



NVM Express®

Zoned Namespace Command Set Specification

Revision 1.1c
October 3rd, 2022

Please send comments to info@nvmexpress.org

NVM Express® Zoned Namespace Command Set Specification, Revision 1.1c is available for download at <http://nvmexpress.org>. The NVM Express Zoned Namespace Command Set Specification, Revision 1.1c incorporates the NVM Express Zoned Namespace Command Set Specification, Revision 1.1 (refer to the Zoned Namespace Command Set Specification change list <https://nvmexpress.org/changes-in-nvm-express-revision-2-0> for details), ECN 001, ECN102, ECN105, ECN106, ECN108, ECN109, and ECN110.

SPECIFICATION DISCLAIMER

LEGAL NOTICE:

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

This NVM Express Zoned Namespace Command Set Specification, Revision 1.1c 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 Zoned Namespace Command Set Specification, Revision 1.1c 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.

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

1	INTRODUCTION.....	5
1.1	Overview.....	5
1.2	Scope.....	5
1.3	Conventions.....	5
1.4	Definitions	5
1.5	References	7
2	ZONED NAMESPACE COMMAND SET MODEL.....	8
2.1	Theory of operation.....	8
2.2	I/O Controller Requirements.....	16
3	I/O COMMANDS FOR THE ZONED NAMESPACE COMMAND SET.....	17
3.1	Submission Queue Entry and Completion Queue Entry	17
3.2	Zoned Namespace Command Set Commands.....	17
3.3	NVM Command Set I/O Commands	18
3.4	Zoned Namespace Command Set I/O Commands.....	21
4	ADMIN COMMANDS FOR THE ZONED NAMESPACE COMMAND SET.....	33
4.1	Admin Command behavior for the Zoned Namespace Command Set.....	33
4.2	I/O Command Set Specific Admin commands	40
5	EXTENDED CAPABILITIES	41
5.1	Reservations.....	41
5.2	Directives	41
5.3	Zone Descriptor Extension	41
5.4	Reset Zone Recommended.....	41
5.5	Finish Zone Recommended	42
5.6	Zone Active Excursions	42
ANNEX A.	ZONED NAMESPACE HOST CONSIDERATIONS (NORMATIVE).....	44
A.1	Introduction	44
A.2	Writing to Zones.....	44
A.3	Open Zone Considerations.....	44
A.4	Partial Failures.....	45
A.5	Capacity and Sizes	45

Table of Figures

Figure 1: NVMe Family of Specifications	5
Figure 2: Zones in a Zoned Namespace	8
Figure 3: Summary of Zone Descriptor Attributes	8
Figure 4: Zone Characteristics	9
Figure 5: Write Pointer in an Empty Zone	10
Figure 6: Write Pointer in a Partially Written Zone	10
Figure 7: Zone State Machine	12
Figure 8: Zone Resources	14
Figure 9: I/O Controller – Zoned Namespace Command Set Support	16
Figure 10: I/O Controller – Zoned Namespace Command Set Specific Log Page Support	16
Figure 11: Status Code – Command Specific Status Values, Zoned Namespace Command Set	17
Figure 12: Opcodes for Zoned Namespace Command Set I/O Commands	17
Figure 13: Compare – Command Specific Status Values	18
Figure 14: Copy – Command Specific Status Values	19
Figure 15: Dataset Management – Command Specific Status Values	19
Figure 16: Flush – Command Specific Status Values	20
Figure 17: Read – Command Specific Status Values	20
Figure 18: Verify – Command Specific Status Values	20
Figure 19: Write – Command Specific Status Values	20
Figure 20: Write Uncorrectable – Command Specific Status Values	21
Figure 21: Write Zeroes – Command Specific Status Values	21
Figure 22: Zone Append – Metadata Pointer	22
Figure 23: Zone Append – Data Pointer	22
Figure 24: Zone Append – Command Dword 2 and Dword 3	22
Figure 25: Zone Append – Command Dword 10 and Command Dword 11	22
Figure 26: Zone Append – Command Dword 12	22
Figure 27: Zone Append – Command Dword 13	23
Figure 28: Zone Append – Command Dword 14	23
Figure 29: Zone Append – Command Dword 15	23
Figure 30: Zone Append – Command Specific Status Values	24
Figure 31: Zone Management Receive – Data Pointer	24
Figure 32: Zone Management Receive – Command Dword 10 and Command Dword 11	25
Figure 33: Zone Management Receive – Command Dword 12	25
Figure 34: Zone Management Receive – Command Dword 13	25
Figure 35: Report Zones Data Structure	26
Figure 36: Extended Report Zones Data Structure	27
Figure 37: Zone Descriptor Data Structure	27
Figure 38: Zone Management Send – Command Dword 10 and Command Dword 11	29
Figure 39: Zone Management Send – Command Dword 13	29
Figure 40: Zone Management Send – Command Specific Status Values	32
Figure 41: Zone Management Send – Completion Queue Entry Dword 0	32
Figure 42: Asynchronous Event Information – Notice, Zoned Namespace Command Set	33
Figure 43: Asynchronous Event Request – Completion Queue Entry Dword 1	33
Figure 44: Asynchronous Event Configuration – Command Dword 11	34
Figure 45: Log Page Identifiers	34
Figure 46: Changed Zone List Log Page	35
Figure 47: CNS Values	35
Figure 48: I/O Command Set Specific Identify Namespace Data Structure for the Zoned Namespace Command Set	36
Figure 49: LBA Format Extension Data Structure	38
Figure 50: I/O Command Set Specific Identify Controller Data Structure for the Zoned Namespace Command Set	38
Figure 51: Sanitize Behavior for the Zoned Namespace Command Set	39
Figure 52: Command Behavior in the Presence of a Reservation	41
Figure 53: Size and Capacity Fields	45
Figure 54: Zone Size Relationships	46

1 Introduction

1.1 Overview

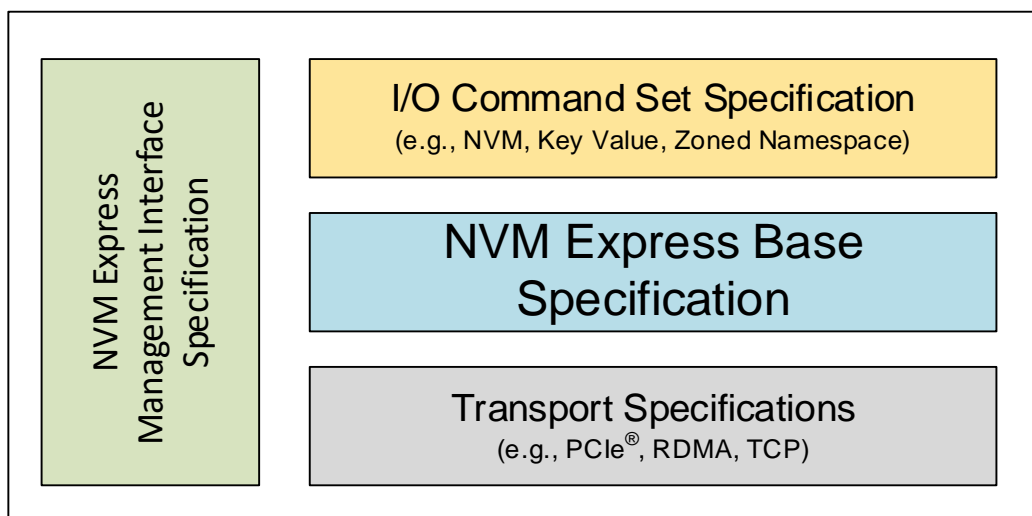
The NVM Express® (NVMe®) Base Specification defines an interface for host software to communicate with a non-volatile memory subsystem over a variety of memory based transports and message based transports.

This document defines a specific NVMe I/O Command Set, the Zoned Namespace Command Set, which extends the NVM Express Base Specification and the NVM Command Set Specification.

1.2 Scope

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

Figure 1: NVMe Family of Specifications



This specification supplements the NVM Express 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 Zoned Namespace Command Set. Functionality that is applicable generally to NVMe or that is applicable across multiple I/O Command Sets is defined in the NVM Express 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. NVM Express Base Specification
3. NVMe transport specifications
4. NVMe I/O command set specifications
5. NVM Express Management Interface Specification

1.3 Conventions

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

1.4 Definitions

1.4.1 Definitions from the NVM Express Base Specification

This specification uses the definitions in the NVM Express Base Specification.

1.4.2 Terms in the NVM Express Base Specification specified in the Zoned Namespace Command Set

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

1.4.2.1 Endurance Group Host Read Command

An Endurance Group Host Read Command as defined in the NVM Command Set Specification.

1.4.2.2 Format Index

A value used to index into the LBA Formats list (refer to the NVM Command Set Specification), the Extended LBA Formats list (refer to the NVM Command Set Specification), and the LBA Format Extensions list (refer to Figure 48).

1.4.2.3 SMART Data Units Read Command

A SMART Data Units Read Command as defined in the NVM Command Set Specification.

1.4.2.4 SMART Host Read Command

A SMART Host Read Command as defined in the NVM Command Set Specification.

1.4.2.5 User Data Out Command

A User Data Out Command as defined in the NVM Command Set Specification and the Zone Append command.

1.4.3 Definitions from the NVM Command Set Specification

The following terms are defined in the NVM Command Set Specification and are used in this specification:

- a) LBA range
- b) logical block
- c) logical block address (LBA)
- d) User Data Format

1.4.4 Definitions specific to the Zoned Namespace Command Set

This section defines terms that are specific to this specification.

1.4.4.1 active zone

A zone that is in the ZSIO:Implicitly Opened state, the ZSEO:Explicitly Opened state, or the ZSC:Closed state.

1.4.4.2 Address-Specific Write Command

A Write command, Write Uncorrectable command, Write Zeroes command, or Copy command. Address-Specific Write Commands specify a range of logical block addresses in command parameters as part of the Submission Queue Entry or in data structures pointed to by the command parameters.

1.4.4.3 open zone

A zone that is in the ZSIO:Implicitly Opened state, the ZSEO:Explicitly Opened state.

1.4.4.4 write operation

An operation initiated by a Write command, a Write Zeroes command, a Write Uncorrectable command, a Copy command, or a Zone Append command.

1.4.4.5 zone

A contiguous range of logical block addresses that are managed as a single unit.

1.4.4.6 Zone Descriptor

The data structure that contains information about a zone.

1.4.4.7 Zone Descriptor Extension

Host defined data that is associated with a zone.

