Revision 1.0b

January 6th, 2022

Please send comments to info@nvmexpress.org

NVM Express® Key Value Command Set specification revision 1.0b is available for download at http://nvmexpress.org. The NVM Express® Key Value Command Set specification revision 1.0a incorporates NVM Express® Key Value Command Set specification, revision 1.0 (refer to the Key Value Command Set Specification change list https://nvmexpress.org/changes-in-nvm-express-revision-2-0 for details), ECN 001, and ECN102.

SPECIFICATION DISCLAIMER

LEGAL NOTICE:

This NVM Express Key Value Command Set revision 1.0b 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 Key Value Command Set revision 1.0b 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: **"© 2007-2022 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.

The NVM Express® design mark is a registered trademark of NVM Express, Inc.
PCIe® is a registered trademark of PCI-SIG.

NVM Express Workgroup
c/o VTM, Inc.
3855 SW 153rd Drive
Beaverton, OR 97003 USA
info@nvmexpress.org

Table of Contents

Table of Figures

1 Introduction

1.1 Overview

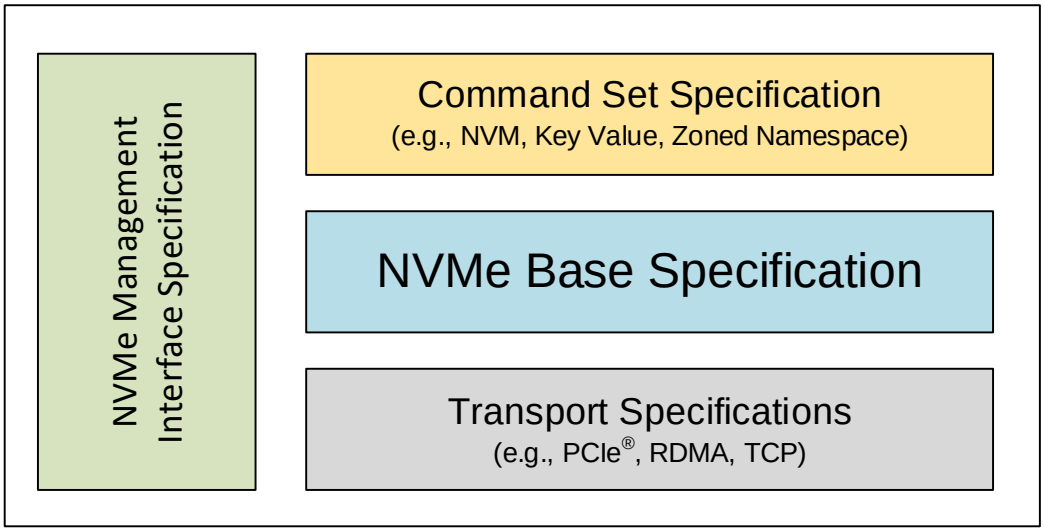
NVM Express® (NVMe®) Base Specification defines an interface for host software to communicate with non-volatile memory subsystems over a variety of memory based transports and message based transports.

This document defines a specific NVMe I/O Command Set, the Key Value Command Set, which extends the NVMe Base Specification.

1.2 Scope

Figure 1 shows the relationship of the NVM Express® Key Value Command Set Specification to other specifications within the NVMe Family of Specifications.

Figure 1: NVMe Family of Specifications



This specification supplements the NVMe Base Specification. This specification defines additional Data Structures, Features, log pages, commands, and status values. This specification also defines extensions to existing data structures, features, log pages, commands, and status values. This specification defines requirements and behaviors that are specific to the Key Value Command Set. Functionality that is applicable generally to NVMe or that is applicable across multiple I/O Command Sets is defined in the NVMe Base Specification.

If a conflict arises among requirements defined in different specifications, then a lower-numbered specification in the following list shall take precedence over a higher-numbered specification:

1. Non-NVMe specifications
2. NVMe Base Specification
3. NVMe transport specifications
4. NVMe I/O command set specifications
5. NVMe-MI specification

1.3 Conventions

This specification conforms to the Conventions section, Keywords section, and Byte, Word, and Dword Relationships section of the NVMe Base Specification.

1.4 Definitions

1.4.1 Definitions from the NVMe Base Specifications

This specification uses the definitions in the NVMe Base Specification.

1.4.2 Definitions in the NVMe Base Specification specified in the Key Value Command Set

The following terms used in this specification and the NVMe Base Specification are as defined here.

1.4.2.1 Endurance Group Host Read Command

A Retrieve command

1.4.2.2 Format Index

A value used to index into the KV Format data structure.

1.4.2.3 SMART Data Units Read Command

A Retrieve command

1.4.2.4 SMART Host Read Command

A Retrieve command.

1.4.2.5 User Data Format

The layout of the data on the NVM media as described by the Key Value Format of the namespace.

1.4.2.6 User Data Out Command

A Store command

1.4.3 Definitions specific to the Key Value Command Set

This section defines terms that are specific to this specification.

1.4.3.1 key value pair

An associated KV key and KV value that may be stored on media where the KV key identifies the associated KV value.