1.4.4.8 zoned namespace

A namespace that is divided into zones and is associated with the Zoned Namespace Command Set.

1.5 References

NVMe® Base Specification, Revision 2.0. Available from <https://www.nvmexpress.org>.

NVMe® NVM Command Set Specification, Revision 1.0. Available from <https://www.nvmexpress.org>.

2 Zoned Namespace Command Set Model

The NVM Express Base Specification defines an interface for host software to communicate with a non-volatile memory subsystem. This specification defines additional functionality for the Zoned Namespace Command Set.

The Command Set Identifier (CSI) value for this Command Set is 02h.

2.1 Theory of operation

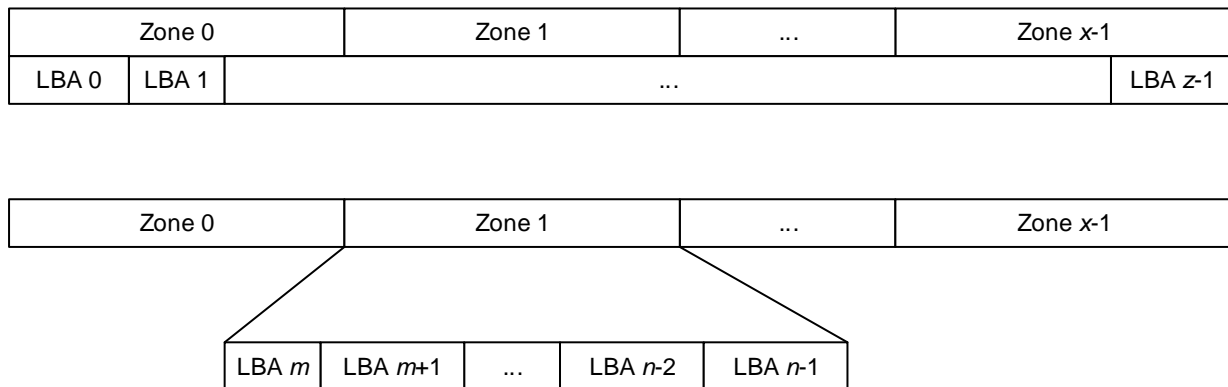
This section defines the operation of the Zoned Namespace Command Set.

2.1.1 Namespaces

A namespace is a collection of NVM and is defined in the NVM Express Base Specification and in the NVM Command Set Specification, as modified by this specification.

A zoned namespace is a namespace that is associated with the Zoned Namespace Command Set. A zoned namespace is divided into a set of equally-sized zones, which are contiguous non-overlapping ranges of logical block addresses. Figure 2 shows a zoned namespace with x zones and z LBAs where LBA 0 is the lowest LBA of zone 0, LBA $z-1$ is the highest LBA of zone $x-1$, and for Zone 1, m is the lowest LBA and $n-1$ is its highest LBA.

Figure 2: Zones in a Zoned Namespace



The Zoned Namespace Command Set is based on the NVM Command Set (refer to the NVM Command Set Specification).

Each zone has an associated Zone Descriptor that contains a set of attributes. A Zone Management Receive command may be used to retrieve one or more Zone Descriptors.

2.1.1.1 Zone Descriptor

The attributes associated with a zone are summarized in Figure 3. These attributes are reported in the Zone Descriptor data structure as defined in Figure 37.

Figure 3: Summary of Zone Descriptor Attributes

Attribute	Description	
Zone Type	The zone type attribute defines the rules for reading and writing to a zone.	
	Zone Type	Reference Section
	Sequential Write Required	2.1.1.2.1

Figure 3: Summary of Zone Descriptor Attributes

Attribute	Description																		
Zone State	Each zone has an associated state machine. That state machine has a set of states and each state, together with the zone type, defines the operational characteristics of that zone.																		
	<table><tr><th>Zone Type</th><th>Zone State</th><th>Reference Section</th></tr><tr><td rowspan="7">Sequential Write Required</td><td>ZSE:Empty</td><td>2.1.1.3.1</td></tr><tr><td>ZSIO:Implicitly Opened</td><td>2.1.1.3.2</td></tr><tr><td>ZSEO:Explicitly Opened</td><td>2.1.1.3.3</td></tr><tr><td>ZSC:Closed</td><td>2.1.1.3.4</td></tr><tr><td>ZSF:Full</td><td>2.1.1.3.5</td></tr><tr><td>ZSRO:Read Only</td><td>2.1.1.3.6</td></tr><tr><td>ZSO:Offline</td><td>2.1.1.3.7</td></tr></table>	Zone Type	Zone State	Reference Section	Sequential Write Required	ZSE:Empty	2.1.1.3.1	ZSIO:Implicitly Opened	2.1.1.3.2	ZSEO:Explicitly Opened	2.1.1.3.3	ZSC:Closed	2.1.1.3.4	ZSF:Full	2.1.1.3.5	ZSRO:Read Only	2.1.1.3.6	ZSO:Offline	2.1.1.3.7
	Zone Type	Zone State	Reference Section																
	Sequential Write Required	ZSE:Empty	2.1.1.3.1																
		ZSIO:Implicitly Opened	2.1.1.3.2																
		ZSEO:Explicitly Opened	2.1.1.3.3																
		ZSC:Closed	2.1.1.3.4																
		ZSF:Full	2.1.1.3.5																
ZSRO:Read Only		2.1.1.3.6																	
ZSO:Offline		2.1.1.3.7																	
Write Pointer	The Write Pointer attribute defines the lowest numbered writeable logical block address in that zone. The validity of the write pointer is zone state specific and is defined per zone type (refer to section 2.1.1.2).																		
Zone Start Logical Block Address	The Zone Start Logical Block Address (ZSLBA) attribute defines the lowest logical block address for that zone.																		
Zone Capacity	The Zone Capacity attribute defines the writeable capacity of that zone.																		
Zone Descriptor Extension Valid	The Zone Descriptor Extension Valid attribute defines the validity of the Zone Descriptor Extension data of that zone. Refer to section 5.3.																		
Reset Zone Recommended	The Reset Zone Recommended attribute indicates that the controller recommends that the host resets that zone. Refer to section 5.4.																		
Finish Zone Recommended	The Finish Zone Recommended attribute indicates that the controller recommends that the host finishes that zone. Refer to section 5.5.																		
Zone Finished by Controller	The Zone Finished by Controller attribute indicates that the controller finished that zone due to a Zone Active Excursion. Refer to section 5.6.																		

2.1.1.2 Zone Types

2.1.1.2.1 Sequential Write Required Zones

A zone type of Sequential Write Required requires the set of logical block addresses of a zone to be written sequentially.

2.1.1.2.1.1 Writing in Sequential Write Required Zones

The following commands may be used to write to logical blocks in a specific zone of zone type Sequential Write Required:

- Write command;
- Write Zeroes command;
- Write Uncorrectable command;
- Copy command; and
- Zone Append command.

A write pointer is maintained for each zone in the zoned namespace that indicates the next writeable logical block address in that zone. The write pointer is valid for a subset of the zone states as defined in Figure 4.

Figure 4: Zone Characteristics

State	Zone Characteristics		
	Valid Write Pointer ¹	Active Resources ²	Open Resources ²
ZSE:Empty	Yes	No	No
ZSIO:Implicitly Opened	Yes	Yes	Yes
ZSEO:Explicitly Opened	Yes	Yes	Yes
ZSC:Closed	Yes	Yes	No
ZSF:Full	No	No	No
ZSRO:Read Only	No	No	No
ZSO:Offline	No	No	No

Figure 4: Zone Characteristics

State	Zone Characteristics		
	Valid Write Pointer ¹	Active Resources ²	Open Resources ²
Notes: 1. A valid write pointer (i.e., Yes) indicates that the write pointer zone attribute within that zone contains a valid logical block address. An invalid write pointer (i.e., No) provides no information. 2. Resources associated with a zone are defined in section 2.1.1.4.			

The host may use the Zone Management Receive command to determine the current write pointer for a zone.

The write pointer for a zone in the ZSE:Empty state, the ZSIO:Implicitly Opened state, the ZSEO:Explicitly Opened state, or the ZSC:Closed state shall be increased by the number of logical blocks written on successful completion of a write operation.

If the controller is not able to successfully write to all logical blocks specified by a command that initiates a write operation, then the write pointer shall:

- a) be set to a value within the range of LBAs specified in that command ;
- b) be set to one greater than the last LBA in the range of LBAs specified in that command; or
- c) become invalid (i.e., due to transitioning to the ZSRO:Read Only state or the ZSO:Offline state, or due to a Zone Active Excursion (refer to section 5.6)).

The Zone Management Send command with a Zone Send Action of Reset Zone sets the write pointer to the ZSLBA for that zone.

Figure 5 shows an example of a zone in the ZSE:Empty state. LBA m is the ZSLBA attribute, the write pointer indicates ZSLBA, and $n-1$ is the highest LBA of the zone.

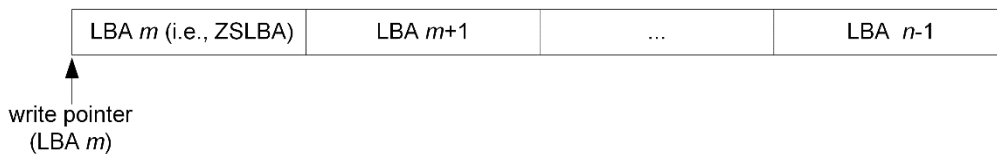
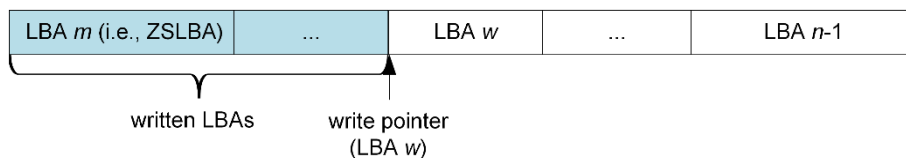
Figure 5: Write Pointer in an Empty Zone

Figure 6 shows an example of a zone in the ZSIO:Implicitly Opened state, the ZSEO:Explicitly Opened state, or the ZSC:Closed state, that has had some logical blocks written. The write pointer, indicated by LBA w , is the lowest-numbered unwritten LBA (i.e., the next LBA to be written) and $n-1$ is the highest LBA of the zone.

Figure 6: Write Pointer in a Partially Written Zone

The controller shall abort a command that writes to a zone that is in the ZSF:Full state with a status code of Zone Is Full.

The controller shall abort a command that writes to a zone that is in the ZSRO:Read Only state, with a status code of Zone Is Read Only.

The controller shall abort a command that writes to a zone that is in the ZSO:Offline state, with a status code of Zone Is Offline.

For a zone in the ZSE:Empty state, the ZSIO:Implicitly Opened state, the ZSEO:Explicitly Opened state, or the ZSC:Closed state:

- a) if an Address-Specific Write Command specifies a Starting LBA field that is not equal to the write pointer for that zone, then the controller shall abort that command with a status code of Zone Invalid Write; and
- b) if a Zone Append command specifies a ZSLBA that is not the lowest logical block address in that zone, then the controller shall abort that command with a status code of Invalid Field in Command.

The controller shall abort a command that initiates a write operation that has a starting LBA in one zone, and for which the Number of Logical blocks exceeds the remaining number of logical blocks in that zone, with a status code of Zone Boundary Error.

2.1.1.2.1.2 Reading in Sequential Write Required Zones

If the zone for a User Data Read Access Command is:

- a) in the ZSO:Offline state, then that command shall be aborted with a status code of Zone Is Offline; or
- b) in any other state, then that User Data Read Access Command is performed as described in this section, in section 2.1.1.5, and in the Deallocated or Unwritten Logical Blocks section in the NVM Command Set Specification.

If the Read Across Zone Boundaries bit is set to '1' in the Zoned Namespace Command Set specific Identify Namespace data structure (refer to Figure 48), then User Data Read Access Commands are allowed to specify an LBA range containing logical blocks in more than one zone.

If the Read Across Zone Boundaries bit is cleared to '0', then commands that perform read operations that specify an LBA range containing logical blocks in more than one zone shall be aborted with a status code of Zone Boundary Error.

2.1.1.3 Zone State Machine

There is a state machine associated with each zone. The state machine controls the operational characteristics of each zone. The state machine consists of the following states: ZSE:Empty, ZSIO:Implicitly Opened, ZSEO:Explicitly Opened, ZSC:Closed, ZSF:Full, ZSRO:Read Only, and ZSO:Offline.

The initial state of a zone state machine is set as a result of:

- a) an NVM Subsystem Reset; or
- b) all controllers in the NVM subsystem reporting controller shutdown processing complete (i.e., the Shutdown Type (ST) bit cleared to '0' and the Shutdown Status (SHST) field set to 10b, refer to the NVM Express Base Specification).

The initial state for each zone is the:

- a) ZSE:Empty state, if the write pointer is valid, the write pointer points to the lowest LBA in the zone, and the Zone Descriptor Extension Valid bit is cleared to '0';
- b) ZSC:Closed state:
 - a) if the write pointer is valid and does not point to the lowest LBA in the zone; or
 - b) if the write pointer is valid and the Zone Descriptor Extension Valid bit is set to '1';
- c) ZSF:Full state:
 - a) if the most recent state was the ZSF:Full; or
 - b) if the zone state was transitioned to the ZSF:Full state as a result of the NVM Subsystem Reset;
- d) ZSRO:Read Only state, if the most recent zone state was the ZSRO:Read Only state; and
- e) ZSO:Offline state, if the most recent zone state was the ZSO:Offline state.

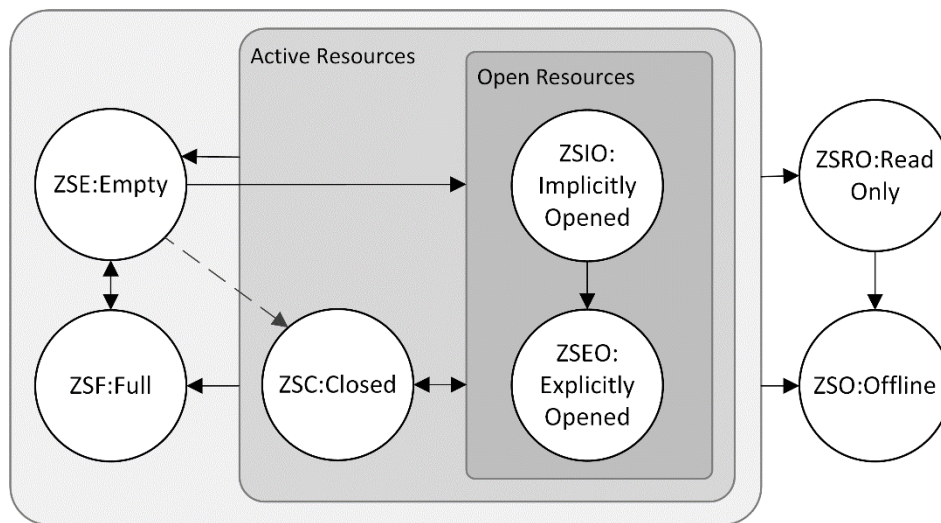
Transitions between zone states cause reporting of a Zone Descriptor Changed asynchronous event as defined in Figure 42.

If the zoned namespace is formatted with a Format NVM command or created with a Namespace Management command, the zones in the zoned namespace are initialized to either the ZSE:Empty state or the ZSO:Offline state.

Zones that are in the ZSE:Empty state, the ZSIO:Implicitly Opened state, the ZSEO:Explicitly Opened state, the ZSC:Closed state, or the ZSF:Full state may transition to the ZSRO:Read Only state or the ZSO:Offline state by mechanisms outside the scope of this specification.

Figure 7 shows the valid transitions between each zone state from the time a zoned namespace is formatted or created. The transition (dotted line) from the ZSE:Empty state to the ZSC:Closed state is only valid when the zoned namespace is formatted with Zone Descriptor Extension support (refer to section 5.3). An arrow to or from a shaded area indicates transitions to or from all states in that area.

Figure 7: Zone State Machine



The processing of a command may cause multiple state machine transitions (e.g., the processing of a Write command may cause a transition from the ZSC:Closed state to the ZSIO:Implicitly Opened state and then cause a transition from the ZSIO:Implicitly Opened state to the ZSF:Full state).

2.1.1.3.1 ZSE:Empty state

Transition ZSE:ZSIO: The zone shall transition from the ZSE:Empty state to the ZSIO:Implicitly Opened state if there are available resources as defined in section 2.1.1.4 and a write operation writes one or more logical blocks of that zone.

Transition ZSE:ZSEO: The zone shall transition from the ZSE:Empty state to the ZSEO:Explicitly Opened state if there are available resources as defined in section 2.1.1.4 and a Zone Management Send command with a Zone Send Action of Open Zone completes successfully.

Transition ZSE:ZSC: The zone shall transition from the ZSE:Empty state to the ZSC:Closed state if there are available active resources as defined in section 2.1.1.4 and a Zone Management Send command with a Zone Send Action of Set Zone Descriptor Extension completes successfully.

Transition ZSE:ZSF: The zone shall transition from the ZSE:Empty state to the ZSF:Full state as a result of a Zone Management Send command with a Zone Send Action of Finish Zone completes successfully.

2.1.1.3.2 ZSIO:Implicitly Opened state

Transition ZSIO:ZSE: The zone shall transition from the ZSIO:Implicitly Opened state to the ZSE:Empty state as a result of successful completion of a Zone Management Send command with a Zone Send Action of Reset Zone.

Transition ZSIO:ZSEO: The zone shall transition from the ZSIO:Implicitly Opened state to the ZSEO:Explicitly Opened state as a result of successful completion of a Zone Management Send command with a Zone Send Action of Open Zone.

Transition ZSIO:ZSC: The zone shall transition from the ZSIO:Implicitly Opened state to the ZSC:Closed state as a result of:

- a) successful completion of a Zone Management Send command with a Zone Send Action of Close Zone; and
- b) the controller initiating the transition as defined in section 2.1.1.4.1.

Transition ZSIO:ZSF: The zone shall transition from the ZSIO:Implicitly Opened state to the ZSF:Full state:

- a) as a result of successful completion of a Zone Management Send command with a Zone Send Action of Finish Zone;
- b) as a result of a write operation that writes one or more logical blocks that causes the zone to reach its writeable zone capacity;
- c) due to a Zone Active Excursion (refer to section 5.6); and
- d) as a result of the zoned namespace becoming write protected (refer to the Namespace Write Protection section in the NVM Express Base Specification).

2.1.1.3.3 ZSEO:Explicitly Opened state

Transition ZSEO:ZSE: The zone shall transition from the ZSEO:Explicitly Opened state to the ZSE:Empty state as a result of successful completion of a Zone Management Send command with a Zone Send Action of Reset Zone.

Transition ZSEO:ZSC: The zone shall transition from the ZSEO:Explicitly Opened state to the ZSC:Closed state upon successful completion of a Zone Management Send command with a Zone Send Action of Close Zone.

Transition ZSEO:ZSF: The zone shall transition from the ZSEO:Explicitly Opened state to the ZSF:Full state:

- a) as a result of successful completion of a Zone Management Send command with a Zone Send Action of Finish Zone;
- b) as a result of a write operation that writes one or more logical blocks that causes the zone to reach its writeable zone capacity;
- c) due to a Zone Active Excursion (refer to section 5.6); or
- d) as a result of the zoned namespace becoming write protected (refer to the Namespace Write Protection section in the NVM Express Base Specification).

2.1.1.3.4 ZSC:Closed state

Transition ZSC:ZSE: The zone shall transition from the ZSC:Closed state to the ZSE:Empty state as a result of successful completion of a Zone Management Send command with a Zone Send Action of Reset Zone.

Transition ZSC:ZSIO: The zone shall transition from the ZSC:Closed state to the ZSIO:Implicitly Opened state, if there are available resources as defined in section 2.1.1.4 and a write operation writes one or more logical blocks of that zone.

Transition ZSC:ZSEO: The zone shall transition from the ZSC:Closed state to the ZSEO:Explicitly Opened state if there are available resources as defined in section 2.1.1.4 and a Zone Management Send command with a Zone Send Action of Open Zone completes successfully.

Transition ZSC:ZSF: The zone shall transition from the ZSC:Closed state to the ZSF:Full state:

- a) as a result of successful completion of a Zone Management Send command with a Zone Send Action of Finish Zone;
- b) due to a Zone Active Excursion (refer to section 5.6); and
- c) as a result of the zoned namespace becoming write protected (refer to the Namespace Write Protection section in the NVM Express Base Specification).