1.4.3.2 KV key

The part of a key value pair that is used to identify that key value pair.

1.4.3.3 KV value

The value that is associated with a key value pair.

1.5 References

NVMe Base Specification, revision 2.0. Available from http://www.nvmexpress.org.

2 Key Value Command Set Model

The NVMe Base Specification defines a register level interface for host software to communicate with a non-volatile memory subsystem. This specification defines additional functionality for the Key Value Command Set.

Each I/O Command Set is assigned a specific Command Set Identifier (CSI) value by the NVMe Base Specification. The Key Value Command Set is assigned a CSI value of 01h.

2.1 Theory of operation

An NVM subsystem may contain controllers that implement the Key Value Command Set. Key Value storage is measured in bytes. The amount of storage required to store a key value pair is the sum of the KV key size and the KV value size. A KV value is allowed to have a length of zero bytes, Supported KV key and KV value sizes are reported in the I/O Command Set specific Identify Namespace data structure for the Key Value Command Set.

A device that implements the Key Value Command Set provides access to data identified by a KV key. The KV key may be variable length and the length of the KV key is specified in the command. Two KV keys that have different lengths are not the same. The KV value that is associated with a KV key has a length in bytes that is specified in the command that stores that KV value. The length in bytes of a KV value is indicated in the response to a query about that KV value (e.g., Retrieve command, Exist command). The length in bytes of a KV key is indicated in the response to a List command that returns that KV key.

While a controller may perform operations (e.g., compression) on data before the data is stored on the media and perform the reverse of that operation (e.g., decompression) when retrieving the data from the media, this functionality is outside of the scope of this specification.

The maximum size of any KV key and the maximum size of any KV value in a namespace is specified when the namespace is formatted and is selected from the matrix of KV formats in the I/O Command Set specific Identify Namespace data structure.

2.1.1 Namespaces

A namespace is a collection of NVM and is as defined in the NVMe Base Specification.

The number of bytes required to store a key value pair is related to the KV key size and the KV value size. Supported KV key sizes and KV value sizes are reported in the KV Format data structures in the Identify Namespace data structure.

The number of bytes required to store a given key value pair is greater than or equal to the sum of the size of the KV key and the size of the KV value. Namespace Size and Namespace Utilization reflect the number of bytes required to store the KV value and KV key.

The Key Value Command Set specific Identify Namespace data structure (refer to 4.1.5.1) contains related fields reporting the Namespace Size, Capacity and Utilization:

- The Namespace Size (NSZE) field defines the total size of the namespace in bytes.
- The Namespace Utilization (NUSE) field defines the number of bytes currently allocated in the namespace.

The Namespace Utilization (NUSE) field shall be less than or equal to the Namespace Size (NSZE) field.

2.1.2 Command Ordering Requirements

Each command is processed as an independent entity without reference to other commands submitted to the same I/O Submission Queue or to commands submitted to other I/O Submission Queues. Specifically, the controller is not responsible for checking the KV key of a Retrieve or Store command to ensure any type of ordering between commands. For example, if a Retrieve command is submitted for KV key *x* and there is a Store command also submitted for KV key *x*, then there is no guarantee of the order of completion for those commands (the Retrieve command may finish first or the Store command may finish first). If there are ordering requirements between these commands, the host enforces those requirements above the level of the controller.

2.1.3 Fused Operation

The Key Value Command Set does not support any Fused Operations.

2.1.4 Atomic Operation

All Store commands and Delete commands are atomic with respect to the associated key value pair.

2.1.5 Key Size implications

The maximum KV key size is 16 bytes.

The KV key is specified in Command Dword 2, Command Dword 3, Command Dword 14, and Command Dword 15.

If a command specifies a KV key size greater than 16 bytes, that command is aborted with a status code of Invalid Field in Command.

2.1.6 Command Operation

2.1.6.1 Delete command

The Delete command requests the controller to delete the specified key value pair from the namespace.

2.1.6.2 List command

The host may request a list of the KV keys in that namespace. This is accomplished using the List command. The KV keys in the data structure returned from the List command are not in any specified order, but in the absence of Sanitize, Format NVM, Store, and Delete commands the order of the KV keys in the list shall be constant. The KV key that is sent in that command specifies the starting point in the list of KV keys. If that KV key exists, then that KV key is the first key returned in the data structure. If that KV key does not exist, then the device returns KV keys where the first KV key returned is vendor specific, but in the absence of Sanitize, Format NVM, Store, and Delete commands the first KV key returned shall not change.

2.1.6.3 Exist command

The Exist command is used to determine if a specified KV key exists in the namespace. The existence of the KV key is indicated by the value returned in the CQE for that command.

2.1.6.4 Store command

The Store command is used to store a key value pair to the namespace. The length of the KV value is specified in the Store command and the location of the KV value to be stored is specified by either the SGL or the PRP in the command. The Store command is an atomic command (e.g., following a Store command

of a key value pair that existed prior to that command, a Retrieve command returns either all of the previous KV value or all of the KV value in that Store command but shall not return a combination of previous data and data from that Store command).

2.1.6.5 Retrieve command

The Retrieve command is used to retrieve a key value pair from the namespace. The length to be retrieved of the KV value is specified in the Retrieve command and the location to transfer the KV value to is specified by either the SGL or the PRP in the command. If the length specified in the command is less than the length of the KV value that is being retrieved, then the device returns the requested portion of the KV value and the full length of the KV value is returned in the CQE. If the length specified in the command is greater than the length of the KV value that is being retrieved, then the device returns the data from the media and the length of that KV value is returned in the CQE.

2.2 I/O Controller Requirements

2.2.1 Command Support

This specification implements the command support requirements for I/O Controllers defined in the NVMe Base Specification. Figure 2 defines Key Value Command Set specific definitions for I/O commands that are mandatory, optional, and prohibited for an I/O controller that supports the Key Value Command Set.

Figure 2: I/O Controller – Key Value Command Set Support

Command	Command Support Requirements [1]
Store	M
Retrieve	M
Delete	M
Exist	M
List	M
Notes: 1. O = Optional, M = Mandatory, P = Prohibited	

2.2.2 Log Page Support

This specification implements the log page support requirements for I/O Controllers defined in the NVMe Base Specification. There are no additional Key Value Command Set specific definitions for log pages that are mandatory, optional, and prohibited for an I/O controller that supports the Key Value Command Set.

2.2.3 Features Support

This specification implements the feature support requirements for I/O Controllers defined in the NVMe Base Specification. Figure 3 defines Key Value Command Set specific definitions for features that are mandatory, optional, prohibited, and not recommended for an I/O Controller that supports the Key Value Command Set.

Figure 3: I/O Controller – Feature Support

Feature Name	Feature Support Requirements [1]	Logged in Persistent Event Log
Key Value Configuration	M	Yes
Notes: 1. O = Optional, M = Mandatory, P = Prohibited, NR = Not Recommended		

3 I/O Commands for the Key Value Command Set

This section specifies the Key Value Command Set I/O Commands.

3.1 Submission Queue Entry and Completion Queue Entry

The Submission Queue Entry (SQE) structure and the fields that are common to all NVMe I/O Command Sets are defined in the Submission Queue Entry – Command Format section in the NVMe Base Specification. The Completion Queue Entry (CQE) structure and the fields that are common to all NVMe I/O Command Sets are defined in the Completion Queue Entry section in the NVMe Base Specification.

The Key Value Command Set uses the Common Command Format as defined in the NVMe Base Specification.

Command Dword 0, Namespace Identifier, Metadata Pointer, PRP Entry 1, PRP Entry 2, SGL Entry 1, and Metadata SGL Segment Pointer have common definitions for all Admin commands and I/O commands and are described in the Submission Queue Entry – Command Format section in the NVMe Base Specification.

The command specific fields in the SQE structure (i.e., SQE Command Dword2, Command Dword 3, Command Dwords 10-15) and the CQE structure (i.e., CQE Dword 0, and Dword 1) for the Key Value Command Set are defined in this section.

3.1.1 Common Command Format

The Common Command Format is as defined in the NVMe Base Specification.

SQE Command Dword 2 and Command Dword 3 contain KV key bytes [7:0]. SQE Command Dword 14 and Command Dword 15 contain KV key [15:8].

3.1.2 Key Value Command Set Specific Status Values

This specification supports the Command Specific status values defined in the NVMe Base Specification. Command Specific status values that are specific to the Key Value Command Set specification are defined in this section. Figure 4 defines the status values specific to the Key Value Command Set.

Figure 4: Status Code – Command Specific Status Values, Key Value Command Set

Value	Description	Commands Affected
81h	Capacity Exceeded	Store
82h	Namespace Not Ready	Delete, Exist, Retrieve, Store
83h	Reservation Conflict	Delete, Store, Retrieve
84h	Format In Progress	Delete, Exist, List, Retrieve, Store
85h	Invalid Value Size	Store
86h	Invalid Key Size	List, Retrieve, Store
87h	KV Key Does Not Exist	Delete, Exist, Retrieve, Store
88h	Unrecovered Error	Retrieve
89h	Key Exists	Store

3.2 Key Value Command Set Commands

The Key Value Command Set includes the commands listed in Figure 5. Section 3.2 describes the definition for each of the commands defined by this specification. Commands are submitted as described in the NVMe Base Specification.

Figure 5: Opcodes for Key Value Command Set Commands

Opcode by Field			Combined Opcode[1]	Command[2]	Reference
(07)	(06:02)	(01:00)			
Standard Command	Function	Data Transfer[3]			
Refer to the NVMe Base Specification				Flush	NVMe Base Specification
Refer to the NVMe Base Specification				Reservation Register	NVMe Base Specification
Refer to the NVMe Base Specification				Reservation Report	NVMe Base Specification
Refer to the NVMe Base Specification				Reservation Acquire	NVMe Base Specification
Refer to the NVMe Base Specification				Reservation Release	NVMe Base Specification
0b	000 00b	01b	01h	Store	3.2.5
0b	000 00b	10b	02h	Retrieve	3.2.3
0b	001 00b	00b	10h	Delete	3.2.1
0b	001 01b	00b	14h	Exist	3.2.4
0b	000 01b	10b	06h	List	3.2.2

NOTES:
1. Opcodes not listed are defined in the NVMe Base Specification.
2. All Key Value Command Set Commands use the Namespace Identifier (NSID) field. The value FFFFFFFFh is not supported in this field unless footnote 4 in this figure indicates that a specific command does support that value.
3. 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.
4. This command may support the use of the Namespace Identifier (NSID) field set to FFFFFFFFh.