2.1.1.3.5 ZSF:Full state

Transition ZSF:ZSE: The zone shall transition from the ZSF:Full state to the ZSE:Empty state as a result of successful completion of a Zone Management Send command with a Zone Send Action of Reset Zone.

2.1.1.3.6 ZSRO:Read Only state

Transition ZSRO:ZSO: The zone shall transition from the ZSRO:Read Only state to the ZSO:Offline state as a result of successful completion of a Zone Management Send command with a Zone Send Action of Offline Zone.

2.1.1.3.7 ZSO:Offline state

There are no transitions from the ZSO:Offline state to any other zone state.

2.1.1.4 Zone Resources

Zones may have associated Active Resources and associated Open Resources. The resources limit the number of zones that are allowed to be in each zone state. The resource relationship is defined in Figure 8.

Figure 8: Zone Resources

Resource	States	Comment
Active	ZSIO:Implicitly Opened, ZSEO:Explicitly Opened, ZSC:Closed	Zones in zone states associated with this resource are limited by the Maximum Active Resources field.
Open	ZSIO:Implicitly Opened, ZSEO:Explicitly Opened	Zones in zone states associated with this resource are limited by the Maximum Open Resources field.

Zones that have associated Open Resources are a subset of the zones that are associated with Active Resources. The Maximum Open Resources field shall be less than or equal to the Maximum Active Resources field in the I/O Command Set specific Identify Namespace Data Structure for the Zoned Namespace Command Set (refer to Figure 48).

2.1.1.4.1 Managing resources

The controller associates Active Resources with zones in the ZSIO:Implicitly Opened state, the ZSEO:Explicitly Opened state, and the ZSC:Closed state. The controller associates Open Resources with zones in the ZSIO:Implicitly Opened state and the ZSEO:Explicitly Opened state. The resource management is as follows:

- A transition from the ZSE:Empty state to the ZSIO:Implicitly Opened state or the ZSEO:Explicitly Opened state increases the resource usage of Active Resources and Open Resources by 1.
- A transition from the ZSE:Empty state to the ZSC:Closed state increases the resource usage of Active Resources by 1.
- A transition from the ZSIO:Implicitly Opened state or the ZSEO:Explicitly Opened state to the ZSC:Closed state decreases the resource usage of Open Resources by 1.
- A transition from the ZSC:Closed state to the ZSIO:Implicitly Opened state or the ZSEO:Explicitly Opened state increases the resource usage of Open Resources by 1.
- A transition from the ZSIO:Implicitly Opened state or the ZSEO:Explicitly Opened state to the ZSE:Empty state, the ZSF:Full state, the ZSRO:Read Only state, or the ZSO:Offline state decreases the resource usage of Active Resources and Open Resources by 1.
- A transition from the ZSC:Closed state to the ZSE:Empty state, the ZSF:Full state, the ZSRO:Read Only state, or the ZSO:Offline state decreases the resource usage of Active Resources by 1.
- A transition from the ZSE:Empty state to the ZSF:Full state shall not impact resource usage of Active Resources and Open Resources.

A controller processing a command that requests a zone to transition to the ZSIO:Implicitly Opened state, the ZSEO:Explicitly Opened state, or the ZSC:Closed state shall, if resources are not available and:

- a) the Maximum Active Resources field is greater than the Maximum Open Resources field and:
 - i. as a result of the requested transition the resource that is not available is Active Resources, then abort the command with a status code of Too Many Active Zones; or
 - ii. as a result of the requested transition the resource that is not available is Open Resources, and:
 - a. one or more zones are in the ZSIO:Implicitly Opened state, then the controller shall select one of the zones in the ZSIO:Implicitly Opened state, transition that zone to the ZSC:Closed state, and proceed to process the command; or
 - b. no zones are in the ZSIO:Implicitly Opened state, then the controller shall abort the command with a status code of Too Many Open Zones;
- or
- b) the Maximum Active Resources field is equal to the Maximum Open Resources field and as a result of the requested transition the resource that is not available is Active Resources, then abort the command with a status code of Too Many Active Zones.

Zones that have associated Active Resources are transitioned to the ZSF:Full state when the zoned namespace becomes write protected. Refer to the Namespace Write Protection section in the NVM Express Base Specification.

The coordination of host software usage of resources associated with shared zoned namespaces is outside the scope of this specification.

2.1.1.5 Logical Block Allocation and Capacity Management

A logical block shall be marked as allocated when that logical block is written with:

- a) a Write command;
- b) a Write Uncorrectable command;
- c) a Write Zeroes command that does not deallocate the logical block (refer to the Deallocated or Unwritten Logical Blocks section in the NVM Command Set Specification);
- d) a Copy command; and
- e) a Zone Append command.

A logical block may be marked as allocated as the result of:

- a) a Write command not addressing that logical block (e.g., a write of LBA n causes allocation of LBAs n and n+1);
- b) a Write Uncorrectable command not addressing that logical block;
- c) a Write Zeroes command not addressing that logical block (refer to the Write Zeroes command in the NVM Command Set Specification);
- d) a Copy command not addressing that logical block;
- e) a Zone Management Send command that transitions a zone containing that logical block to the ZSEO:Explicitly Opened state or the ZSC:Closed state; and
- f) a Zone Append command.

A logical block may be marked deallocated as the result of:

- a) a Dataset Management command;
- b) a Write Zeroes command addressing that deallocates logical blocks (refer to the Deallocated or Unwritten Logical Blocks section in the NVM Command Set Specification); and
- c) a sanitize operation.

All logical blocks in a zone shall be marked as deallocated when the zone is in the ZSE:Empty state.

2.1.2 Command Ordering Requirements

Commands submitted to a zoned namespace are processed as specified in the NVM Express Base Specification and in the NVM Command Set Specification. As described in the NVM Command Set Specification, the controller is not responsible for checking the LBA of an Address-specific Write command to ensure any type of ordering between commands. For example, if an Address-specific Write command is submitted for LBA x and an Address-specific Write command is also submitted for LBA x

+ 1, there is no guarantee of the order of processing for those commands. If there are ordering requirements between these commands (e.g., zone type), host software is required to enforce that ordering during the submission of each of those commands.

2.1.3 Fused Operation

No fused operations are supported for this specification (e.g., the fused operations defined in the NVM Command Set Specification are not supported for this specification).

2.1.4 Atomic Operation

The AWUN, NAWUN, NABSN, AWUPF, NAWUPF, NABSPF atomicity parameters apply as defined in the Atomic Operations section in the NVM Command Set Specification.

2.1.5 End-to-end Protection Information

End-to-end protection information operates as defined in the NVM Command Set Specification, with requirements for the Zone Append command defined in section 3.4.1.1.

2.1.6 Metadata Region (MR)

The Metadata Region is as defined in the NVM Command Set Specification.

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 NVM Express Base Specification and in the NVM Command Set Specification. Additionally, Figure 9 defines Zoned Namespace Command Set specific commands that are mandatory, optional, and prohibited for an I/O controller that supports the Zoned Namespace Command Set Specification.

Figure 9: I/O Controller – Zoned Namespace Command Set Support

Command	Command Support Requirements ¹
Zone Management Send	M
Zone Management Receive	M
Zone Append	O
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 NVM Express Base Specification and in the NVM Command Set Specification. Additionally, Figure 10 defines Zoned Namespace Command Set specific log pages that are mandatory, optional, and prohibited for an I/O controller that supports the Zoned Namespace Command Set Specification.

Figure 10: I/O Controller - Zoned Namespace Command Set Specific Log Page Support

Log Page	Log Page Support Requirements ¹
Changed Zone List	O
Notes:	
1. O = Optional, M = Mandatory, P = Prohibited	

2.2.3 Features Support

This specification implements the feature support requirements for I/O controllers defined in the NVM Express Base Specification and in the NVM Command Set Specification. Additional requirements are defined in section 4.1.3.

3 I/O Commands for the Zoned Namespace Command Set

This section defines the Zoned Namespace Command Set I/O Commands.

3.1 Submission Queue Entry and Completion Queue Entry

The Submission Queue Entry (SQE) structure and the Completion Queue Entry (CQE) structure are as defined in the NVM Command Set Specification, with changes as defined in this section.

3.1.1 Common Command Format

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

3.1.2 Command Specific Status Values

This specification supports the Command Specific status values defined in the NVM Express Base Specification and in the NVM Command Set Specification. Command Specific status values that are specific to the Zoned Namespace Command Set Specification are defined in this section.

Figure 11 defines the status values specific to the Zoned Namespace Command Set.

Figure 11: Status Code – Command Specific Status Values, Zoned Namespace Command Set

Value	Description	Commands Affected
B8h	Zone Boundary Error	Compare ¹ , Copy ¹ , Read ¹ , Verify ¹ , Write, Write Uncorrectable, Write Zeroes, Zone Append
B9h	Zone Is Full	Copy, Write, Write Uncorrectable, Write Zeroes, Zone Append
BAh	Zone Is Read Only	Copy, Write, Write Uncorrectable, Write Zeroes, Zone Append, Zone Management Send
BBh	Zone Is Offline	Compare, Copy, Read, Verify, Write, Write Uncorrectable, Write Zeroes, Zone Append, Zone Management Send
BCh	Zone Invalid Write	Copy, Write, Write Uncorrectable, Write Zeroes
BDh	Too Many Active Zones	Copy, Write, Write Uncorrectable, Write Zeroes, Zone Append, Zone Management Send
BEh	Too Many Open Zones	Copy, Write, Write Uncorrectable, Write Zeroes, Zone Append, Zone Management Send
BFh	Invalid Zone State Transition	Zone Management Send
Notes:		
1. This command is affected if the Read Across Zone Boundaries bit in the Zoned Namespace Command Set specific Identify Namespace data structure (refer to section 4.1.5.1) is cleared to '0'.		

3.2 Zoned Namespace Command Set Commands

This specification includes the commands listed in Figure 12. Section 3.3 describes the Zoned Namespace Command Set specific behavior for NVM Command Set I/O commands. Section 3.4 describes the commands defined by this specification. Commands are submitted as defined in the NVM Express Base Specification.

Figure 12: Opcodes for Zoned Namespace Command Set I/O Commands

Opcode by Field			Combined Opcode ¹	Command ²	Reference
(07)	(06:02)	(01:00)			
Standard Command	Function	Data Transfer ³			
NVM Express Base Specification I/O commands implemented by this specification					
Refer to the NVM Express Base Specification				Flush ⁴	NVM Express Base Specification
Refer to the NVM Express Base Specification				Reservation Register	NVM Express Base Specification
Refer to the NVM Express Base Specification				Reservation Report	NVM Express Base Specification

Figure 12: Opcodes for Zoned Namespace Command Set I/O Commands

Opcode by Field			Combined Opcode ¹	Command ²	Reference
(07)	(06:02)	(01:00)			
Standard Command	Function	Data Transfer ³			
Refer to the NVM Express Base Specification				Reservation Acquire	NVM Express Base Specification
Refer to the NVM Express Base Specification				Reservation Release	NVM Express Base Specification
NVM Command Set commands implemented by this specification					
Refer to the NVM Command Set Specification				Dataset Management	NVM Command Set Specification
NVM Command Set commands modified by this specification					
Refer to the NVM Command Set Specification				Write	3.3.7
Refer to the NVM Command Set Specification				Read	3.3.3
Refer to the NVM Command Set Specification				Write Uncorrectable	3.3.8
Refer to the NVM Command Set Specification				Compare	3.3.1
Refer to the NVM Command Set Specification				Write Zeroes	3.3.9
Refer to the NVM Command Set Specification				Verify	3.3.6
Refer to the NVM Command Set Specification				Copy	3.3.2
I/O commands defined in this specification					
0b	111 10b	01b	79h	Zone Management Send	3.4.3
0b	111 10b	10b	7Ah	Zone Management Receive	3.4.2
0b	111 11b	01b	7Dh	Zone Append	3.4.1
Notes: 1. Opcodes not listed are defined in the NVM Express Base Specification and in the NVM Command Set Specification. 2. All Zoned Namespace 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.					

3.3 NVM Command Set I/O Commands

The Zoned Namespace Command Set section implements the NVM Command Set I/O commands, with changes as defined in this section.

3.3.1 Compare command

The Compare command operates as defined in the NVM Command Set Specification, with the additional requirements associated with the zone type of the specified zones that the command operates on (refer to section 2.1.1.2).

3.3.1.1 Command Completion

Command Completion is as defined in the NVM Command Set Specification, with the additional Zoned Namespace Command Set Command Specific status values that are defined in Figure 13.

Figure 13: Compare – Command Specific Status Values

Value	Description
B8h	Zone Boundary Error: The command specifies logical blocks in more than one zone.
B9h	Zone Is Full: The accessed zone is in the ZSF:Full state.
BAh	Zone Is Read Only: The accessed zone is in the ZSRO:Read Only state.
BBh	Zone Is Offline: The accessed zone is in the ZSO:Offline state.
BCh	Zone Invalid Write: The write to a zone was not at the write pointer.
BDh	Too Many Active Zones: The controller does not allow additional active zones.

Figure 13: Compare – Command Specific Status Values

Value	Description
BEh	Too Many Open Zones: The controller does not allow additional open zones.

3.3.2 Copy command

The Copy command operates as defined in the NVM Command Set Specification, with the additional requirements associated with the zone type of the specified zones that the command operates on (refer to section 2.1.1.2).

3.3.2.1 Command Completion

Command Completion is as defined in the NVM Command Set Specification, with the additional Zoned Namespace Command Set Command Specific status values that are defined in Figure 14.

Figure 14: Copy – Command Specific Status Values

Value	Description
B8h	Zone Boundary Error: The command specifies a Source Range Entry that contains logical blocks in more than one zone or a destination LBA range that contains logical blocks in more than one zone.
B9h	Zone Is Full: The zone specified by the destination LBA range is in the ZSF:Full state.
BAh	Zone Is Read Only: The zone specified by the destination LBA range is in the ZSRO:Read Only state.
BBh	Zone Is Offline: A zone specified by a Source Range Entry or the zone specified by the destination LBA range is in the ZSO:Offline state.
BCh	Zone Invalid Write: The write to the zone was not at the write pointer.
BDh	Too Many Active Zones: The controller does not allow additional active zones.
BEh	Too Many Open Zones: The controller does not allow additional open zones.

3.3.3 Dataset Management command

The Dataset Management command operates as defined in the NVM Command Set Specification, with the additional requirements defined in this section.

If the specified zone is in the ZSO:Offline state then the command shall be aborted with a status code of Zone is Offline.

3.3.3.1 Command Completion

Command Completion is as defined in the NVM Command Set Specification, with the additional Zoned Namespace Command Set Command Specific status values that are defined in Figure 15.

Figure 15: Dataset Management – Command Specific Status Values

Value	Description
BBh	Zone Is Offline: A zone specified by a Source Range Entry or the zone specified by the destination LBA range is in the ZSO:Offline state.

3.3.4 Flush command

The Flush command operates as defined in the NVM Command Set Specification, with the additional requirements defined in this section.

If:

- a) a volatile write cache is present;
- b) a volatile write cache is enabled; and
- c) the specified zone is in the ZSO:Offline state,

then the command shall be aborted with a status code of Zone Is Offline.

3.3.4.1 Command Completion

Command Completion is as defined in the NVM Command Set Specification, with the additional Zoned Namespace Command Set Command Specific status values that are defined in Figure 16.

Figure 16: Flush – Command Specific Status Values

Value	Description
BBh	Zone Is Offline: A zone specified by a Source Range Entry or the zone specified by the destination LBA range is in the ZSO:Offline state.

3.3.5 Read command

The Read command operates as defined in the NVM Command Set Specification, with the additional requirements associated with the zone type of the specified zones that the command operates on (refer to section 2.1.1.2).

3.3.5.1 Command Completion

Command Completion is as defined in the NVM Command Set Specification, with the additional Zoned Namespace Command Set Command Specific status values that are defined in Figure 17.

Figure 17: Read – Command Specific Status Values

Value	Description
B8h	Zone Boundary Error: The command specifies logical blocks in more than one zone.
BBh	Zone Is Offline: The accessed zone is in the ZSO:Offline state.

3.3.6 Verify command

The Verify command operates as defined in the NVM Command Set Specification, with the additional requirements associated with the zone type of the specified zones that the command operates on (refer to section 2.1.1.2).

3.3.6.1 Command Completion

Command Completion is as defined in the NVM Command Set Specification, with the additional Zoned Namespace Command Set Command Specific status values that are defined in Figure 18.

Figure 18: Verify – Command Specific Status Values

Value	Description
B8h	Zone Boundary Error: The command specifies logical blocks in more than one zone.
BBh	Zone Is Offline: The accessed zone is in the ZSO:Offline state.

3.3.7 Write command

The Write command operates as defined in the NVM Command Set Specification, with the additional requirements associated with the zone type of the specified zones that the command operates on (refer to section 2.1.1.2).

3.3.7.1 Command Completion

Command Completion is as defined in the NVM Command Set Specification, with the additional Zoned Namespace Command Set Command Specific status values that are defined in Figure 19.

Figure 19: Write – Command Specific Status Values

Value	Description
B8h	Zone Boundary Error: The command specifies logical blocks in more than one zone.
B9h	Zone Is Full: The accessed zone is in the ZSF:Full state.
BAh	Zone Is Read Only: The accessed zone is in the ZSRO:Read Only state.
BBh	Zone Is Offline: The accessed zone is in the ZSO:Offline state.
BCh	Zone Invalid Write: The write to a zone was not at the write pointer.

Figure 19: Write – Command Specific Status Values

Value	Description
BDh	Too Many Active Zones: The controller does not allow additional active zones.
BEh	Too Many Open Zones: The controller does not allow additional open zones.

3.3.8 Write Uncorrectable command

The Write Uncorrectable command operates as defined in the NVM Command Set Specification, with the additional requirements associated with the zone type of the specified zones that the command operates on (refer to section 2.1.1.2).

3.3.8.1 Command Completion

Command Completion is as defined in the NVM Command Set Specification, with the additional Zoned Namespace Command Set Command Specific status values that are defined in Figure 20.

Figure 20: Write Uncorrectable – Command Specific Status Values

Value	Description
B8h	Zone Boundary Error: The command specifies logical blocks in more than one zone.
B9h	Zone Is Full: The accessed zone is in the ZSF:Full state.
BAh	Zone Is Read Only: The accessed zone is in the ZSRO:Read Only state.
BBh	Zone Is Offline: The accessed zone is in the ZSO:Offline state.
BCh	Zone Invalid Write: The write to a zone was not at the write pointer.
BDh	Too Many Active Zones: The controller does not allow additional active zones.
BEh	Too Many Open Zones: The controller does not allow additional open zones.

3.3.9 Write Zeroes command

The Write Zeroes command operates as defined in the NVM Command Set Specification, with the additional requirements associated with the zone type of the specified zones that the command operates on (refer to section 2.1.1.2).

3.3.9.1 Command Completion

Command Completion is as defined in the NVM Command Set Specification, with the additional Zoned Namespace Command Set Command Specific status values that are defined in Figure 21.

Figure 21: Write Zeroes – Command Specific Status Values

Value	Description
B8h	Zone Boundary Error: The command specifies logical blocks in more than one zone.
B9h	Zone Is Full: The accessed zone is in the ZSF:Full state.
BAh	Zone Is Read Only: The accessed zone is in the ZSRO:Read Only state.
BBh	Zone Is Offline: The accessed zone is in the ZSO:Offline state.
BCh	Zone Invalid Write: The write to a zone was not at the write pointer.
BDh	Too Many Active Zones: The controller does not allow additional active zones.
BEh	Too Many Open Zones: The controller does not allow additional open zones.

3.4 Zoned Namespace Command Set I/O Commands

3.4.1 Zone Append command

The Zone Append command writes data and metadata, if applicable, to the I/O controller for the zone indicated by the ZSLBA field. The controller assigns the data and metadata, if applicable, to a set of logical blocks within the zone. The lowest LBA of the set of logical blocks written is returned in the completion queue entry (refer to section 3.4.1.2). The host may also specify protection information to include as part of the operation.

This 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.

Write ordering in the case of multiple outstanding Zone Append commands to a zone is undefined and left to the controller.

If the zone specified by the Zone Append command is not a Sequential Write Required zone, then the command shall be aborted with a status code of Invalid Field in Command.

If the ZSLBA field in the Zone Append command does not specify the lowest logical block for a zone, then the command shall be aborted with a status code of Invalid Field in Command.

The AWUN, NAWUN, NABSN, AWUPF, NAWUPF, NABSPF atomicity parameters apply as defined in the Atomic Operations section in the NVM Command Set Specification to the Zone Append command.