3.2.1 Delete command

The Delete command deletes the KV key and the associated KV value for the specified KV namespace.

The command uses Command Dword 2, Command Dword 3, Command Dword 11, Command Dword 14, and Command Dword 15 fields. All other command specific fields are reserved.

If the value in the Key Length field is greater than 16, then the controller shall abort the command with Invalid Field in Command.

Figure 6: Delete – Command Dword 11

Bit	Description
31:8	**Reserved**
7:0	**Key Length (KL):** Specifies the length of the KV key in bytes.

Figure 7: Delete – Command Dword 2 and Command Dword 3

Bit	Description
63:0	**KV key[63:00]:** This field specifies the least-significant 64-bits of the KV key to be used for the command. Command Dword 2 contains bits 31:00; Command Dword 3 contains bits 63:32.

Figure 8: Delete – Command Dword 14 and Command Dword 15

Bit	Description
63:0	**KV key[127:64]:** This field specifies the most-significant 64-bits of the KV key to be used for the command. Command Dword 14 contains bits 95:64; Command Dword 15 contains bits 127: 96.

3.2.1.1 Command Completion

Upon completion of the Delete command, the controller posts a completion queue entry (CQE) to the associated I/O Completion Queue. If the status code returned is 00h, then the KV key and its associated KV value have been deleted.

Delete command specific status values are defined in Figure 9

Figure 9: Delete – Command Specific Status Values

Value	Description
87h	KV Key Does Not Exist**:** The KV key does not exist
0Bh	**Invalid Namespace or Format:** The namespace or the format of that namespace is invalid or the namespace is not associated with the KV Command Set..

3.2.2 List command

The List command retrieves a list of KV keys that exist for the specified KV namespace starting at the KV key specified. The number of keys returned are the minimum of:

a) the number of keys in the controller; or
b) the number of complete keys that fit in the buffer provided by the host.

The command uses Command Dword 2, Command Dword 3, Command Dword 10, Command Dword 11, 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

If the value in the Key Length field is greater than 16, then the controller shall abort the command with Invalid Field in Command.

Figure 10: List – Command Dword 10

Bit	Description
31:00	**Host Buffer Size (HBS):** This field indicates the host buffer size in bytes.

Figure 11: List – Command Dword 11

Bit	Description
31:8	**Reserved**
7:0	**Key Length (KL):** Specifies the length of the KV key in bytes.

Figure 12: List – Command Dword 2 and Command Dword 3

Bit	Description
63:0	**KV key[63:00]:** This field specifies least-significant 64-bits of the KV key to be used for the command. Command Dword 2 contains bits 31:00; Command Dword 3 contains bits 63:32.

Figure 13: List – Command Dword 14 and Command Dword 15

Bit	Description
63:0	**KV key[127:64]:** This field specifies the most-significant 64-bits of the KV key to be used for the command. Command Dword 14 contains bits 95:64; Command Dword 15 contains bits 127: 96.

3.2.2.1 Command Completion

Upon completion of the List command, the controller shall post a completion queue entry to the associated I/O Completion Queue indicating the status for the command.

The command returns a list of KV keys that exist as described in 3.2.2.2.

List command specific status values are defined in Figure 14.

Figure 14: List – Command Specific Status Values

Value	Description
86h	**Invalid Key Size:** The KV key size is not valid
0Bh	**Invalid Namespace or Format:** The namespace or the format of that namespace is invalid.

3.2.2.2 List command return data structure

The data structure returned for the list command is as defined in Figure 15.

Figure 15: List – Return data structure

Byte	Description
03:00	**Number of Returned Keys (NRK):** This value reflects how many KV keys are returned in this data structure.
	Key data structure 1 (refer to Figure 16)
	Key data structure 2 (refer to Figure 16)
	…
	Key data structure n (refer to Figure 16

Figure 16: Key data structure

Byte	Description
01:00	**Key Length (KL):** indicates the length of the KV key in bytes that this data structure represents.
n:02	**Key:** KV key that this entry describes.
m:n	**Pad:** Pad necessary, if any to end the data structure on a 4 byte boundary.

3.2.3 Retrieve command

The Retrieve command retrieves a KV value from 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 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.

If the value in the Key Length field is greater than 16, then the controller shall abort the command with Invalid Field in Command.

Figure 17: Retrieve – Data Pointer

Bit	Description
127:00	**Data Pointer (DPTR):** This field specifies where data is transferred to. Refer to the NVMe Base Specification for the definition of this field.

Figure 18: Retrieve – Command Dword 10

Bit	Description
31:00	**Host Buffer Size (HBS):** This field indicates the host buffer size in bytes.

Figure 19: Retrieve – Command Dword 11

Bit	Description
31:16	Reserved
15:8	**Retrieve Option (RO):** This field specifies the retrieve option. Bits 15:9 are reserved. Bit 8 if set to '1' specifies that the controller shall return raw data (i.e., no decompression is performed on the data). Bit 8 if cleared to '0' specifies that the controller shall return decompressed data if compression is supported. Control of compression algorithms, if any, and their use by the controller is outside the scope of this specification. If the controller does not compress data then this bit is ignored.
7:0	**Key Length (KL):** Specifies the length of the KV key in bytes.

Figure 20: Retrieve – Command Dword 2 and Command Dword 3

Bit	Description
63:0	**KV key[63:00]:** This field specifies the least-significant 64-bits of the KV key to be used for the command. Command Dword 2 contains bits 31:00; Command Dword 3 contains bits 63:32.

Figure 21: Retrieve –Command Dword 14 and Command Dword 15

Bit	Description
63:0	**KV key[127:64]:** This field specifies the most-significant 64-bits of the KV key to be used for the command. Command Dword 14 contains bits 95:64; Command Dword 15 contains bits 127: 96.

3.2.3.1 Command Completion

Upon completion of the Retrieve command, the controller shall post a completion queue entry to the associated I/O Completion Queue indicating the status for the command. On successful completion of the command, Dword 0 of the completion queue entry contains the KV value size in bytes.

If the host buffer size is less than the size of the KV value then the portion of the KV value that fits in the host buffer shall be returned starting at the beginning of the KV value. If the host requires the entire value, then the host should issue a subsequent Retrieve command with a buffer large enough to retrieve the KV value length returned in the I/O Completion Queue.

Retrieve command specific status values are defined in Figure 22.

Figure 22: Retrieve – Command Specific Status Values

Value	Description
86h	**Invalid Key Size:** The KV key size is not valid
0Bh	**Invalid Namespace or Format:** The namespace or the format of that namespace is invalid.
87h	KV Key Does Not Exist**:** The KV key does not exist
88h	**Unrecovered Error:** There was an unrecovered error when reading from the medium

3.2.4 Exist command

The Exist command returns a status indicating if the specified KV key exists.

The command uses Command Dword 2, Command Dword 3, Command Dword 11, Command Dword 14, and Command Dword 15 fields. All other command specific fields are reserved.

If the value in the Key Length field is greater than 16, then the controller shall abort the command with Invalid Field in Command.

Figure 23: Exist – Command Dword 11

Bit	Description
31:8	**Reserved**
7:0	**Key Length (KL):** Specifies the length of the KV key in bytes.

Figure 24: Exist – Command Dword 2 and Command Dword 3

Bit	Description
63:0	**KV key[63:00]:** This field specifies the least-significant 64-bits of the KV key to be used for the command. Command Dword 2 contains bits 31:00; Command Dword 3 contains bits 63:32.

Figure 25: Exist – Command Dword 14 and Command Dword 15

Bit	Description
63:0	**KV key[127:64]:** This field specifies most-significant 64-bits of the KV key to be used for the command. Command Dword 14 contains bits 95:64; Command Dword 15 contains bits 127: 96.

3.2.4.1 Command Completion

Upon completion of the Exist command, the controller posts a completion queue entry (CQE) to the associated I/O Completion Queue. If the status code returned is 00h, then the KV key exists. The Exist command specific status values are defined in Figure 26.

Figure 26: Exist – Command Specific Status Values

Value	Description
87h	**KV Key Does Not Exist**:The KV key does not exist.

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 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.

Figure 27: Store – Data Pointer

Bit	Description
127:00	**Data Pointer (DPTR):** This field specifies the location of a data buffer from which data is transferred. Refer to the NVMe Base Specification for the definition of this field.

Figure 28: Store – Command Dword 10

Bit	Description
31:00	**Value Size (VS):** This field indicates the KV value size in bytes. A KV value of 0h specifies that there is no value associated with this KV key but that the KV key exists.

Figure 29: Store – Command Dword 11

Bit	Description
31:16	**Reserved**
15:8	**Store Option (SO):** Specifies the store option Bits 15:11 are reserved Bit 10 if set to '1' specifies that the controller shall not compress the KV value. Bit 10 if cleared to '0' specifies that the controller shall compress the KV value if compression is supported. Bit 9 if set to '1' specifies that the controller shall not store the KV value if the KV key exists. Bit 9 if cleared to '0' specifies that the controller shall store the KV value if other Store Options are met. Bit 8 if set to '1' specifies that the controller shall not store the KV value if the KV key does not exists. Bit 8 if cleared to '0' specifies that the controller shall store the KV value if other Store Options are met.
7:0	**Key Length (KL):** Specifies the length of the KV key in bytes.

Figure 30: Store – Command Dword 2 and Command Dword 3

Bit	Description
63:0	**KV key[63:00]:** This field specifies Figure 16he least-significant 64-bits of the KV key to be used for the command. Command Dword 2 contains bits 31:00; Command Dword 3 contains bits 63:32.