Figure 22: Zone Append – Metadata Pointer

Bits	Description
63:00	Metadata Pointer (MPTR): This field contains the Metadata Pointer, if applicable. Refer to the Common Command Format figure in the NVM Express Base Specification for the definition of this field.

Figure 23: Zone Append – Data Pointer

Bits	Description
127:00	Data Pointer (DPTR): This field specifies the location of a data buffer where data is transferred from. Refer to the Common Command Format figure in the NVM Express Base Specification for the definition of this field.

Figure 24: Zone Append – Command Dword 2 and Dword 3

Bits	Description
63:48	Reserved
47:00	This field and Command Dword 14 specify the variable sized Logical Block Storage Tag (LBST) and Initial Logical Block Reference Tag (ILBRT) fields, which are defined in the NVM Express Base Specification. If the namespace is not formatted to use end-to-end protection information, then this field is ignored.

Figure 25: Zone Append – Command Dword 10 and Command Dword 11

Bits	Description
63:00	Zone Start Logical Block Address (ZSLBA): This field indicates the 64-bit address of the lowest logical block of the zone in which the data and metadata, if applicable, associated with this command is to be stored. Command Dword 10 contains bits 31:00; Command Dword 11 contains bits 63:32.

Figure 26: Zone Append – Command Dword 12

Bits	Description
31	Limited Retry (LR): If set to '1', the controller should apply limited retry efforts. If cleared to '0', the controller should apply all available error recovery means to write the data to the NVM, as defined in the NVM Command Set Specification.
30	Force Unit Access (FUA): If set to '1', then for data and metadata, if any, associated with logical blocks specified by the Zone Append command, the controller shall write that data and metadata, if any, to non-volatile media before indicating command completion. There is no implied ordering with other commands. If cleared to '0', then this bit has no effect, as defined in the NVM Command Set Specification.
29:26	Protection Information (PRINFO): Specifies the Protection Information field, as defined in the NVM Command Set Specification.

Figure 26: Zone Append – Command Dword 12

Bits	Description
25	Protection Information Remap (PIREMAP): This bit determines the contents of the reference tag written to the media (refer to section 3.4.1.1). If the specified zoned namespace is formatted for Type 1 protection, the controller shall abort the command with a status of Invalid Protection Information if this bit is cleared to '0'. If the specified zoned namespace is formatted for Type 3 protection, the controller shall abort the command with a status of Invalid Protection Information if this bit is set to '1'.
24	Storage Tag Check (STC): This bit specifies that the Storage Tag field shall be checked as part of end-to-end data protection processing as defined in the NVM Command Set Specification.
23:20	Directive Type (DTYPE): Specifies the Directive Type associated with the Directive Specific field (refer to the NVM Express Base Specification).
19:16	Reserved
15:00	Number of Logical Blocks (NLB): This field indicates the number of logical blocks to be written. This is a 0's based value.

Figure 27: Zone Append – Command Dword 13

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

Figure 28: Zone Append – Command Dword 14

Bits	Description
31:00	This field and bits 47:00 of Command Dword 2 and Dword 3 specify the variable sized Logical Block Storage Tag (LBST) and Initial Logical Block Reference Tag (ILBRT) fields, which are defined in the NVM Command Set Specification. If the namespace is not formatted to use end-to-end protection information, then this field is ignored.

Figure 29: Zone Append – Command Dword 15

Bits	Description
31:16	Logical Block Application Tag Mask (LBATM): This field specifies the Application Tag Mask value. This field is ignored if the zoned namespace is not formatted to use end-to-end protection information. Refer to the End-to-end Data Protection section in the NVM Command Set Specification.
15:00	Logical Block Application Tag (LBAT): This field specifies the Application Tag value. This field is ignored if the zoned namespace is not formatted to use end-to-end protection information. Refer to the End-to-end Data Protection section in the NVM Command Set Specification.

3.4.1.1 Protection Information

For the Zone Append command, the actual LBA where data is written by the command is not known to the host until the Zone Append command completes (refer to the ALBA field defined in section 3.4.1.2). The host is not able to provide the actual LBA of the data in the reference tags in the Protection Information at the time the Zone Append command is issued (refer to the End-to-end Data Protection section of the NVM Command Set Specification). As a result, handling of LBA based reference tags in the Protection Information is handled as defined in this section.

Unless otherwise specified, for the protection information received by the controller from the host, the checks performed by the controller shall be performed as defined in the Control of Protection Information Checking – PRCHK section of the NVM Command Set Specification.

For the Reference Tag written to the media, if the PIREMAP bit is:

- a) cleared to '0', then the controller shall write the Reference Tag to the media per the End-to-end Protection Information section of the NVM Command Set Specification without modification; or
- b) set to '1', then the controller shall write the Logical Block Reference Tag to media for the first and subsequent LBAs as follows:
 - A) Media Reference Tag[0] = ILBRT + (ALBA - ZSLBA); and
 - B) Media Reference Tag[n+1] = Media Reference Tag[n] + 1

Where: Media Reference Tag [0] is the Logical Block Reference Tag written to media for the first LBA, and Media Reference Tag [n+1] is the Logical Block Reference Tag written to media for all other LBAs.

Note: When writing protection information using the Zone Append command with the PIREMAP bit set to '1', the Logical Block Reference Tag written to media is not the value transferred from the host to the controller, but contains the value calculated as defined in this section.

The controller shall write all protection information received from the host, other than the Logical Block Reference Tag, to the media without modification.

3.4.1.2 Command Completion

When the command is completed, the controller shall post a completion queue entry (CQE) to the associated I/O Completion Queue indicating the status for the command. If the command is successfully completed, then the lowest LBA containing the data written by the command is returned in the Assigned LBA (ALBA) field of the completion queue entry. Dword 0 contains bits 31:00 of the ALBA field; Dword 1 contains bits 63:32 of the ALBA field. If the command is not successfully completed, then the contents of the ALBA field are undefined.

Figure 30 defines the Zone Append command specific status values.

Figure 30: Zone Append – Command Specific Status Values

Value	Description
81h	Invalid Protection Information: The Protection Information (PRINFO) field settings specified in the command are invalid for the Protection Information with which the namespace was formatted or the ILBRT field is invalid (refer to section 3.4.1).
B8h	Zone Boundary Error: The command specifies logical blocks in more than one zone.
B9h	Zone Is Full: The accessed zone is in the ZSF:Full state.
BAh	Zone Is Read Only: The accessed zone is in the ZSRO:Read Only state.
BBh	Zone Is Offline: The accessed zone is in the ZSO:Offline state.
BDh	Too Many Active Zones: The controller does not allow additional active zones.
BEh	Too Many Open Zones: The controller does not allow additional open zones.

3.4.2 Zone Management Receive command

The Zone Management Receive command returns a data buffer that contains information about zones. That information includes characteristics of the zone, the state of the zone, the capacity of the zone, and other information defined in section 3.4.2.2. The host uses this command to determine the current settings for this information.

If this information changes (e.g., as indicated by the Capacity Changed bit set to '1' in the completion queue entry of a Zone Management Send command, or by a Zone Descriptor Changed event), then the host may use this command to determine the current state of this information (e.g., the current capacity of the zone or the Reset Zone Recommended attribute).

The Zone Management Receive command uses the Data Pointer, Command Dword 10, Command Dword 11, Command Dword 12, and Command Dword 13 fields. All other command specific fields are reserved.

Figure 31: Zone Management Receive – Data Pointer

Bits	Description
127:00	Data Pointer (DPTR): This field specifies the location of a data buffer where data is transferred from. Refer to the NVM Express Base Specification for the definition of this field.

Figure 32: Zone Management Receive – Command Dword 10 and Command Dword 11

Bits	Description
63:00	Starting LBA (SLBA): This field specifies an LBA in the lowest numbered zone that the Zone Receive Action operates on. Command Dword 10 contains bits 31:00; Command Dword 11 contains bits 63:32.

Figure 33: Zone Management Receive – Command Dword 12

Bits	Description
31:00	Number of Dwords: This field specifies the number of dwords to return. If host software specifies a size larger than what the Zone Receive Action data structure returns, then the controller returns the complete result with undefined results for dwords beyond the end of the data of the Zone Receive Action data structure. This is a 0's based value.

Figure 34: Zone Management Receive – Command Dword 13

[illegible]

Figure 34: Zone Management Receive – Command Dword 13

Bits	Description
00h	Report Zones: Reports Zone Descriptor entries through the Report Zones data structure (refer to Figure 35).
01h	Extended Report Zones: Reports Zone Descriptor entries through the Extended Report Zones data structure (refer to Figure 36). This value is supported if the zoned namespace is formatted with a non-zero Zone Descriptor Extension Size. Otherwise, the controller shall abort the command with a status code of Invalid Field in Command.
02h to FFh	Reserved

3.4.2.1 Zone Receive Actions

The Zone Management Receive command Zone Receive Action field specifies what action to perform.

3.4.2.1.1 Report Zones

The Report Zones action returns the Report Zones data structure (refer to Figure 35). The Zone Descriptors of the Report Zones data structure shall:

- report only Zone Descriptors of zones for which the ZSLBA value is greater than or equal to the ZSLBA value of the zone specified by the SLBA value in the command;
- match the criteria in the Zone Receive Action Specific field; and
- be sorted in ascending order by the ZSLBA value of each zone.

3.4.2.1.2 Extended Report Zones

The Extended Report Zones action returns the Extended Report Zones data structure (refer to Figure 36). The Zone Descriptors and Zone Descriptor Extensions of the Extended Report Zones data structure shall:

- report only Zone Descriptors and Zone Descriptor Extensions of zones for which the ZSLBA value of is greater than or equal to the ZSLBA value of the zone specified by the SLBA value in the command;
- match the criteria in the Zone Receive Action Specific field; and
- be sorted in ascending order by the ZSLBA value of each zone.

3.4.2.2 Zone Management Receive Data Structures

3.4.2.2.1 Report Zones Data Structure

Figure 35 defines the Report Zones Data Structure.

Figure 35: Report Zones Data Structure

Bytes	Description
07:00	Number of Zones: If the Partial Report bit (refer to Figure 34) is cleared to '0', then this field indicates the number of zones that match the criteria defined in section 3.4.2.1.1. If the Partial Report bit is set to '1', then this field indicates the number of zones for which complete Zone Descriptors were transferred to the data buffer. Refer to section 3.4.2.1.1 for the content of the data buffer.
63:08	Reserved
127:64	Zone Descriptor 0: Contains the Zone Descriptor for the first zone reported, if any (refer to Figure 37).
191:128	Zone Descriptor 1: Contains the Zone Descriptor for the second zone reported, if any.
...	...
((n+1)*64)+63:(n+1)*64	Zone Descriptor n: Contains the Zone Descriptor for the last zone reported, if any.