Figure 31: Store –Command Dword 14 and Command Dword 15

Bit	Description
63:0	**KV key[127:64]:** This field specifies the most-significant 64-bits of the KV key to be used for the command. Command Dword 14 contains bits 95:64; Command Dword 15 contains bits 127: 96.

3.2.5.1 Command Completion

Upon completion of the Store command, the controller shall post a completion queue entry to the associated I/O Completion Queue indicating the status for the command.

Store command specific errors are defined in Figure 32.

Figure 32: Store – Command Specific Status Values

Value	Description
85h	**Invalid Value Size:** The value size is not valid
86h	**Invalid Key Size:** The KV key size is not valid
0Bh	**Invalid Namespace or Format:** The namespace or the format of that namespace is invalid.
81h	**Capacity Exceeded:** The capacity of the device was exceeded
89h	**Key Exists:** Store Option bit 9 is set to '1' and the KV key exists
87h	**KV Key Does Not Exist:** Store Option bit 8 is set to '1' and the KV key does not exist

4 Admin Commands for the Key Value Command Set

4.1 Admin Command behavior for the Key Value Command Set

The Admin Commands are as defined in the NVMe Base Specification. The Key Value Command Set specific behavior for Admin Commands is described in this section.

4.1.1 Asynchronous Event Request command

The Asynchronous Event Request command operates as defined in the NVMe Base Specification. The Key Value Command Set does not define any additional Asynchronous Events.

4.1.2 Format NVM command – Key Value Command Set Specific

The Format NVM command operates as defined in the NVMe Base Specification. The Format Index indicates a valid KV Format from the KV Format field in the Key Value Command Set specific Identify Namespace data structure.

4.1.3 Get Features & Set Features commands

Figure 33 defines the Features support requirements for I/O Controllers supporting the Key Value Command Set.

Figure 33: Feature Identifiers – Key Value Command Set

Feature Identifier	Persistent Across Power Cycle and Reset[1]	Uses Memory Buffer for Attributes	Description
20h	Yes	No	Key Value Configuration

NOTES:
1. This column is only valid if the feature is not saveable (refer to the NVME Base Specification). If the feature is saveable, then this column is not used and any feature may be configured to be saved across power cycles and reset.

4.1.3.1 Key Value Configuration (Feature Identifier 20h)

This Feature controls behavior of the Key Value Command Set. The scope of this Feature is the namespace.

The attributes are indicated in Command Dword 11.

If a Get Features command is submitted for this Feature, the attributes specified in Figure 34 are returned in Dword 0 of the completion queue entry for that command.

If the capabilities of the Key Value Config Feature Identifier are both changeable and saveable (refer to the NVMe Base Specification), then the host is able to configure this Feature when initially provisioning a device.

Figure 34: Key Value Config – Command Dword 11

Bits	Description
31:01	Reserved

Figure 34: Key Value Config – Command Dword 11

Bits	Description
00	**Error on Delete of Non-Existent KV Key (EDNEK):** This bit defines the response of the controller to a Delete command processed for a KV key that does not exist. If this bit is set to '1' and the controller process a Delete command that specifies a KV key that does not exist, then the controller shall abort the command with a status code of KV Key Does Not Exist. If this bit is cleared to '0' and the controller process a Delete command that specifies a KV key that does not exist, then the controller shall not abort the command with a status code of KV Key Does Not Exist. (i.e., complete the command as if the KV key existed and was deleted).

4.1.4 Get Log Page command

The Get Log Page command operates as defined in the NVMe Base Specification. 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 a status code of Invalid Field in Command.

Log page scope is as defined in the NVMe Base Specification.

The rules for namespace identifier usage are specified in the NVMe Base Specification.

4.1.5 Identify Command

This specification implements the Identify Command and associated Identify data structures defined in the NVMe Base Specification. Additionally, the Key Value Command Set specifies the data structures defined in this section.

Each I/O Command Set is assigned a specific Command Set Identifier (CSI) value by the NVMe Base Specification. The Key Value Command Set is assigned a CSI value of 01h.

Figure 35: Identify – CNS Values

CNS Value	O/M [1]	Definition	NSID [2]	CNTID [3]	CSI [4]	Reference Section
Active Namespace Management						
05h	M [5]	Identify I/O Command Set specific Namespace data structure for the specified NSID for the I/O Command Set specified in the CSI field.	Y	N	Y	4.1.5.1
06h	M	Identify I/O Command Set Specific Controller data structure for the controller processing the command.	Y	N	Y	4.1.5.2

NOTES:
1. O/M definition: O = Optional, M = Mandatory.
2. The NSID field is used: Y = Yes, N = No.
3. The CDW10.CNTID field is used: Y = Yes, N = No.
4. The CDW11.CSI field is used: Y = Yes, N = No.
5. Mandatory for controllers that support the Namespace Management capability (refer to the NVMe Base Specification).

4.1.5.1 I/O Command Set Specific Identify Namespace Data Structure (CNS 05h, CSI 01h)

The I/O Command Set specific Identify Namespace data structure (i.e., CNS 05h) for the Key Value Command Set is defined in Figure 36.

Figure 36: Identify – I/O Command Set Specific Identify Namespace Data Structure, Key Value Type Specific

Bytes	O/M [1]	Description
07:00	M	**Namespace Size (NSZE):** This field indicates the total size of the namespace in bytes. This is the space to store KV keys and KV values. This field is undefined prior to the namespace being formatted.
15:08		Reserved
23:16	M	**Namespace Utilization (NUSE):** This field indicates the current number of bytes allocated in the namespace. This is the space to store KV keys and KV values. This field is less than or equal to the Namespace Size field. A key value pair is allocated when it is written with a Store command. A key value pair is deallocated using the Delete command. If the controller supports Asymmetric Namespace Access Reporting (refer to the CMIC field), and the relationship between the controller and the namespace is in the ANA Inaccessible state (refer to the NVMe Base Specification) or the ANA Persistent Loss state (refer to the NVMe Base Specification), then this field shall be cleared to 0h.
24	M	**Namespace Features (NSFEAT):** This field defines features of the namespace. Bits 7:4 are reserved. Bit 3 if set to '1' indicates that the non-zero NGUID and non-zero EUI64 fields for this namespace are never reused by the controller. If cleared to '0', then the NGUID and EUI64 values may be reused by the controller for a new namespace created after this namespace is deleted. This bit shall be cleared to '0' if both NGUID and EUI64 fields are cleared to 0h. Refer to the NVMe Base Specification. Bits 2:0 are reserved.
25	M	**Number of KV Formats (NKVF):** This field defines the number of KV format descriptors supported by the namespace. KV formats shall be allocated in order (starting with 0) and packed sequentially. This is a 0's based value. The maximum number of KV formats that may be indicated as supported is 16. The supported KV formats are indicated in bytes 72 to 327 in this data structure. The KV Format fields with a Format Index beyond the value set in this field are invalid and not supported. KV Formats that are valid, but not currently available may be indicated by clearing the KV Key Max Length field to 0h and clearing the KV Value Max Length field to 0h for that KV Format.
26	O	**Namespace Multi-path I/O and Namespace Sharing Capabilities (NMIC):** Refer to the NMIC field in the I/O Command Set Independent Identify Namespace data structure in the NVMe Base Specification.
27	O	**Reservation Capabilities (RESCAP):** Refer to the NVMe Base Specification.
28	O	**Format Progress Indicator (FPI):** Refer to the NVMe Base Specification.
31:29		Reserved
35:32	O	**Namespace Optimal Value Granularity (NOVG):** This field indicates the optimal value granularity for this namespace. This field is specified in bytes. The host should construct Store commands that store multiples of NOVG bytes to achieve optimal performance. A value of 0h indicates that no optimal value granularity is reported.
39:36	O	**ANA Group Identifier (ANAGRPID):** Refer to the NVMe Base Specification.
42:40		Reserved
43	O	**Namespace Attributes (NSATTR):** Refer to the NVMe Base Specification.
45:44	O	**NVM Set Identifier (NVMSETID):** Refer to the NVMe Base Specification.
47:46	O	**Endurance Group Identifier (ENDGID):** Refer to the NVMe Base Specification.
63:48	O	**Namespace Globally Unique Identifier (NGUID):** Refer to the NVMe Base Specification.
71:64	O	**IEEE Extended Unique Identifier (EUI64):** Refer to the NVMe Base Specification.

Figure 36: Identify – I/O Command Set Specific Identify Namespace Data Structure, Key Value Type Specific

Bytes	O/M [1]	Description
KV Formats		
87:72	M	KV Format 0 Support (KVF0): This field indicates the KV format 0 that is supported by the controller. The KV format field is defined in Figure 37.
103:88	O	KV Format 1 Support (KVF1): This field indicates the KV format 1 that is supported by the controller. The KV format field is defined in Figure 37.
119:104	O	KV Format 2 Support (KVF2): This field indicates the KV format 2 that is supported by the controller. The KV format field is defined in Figure 37.
135:120	O	KV Format 3 Support (KVF3): This field indicates the KV format 3 that is supported by the controller. The KV format field is defined in Figure 37.
151:136	O	KV Format 4 Support (KVF4): This field indicates the KV format 4 that is supported by the controller. The KV format field is defined in Figure 37.
167:152	O	KV Format 5 Support (KVF5): This field indicates the KV format 5 that is supported by the controller. The KV format field is defined in Figure 37.
183:168	O	KV Format 6 Support (KVF6): This field indicates the KV format 6 that is supported by the controller. The KV format field is defined in Figure 37.
199:184	O	KV Format 7 Support (KVF7): This field indicates the KV format 7 that is supported by the controller. The KV format field is defined in Figure 37.
215:200	O	KV Format 8 Support (KVF8): This field indicates the KV format 8 that is supported by the controller. The KV format field is defined in Figure 37.
231:216	O	KV Format 9 Support (KVF9): This field indicates the KV format 9 that is supported by the controller. The KV format field is defined in Figure 37.
247:232	O	KV Format 10 Support (KVF10): This field indicates the KV format 10 that is supported by the controller. The KV format field is defined in Figure 37.
263:248	O	KV Format 11 Support (KVF11): This field indicates the KV format 11 that is supported by the controller. The KV format field is defined in Figure 37.
279:264	O	KV Format 12 Support (KVF12): This field indicates the KV format 12 that is supported by the controller. The KV format field is defined in Figure 37
295:280	O	KV Format 13 Support (KVF13): This field indicates the KV format 13 that is supported by the controller. The KV format field is defined in Figure 37.
311:296	O	KV Format 14 Support (KVF14): This field indicates the KV format 14 that is supported by the controller. The KV format field is defined in Figure 37.
327:312	O	KV Format 15 Support (KVF15): This field indicates the KV format 15 that is supported by the controller. The KV format field is defined in Figure 37.
3839:328		Reserved
4095:3840	O	Vendor Specific