3.4.2.2.2 Extended Report Zones Data Structure

Figure 36 defines the Extended Report Zones Data Structure.

Figure 36: Extended Report Zones Data Structure

Bytes ¹	Description
07:00	<p>Number of Zones: If the Partial Report bit (refer to Figure 34) is cleared to '0', then this field indicates the number of zones that match the criteria defined in section 3.4.2.1.2.</p> <p>If the Partial Report bit is set to '1', then this field indicates the number of zones for which complete Zone Descriptors and complete Zone Descriptor Extensions were transferred to the data buffer.</p> <p>Refer to section 3.4.2.1.2 for the content of the data buffer.</p>
63:08	Reserved
127:64	Zone Descriptor 0: Contains the Zone Descriptor for the first zone reported, if any (refer to Figure 37).
(ZDES+127):128	Zone Descriptor Extension 0: Contains the Zone Descriptor Extension for the first zone reported, if any. If no Zone Descriptor Extension is associated with the zone, then the contents of this field are undefined.
(ZDES+191):(ZDES+128)	Zone Descriptor 1: Contains the Zone Descriptor for the second zone reported, if any.
((2*ZDES)+191):(ZDES+192)	Zone Descriptor Extension 1: Contains the Zone Descriptor Extension for the second zone reported, if any.
...	...
((n*ZDES)+((n+2)*64)-1): ((n*ZDES)+((n+1)*64))	Zone Descriptor n: Contains the Zone Descriptor for the last zone reported, if any.
((((n+1)*ZDES)+((n+2)*64)-1): ((n*ZDES)+((n+2)*64))	Zone Descriptor Extension n: Contains the Zone Descriptor Extension for the last zone reported, if any.
Notes:	
1. ZDES corresponds to the formatted Zone Descriptor Extension Size field in bytes (i.e. value multiplied by 64) (refer to Figure 49).	

3.4.2.2.3 Zone Descriptor Data Structure

Figure 37 defines the Zone Descriptor data structure.

Figure 37: Zone Descriptor Data Structure

Bytes	Description			
00		Bits	Description	
		7:4	Reserved	
		3:0	Zone Type (ZT): This field indicates the type of the zone.	
			Value	Definition
			2h	Sequential Write Required
All other values	Reserved			
01		Bits	Description	
		7:4	Zone State (ZS): This field indicates the state of the zone.	
			Value	Definition
			0h	Reserved
			1h	Empty
			2h	Implicitly Opened
			3h	Explicitly Opened
			4h	Closed
			5h to Ch	Reserved
			Dh	Read Only
			Eh	Full
			Fh	Offline
			3:0	Reserved

Figure 37: Zone Descriptor Data Structure

Bytes	Description												
02	Zone Attributes (ZA): Indicates attributes for the Zone: <table><tr><th>Bits</th><th>Description</th></tr><tr><td>7</td><td>Zone Descriptor Extension Valid (ZDEV): If this bit is set to '1', then Zone Descriptor Extension data is associated with the zone. If this bit is cleared to '0', then no Zone Descriptor Extension data is associated with the zone. Refer to section 5.1.</td></tr><tr><td>6:3</td><td>Reserved</td></tr><tr><td>2</td><td>Reset Zone Recommended (RZR): If this bit is set to '1', then the controller recommends that this zone be reset. Refer to section 5.4.</td></tr><tr><td>1</td><td>Finish Zone Recommended (FZR): If this bit is set to '1', then the controller recommends that this zone be finished. Refer to section 5.5.</td></tr><tr><td>0</td><td>Zone Finished by Controller (ZFC): If this bit is set to '1', then the controller finished this zone due to a Zone Active Excursion. Refer to section 5.6.</td></tr></table>	Bits	Description	7	Zone Descriptor Extension Valid (ZDEV): If this bit is set to '1', then Zone Descriptor Extension data is associated with the zone. If this bit is cleared to '0', then no Zone Descriptor Extension data is associated with the zone. Refer to section 5.1.	6:3	Reserved	2	Reset Zone Recommended (RZR): If this bit is set to '1', then the controller recommends that this zone be reset. Refer to section 5.4.	1	Finish Zone Recommended (FZR): If this bit is set to '1', then the controller recommends that this zone be finished. Refer to section 5.5.	0	Zone Finished by Controller (ZFC): If this bit is set to '1', then the controller finished this zone due to a Zone Active Excursion. Refer to section 5.6.
	Bits	Description											
	7	Zone Descriptor Extension Valid (ZDEV): If this bit is set to '1', then Zone Descriptor Extension data is associated with the zone. If this bit is cleared to '0', then no Zone Descriptor Extension data is associated with the zone. Refer to section 5.1.											
	6:3	Reserved											
	2	Reset Zone Recommended (RZR): If this bit is set to '1', then the controller recommends that this zone be reset. Refer to section 5.4.											
	1	Finish Zone Recommended (FZR): If this bit is set to '1', then the controller recommends that this zone be finished. Refer to section 5.5.											
0	Zone Finished by Controller (ZFC): If this bit is set to '1', then the controller finished this zone due to a Zone Active Excursion. Refer to section 5.6.												
03	Zone Attributes Information (ZAI): Indicates additional information associated with attributes for the Zone: <table><tr><th>Bits</th><th>Description</th></tr><tr><td>7:4</td><td>Reserved</td></tr><tr><td>3:2</td><td>Reset Zone Recommended Time Limit (RZRTL): If the Reset Zone Recommended bit is set to '1', then the value in this field selects a field in the I/O Command Set specific Identify Namespace data structure for the Zoned Namespace Command Set that indicates amount of time before the NVM subsystem may perform a vendor specific action on a zone after the Reset Zone Recommended zone attribute is set to '1' for that zone:<ul style="list-style-type: none">• A 00b value indicates the Reset Recommended Limit (RRL) field;• A 01b value indicates the Reset Recommended Limit 1 (RRL1) field;• A 10b value indicates the Reset Recommended Limit 2 (RRL2) field; and• A 11b value indicates the Reset Recommended Limit 3 (RRL3) field.</td></tr><tr><td>1:0</td><td>Finish Zone Recommended Time Limit (FZRTL): If the Finish Zone Recommended bit is set to '1', then the value in this field selects a field in the I/O Command Set specific Identify Namespace data structure for the Zoned Namespace Command Set that indicates the amount of time before the NVM subsystem may perform a vendor specific action on a zone after the Finish Zone Recommended zone attribute is set to '1' for that zone:<ul style="list-style-type: none">• A 00b value indicates the Finish Recommended Limit (FRL) field;• A 01b value indicates the Finish Recommended Limit 1 (FRL1) field;• A 10b value indicates the Finish Recommended Limit 2 (FRL2) field; and• A 11b value indicates the Finish Recommended Limit 3 (FRL3) field.</td></tr></table>	Bits	Description	7:4	Reserved	3:2	Reset Zone Recommended Time Limit (RZRTL): If the Reset Zone Recommended bit is set to '1', then the value in this field selects a field in the I/O Command Set specific Identify Namespace data structure for the Zoned Namespace Command Set that indicates amount of time before the NVM subsystem may perform a vendor specific action on a zone after the Reset Zone Recommended zone attribute is set to '1' for that zone: <ul style="list-style-type: none">• A 00b value indicates the Reset Recommended Limit (RRL) field;• A 01b value indicates the Reset Recommended Limit 1 (RRL1) field;• A 10b value indicates the Reset Recommended Limit 2 (RRL2) field; and• A 11b value indicates the Reset Recommended Limit 3 (RRL3) field.	1:0	Finish Zone Recommended Time Limit (FZRTL): If the Finish Zone Recommended bit is set to '1', then the value in this field selects a field in the I/O Command Set specific Identify Namespace data structure for the Zoned Namespace Command Set that indicates the amount of time before the NVM subsystem may perform a vendor specific action on a zone after the Finish Zone Recommended zone attribute is set to '1' for that zone: <ul style="list-style-type: none">• A 00b value indicates the Finish Recommended Limit (FRL) field;• A 01b value indicates the Finish Recommended Limit 1 (FRL1) field;• A 10b value indicates the Finish Recommended Limit 2 (FRL2) field; and• A 11b value indicates the Finish Recommended Limit 3 (FRL3) field.				
	Bits	Description											
	7:4	Reserved											
	3:2	Reset Zone Recommended Time Limit (RZRTL): If the Reset Zone Recommended bit is set to '1', then the value in this field selects a field in the I/O Command Set specific Identify Namespace data structure for the Zoned Namespace Command Set that indicates amount of time before the NVM subsystem may perform a vendor specific action on a zone after the Reset Zone Recommended zone attribute is set to '1' for that zone: <ul style="list-style-type: none">• A 00b value indicates the Reset Recommended Limit (RRL) field;• A 01b value indicates the Reset Recommended Limit 1 (RRL1) field;• A 10b value indicates the Reset Recommended Limit 2 (RRL2) field; and• A 11b value indicates the Reset Recommended Limit 3 (RRL3) field.											
1:0	Finish Zone Recommended Time Limit (FZRTL): If the Finish Zone Recommended bit is set to '1', then the value in this field selects a field in the I/O Command Set specific Identify Namespace data structure for the Zoned Namespace Command Set that indicates the amount of time before the NVM subsystem may perform a vendor specific action on a zone after the Finish Zone Recommended zone attribute is set to '1' for that zone: <ul style="list-style-type: none">• A 00b value indicates the Finish Recommended Limit (FRL) field;• A 01b value indicates the Finish Recommended Limit 1 (FRL1) field;• A 10b value indicates the Finish Recommended Limit 2 (FRL2) field; and• A 11b value indicates the Finish Recommended Limit 3 (FRL3) field.												
07:04	Reserved												
15:08	Zone Capacity (ZCAP): This field contains the maximum number of logical blocks that are available to be written with user data when the zone is in the ZSE:Empty state. This value shall be less than or equal to the Zone Size field (refer to Figure 49). If the Variable Zone Capacity bit is cleared to '0' in the Zone Operation Characteristics field in the Zoned Namespace Command Set specific Identify Namespace data structure (refer to section 4.1.5.1), then this field does not change without a change to the format of the zoned namespace. If the Variable Zone Capacity bit is set to '1' in the Zone Operation Characteristics field in the Zoned Namespace Command Set specific Identify Namespace data structure, then the zone capacity may change upon successful completion of a Zone Management Send command specifying the Zone Send Action of Reset Zone (refer to section 3.4.3.1.4).												
	23:16	Zone Start Logical Block Address (ZSLBA): This field contains the 64-bit address of the lowest logical block for the zone.											
	31:24	Write Pointer (WP): This field is the logical block address where the next write operation for this zone should be issued. Refer to section 2.1.1.2.1 for the behavior of the write pointer.											
63:32	Reserved												

3.4.2.3 Command Completion

When the command is completed, the controller shall post a completion queue entry (CQE) to the associated I/O Completion Queue indicating the status for the command.