NOTES:
1. O/M definition: O = Optional, M = Mandatory.

Figure 37: KV Format Data Structure

Bytes	Description
01:00	**KV Key Max Length:** Maximum length of a KV key in a key value pair in bytes. The maximum value that is supported by the commands in this version of the Key Value Command Set specification is 16.
02	Reserved

Figure 37: KV Format Data Structure

<table>
<tr><th>Bytes</th><th>Description</th></tr>
<tr><td>03</td><td>Additional format options:
Bits 7:2 Reserved
Bits 1:0 Relative Performance (RP): This field indicates the relative performance of the KV format indicated, relative to other KV formats supported by the controller. Depending on the characteristics of the format, there may be performance implications. The performance analysis is based on better performance on a queue depth of 32 with 4KiB KV value reads. The meanings of the values indicated are included in the following table.
<table>
<tr><th>Value</th><th>Definition</th></tr>
<tr><td>00b</td><td>Best performance</td></tr>
<tr><td>01b</td><td>Better performance</td></tr>
<tr><td>10b</td><td>Good performance</td></tr>
<tr><td>11b</td><td>Degraded performance</td></tr>
</table></td></tr>
<tr><td>07:04</td><td>KV Value Max Length: Maximum length in bytes of a KV value in a key value pair.</td></tr>
<tr><td>11:08</td><td>Max Num Keys: Maximum number of KV keys allowed in the namespace. A value of 0h indicates that no maximum number is indicated.</td></tr>
<tr><td>15:12</td><td>Reserved</td></tr>
</table>

4.1.5.2 I/O Command Set Specific Identify Controller Data Structure (CNS 06h, CSI 01h)

The Key Value Command Set does not have an Identify I/O Command Set specific Controller data structure (i.e., CNS 06h). The controller shall return a zero filled data structure for this CNS value.

4.1.6 Namespace Management command

The Namespace Management command operates as defined in the NVMe Base Specification.

The host specified namespace management fields are specific to the I/O Command Set. The data structure passed to the create operation for the Key Value Command Set (CSI 01h) is defined in Figure 38. Fields that are reserved should be cleared to 0h by host software. After successful completion of a Namespace Management command with the create operation, the namespace is formatted with the specified attributes.

Figure 38: Namespace Management – Host Software Specified Fields

Bytes	Description	Host Specified
Fields that are a subset of the I/O Command Set specific Identify Namespace data structure (refer to Figure 36)		
07:00	Namespace Size (NSZE)	Yes
29:08	Reserved	
30	Namespace Multi-path I/O and Namespace Sharing Capabilities (NMIC)	Yes
91:31	Reserved	
95:92	ANA Group Identifier (ANAGRPID) [1]	Yes
99:96	Reserved	
101:100	NVM Set Identifier (NVMSETID) [1]	Yes
103:102	Endurance Group Identifier (ENDGID)	Yes
511:104	Reserved	

Notes:
1. A value of 0h specifies that the controller determines the value to use (refer to the Namespace Management section in the NVMe Base Specification). If the associated feature is not supported, then this field is ignored by the controller.

4.1.7 Sanitize command

The Sanitize command operates as defined in the NVMe Base Specification. There are no Key Value Command Set specific requirements on the Sanitize command.

5 Extended Capabilities

5.1 Namespace Management

Namespace Management operates as defined in the NVMe Base Specification.

5.2 Reservations

Reservations operate as defined the NVMe Base Specification with the additional Command Behavior in the Presence of a Reservation defined in Figure 39.

Figure 39: Command Behavior in the Presence of a Reservation

NVMe 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
Key Value Command Set Read Command Group								
Retrieve	A	A	C	C	A	A	C	A
Key Value Command Set Write Command Group								
Delete Flush Format NVM (Admin) Namespace Attachment (Admin) Namespace Management (Admin) Sanitize (Admin) Security Send (Admin) Store	C	C	C	C	C	A	C	A
Key: A definition: A=Allowed, command processed normally by the controller C definition: C=Conflict, command aborted by the controller with status Reservation Conflict								

5.3 Sanitize Operations

A sanitize operation is performed as defined in the NVMe Base Specification.