3.4.3 Zone Management Send command

The Zone Management Send command requests an action on one or more zones. The command uses the Data Pointer, Command Dword 10, Command Dword 11, and Command Dword 13 fields. All other command specific fields are reserved.

Figure 38: Zone Management Send – Command Dword 10 and Command Dword 11

Bits	Description
63:00	Starting LBA (SLBA): This field specifies the lowest LBA of the zone on which the Zone Send Action is performed. Command Dword 10 contains bits 31:00 of the SLBA; Command Dword 11 contains bits 63:32 of the SLBA.

Figure 39: Zone Management Send – Command Dword 13

Bits	Description																													
31:09	Reserved																													
08	Select All: If this bit is set to ‘1’, then the SLBA field is ignored. If this bit is cleared to ‘0’, then the SLBA field specifies the lowest logical block of the zone. Refer to section 3.4.3.1 for specific behavior for each Zone Send Action.																													
07:00	Zone Send Action (ZSA): Defines the zone action to be performed for Zone Management Send.																													
	Value	Description	Refer to section	00h	Reserved		01h	Close Zone: Close one or more zones.	3.4.3.1.1	02h	Finish Zone: Finish one or more zones.	3.4.3.1.2	03h	Open Zone: Open one or more zones.	3.4.3.1.3	04h	Reset Zone: Reset one or more zones.	3.4.3.1.4	05h	Offline Zone: Offline one or more zones.	3.4.3.1.5	06h to 0Fh	Reserved		10h	Set Zone Descriptor Extension: Attach Zone Descriptor Extension data to a zone.	3.4.3.1.6	11h to FFh	Reserved	
	Value	Description	Refer to section																											
	00h	Reserved																												
	01h	Close Zone: Close one or more zones.	3.4.3.1.1																											
	02h	Finish Zone: Finish one or more zones.	3.4.3.1.2																											
	03h	Open Zone: Open one or more zones.	3.4.3.1.3																											
	04h	Reset Zone: Reset one or more zones.	3.4.3.1.4																											
	05h	Offline Zone: Offline one or more zones.	3.4.3.1.5																											
	06h to 0Fh	Reserved																												
	10h	Set Zone Descriptor Extension: Attach Zone Descriptor Extension data to a zone.	3.4.3.1.6																											
11h to FFh	Reserved																													

If the command completes successfully, depending on the Zone Send Action field and the current states of the zones specified by the command, then that command may affect zones in various ways, including the following:

- a) the zone state may change;
- b) the Zone Descriptor Extension data may change; and
- c) the Zone Descriptor Extension Valid zone attribute bit may change.

If the controller has multiple outstanding Zone Management Send commands that specify one or more of the same zones, then the results are undefined.

If the zoned namespace containing the specified zone is write protected as described in the Namespace Write Protection section in the NVM Express Base Specification, then the controller shall abort the command with a status code of Namespace is Write Protected.

If there are insufficient available Active Resources or insufficient available Open Resources, then the command shall be aborted as defined in section 2.1.1.4, and no zone state transition shall occur.

If the command SLBA field does not specify the starting logical block for a zone in the specified zoned namespace and the Select All bit is cleared to '0', then the controller shall abort the command with a status code of Invalid Field in Command.

If the Zone Send Action field specifies the Set Zone Descriptor Extension Zone Send Action and the Zone Descriptor Extension Size field value in the Zoned Namespace Command Set specific Identify Namespace data structure is cleared to 0h, then the controller shall abort the command with a status code of Invalid Field in Command.

For a shared zoned namespace, the method used by hosts to coordinate Zone Management Send commands is outside the scope of this specification.

3.4.3.1 Zone Send Actions

The Zone Management Send command Zone Send Action field defines what action to perform on one or more zones.

3.4.3.1.1 Close Zone

If the Select All bit in Command Dword 13 is cleared to '0', and the zone specified by the SLBA field is in the:

- a) ZSIO:Implicitly Opened state or the ZSEO:Explicitly Opened state, then the zone shall be transitioned to the ZSC:Closed state;
- b) ZSC:Closed state, then no change shall be made to the zone state; and
- c) ZSE:Empty state, the ZSF:Full state, the ZSRO:Read Only state, or the ZSO:Offline state, then the controller shall abort the command with a status code of Invalid Zone State Transition.

If the Select All bit is set to '1', then the SLBA field shall be ignored and all zones that are in the:

- a) ZSIO:Implicitly Opened state; and
- b) ZSEO:Explicitly Opened state,

shall be transitioned to the ZSC:Closed state.

3.4.3.1.2 Finish Zone

If the Select All bit in Command Dword 13 is cleared to '0', and the zone specified by the SLBA field is in the:

- a) ZSE:Empty state, the ZSIO:Implicitly Opened state, the ZSEO:Explicitly Opened state, or the ZSC:Closed state, then the zone shall be transitioned to the ZSF:Full state;
- b) ZSF:Full state, then no change shall be made to the zone state; and
- c) ZSRO:Read Only state or the ZSO:Offline state, then the controller shall abort the command with a status code of Invalid Zone State Transition.

If the Select All bit is set to '1', then the SLBA field shall be ignored and all zones that are in the:

- a) ZSIO:Implicitly Opened state;
- b) ZSEO:Explicitly Opened state; and
- c) ZSC:Closed state,

shall be transitioned to the ZSF:Full state.

3.4.3.1.3 Open Zone

If the Select All bit in Command Dword 13 is cleared to '0', and the zone specified by the SLBA field is in the:

- a) ZSE:Empty state, the ZSIO:Implicitly Opened state or the ZSC:Closed state, then the zone should be transitioned to the ZSEO:Explicitly Opened state;
- b) ZSEO:Explicitly Opened state, then no change shall be made to the zone state; and
- c) ZSF:Full state, the ZSRO:Read Only state, or the ZSO:Offline state, then the controller shall abort the command with a status code of Invalid Zone State Transition.

If the Select All bit is set to '1', then the SLBA field shall be ignored and all zones that are in the ZSC:Closed state should be transitioned to the ZSEO:Explicitly Opened state.

If there are insufficient available Active Resources or insufficient available Open Resources, then the command shall be aborted as described in section 2.1.1.4 and no zone state transition shall occur.

3.4.3.1.4 Reset Zone

If the Select All bit in Command Dword 13 is cleared to '0', and the zone specified by the SLBA field is in the:

- a) ZSIO:Implicitly Opened state, the ZSEO:Explicitly Opened state, the ZSC:Closed state, or the ZSF:Full state, then the specified zone shall be transitioned to the ZSE:Empty state;

- b) ZSE:Empty state, then no change shall be made to the zone state; and
- c) ZSRO:Read Only state, or the ZSO:Offline state, then the controller shall abort the command with a status code of Invalid Zone State Transition.

If the Select All bit is set to '1', then the SLBA field shall be ignored and all zones that are in the:

- a) ZSIO:Implicitly Opened state;
- b) ZSEO:Explicitly Opened state;
- c) ZSC:Closed state; and
- d) the ZSF:Full state,

shall be transitioned to the ZSE:Empty state.

If the command completes successfully, then for each affected zone:

- a) the Write Pointer zone attribute in the Zone Descriptor shall be set to the ZSLBA of the zone; and
- b) the following zone attribute bits in the Zone Descriptor shall be cleared to '0':
 - a) Zone Descriptor Extension Valid;
 - b) Finish Zone Recommended;
 - c) Reset Zone Recommended; and
 - d) Zone Finished by Controller.

If the Variable Zone Capacity bit is set to '1' in the Zone Operation Characteristics field in the Zoned Namespace Command Set specific Identify Namespace data structure (refer to Figure 48), then the controller may change the Zone Capacity field in the Zone Descriptor of each affected zone. If the Zone Capacity field is changed for one or more zones, then the Zone Capacity Changed bit shall be set to '1' in the completion queue entry Dword 0 (refer to section 3.4.3.2).

3.4.3.1.5 Offline Zone

If the Select All bit in Command Dword 13 is cleared to '0', and the zone specified by the SLBA field is in the:

- a) ZSRO:Read Only state, then the specified zone shall be transitioned to the ZSO:Offline state;
- b) ZSO:Offline state, then no change shall be made to the zone state; and
- c) ZSE:Empty state, the ZSIO:Implicitly Opened state, the ZSEO:Explicitly Opened state, the ZSC:Closed state, or the ZSF:Full state, then the controller shall abort the command with a status code of Invalid Zone State Transition.

If the Select All bit is set to '1', then the SLBA field shall be ignored, and all zones that are in the ZSRO:Read Only state shall be transitioned to the ZSO:Offline state.

3.4.3.1.6 Set Zone Descriptor Extension

If the Select All bit in Command Dword 13 is cleared to '0' and the zone specified by the SLBA field is in:

- a) the ZSE:Empty state, then the zone shall be transitioned to the ZSC:Closed state; and
- b) any state other than the ZSE:Empty state, then the controller shall abort the command with a status code of Invalid Zone State Transition.

If the Select All bit is set to '1', then the command shall be aborted with a status code of Invalid Field in Command.

If there are insufficient available Active Resources, the command shall be aborted as described in section 2.1.1.4 and no zone state transition shall occur.

On successful command completion, the Zone Descriptor Extension of the zone shall be set to the data in the data buffer.

3.4.3.2 Command Completion

When the command is completed, the controller shall post a completion queue entry (CQE) to the associated I/O Completion Queue indicating the status for the command. Figure 40 defines the Zone Management Send command specific status values.

Figure 40: Zone Management Send – Command Specific Status Values

Value	Description
BAh	Zone Is Read Only: Zone is in the ZSRO:Read Only state. This may have occurred during the processing of the command.
BBh	Zone Is Offline: Zone is in the ZSO:Offline state. This may have occurred during the processing of the command.
BDh	Too Many Active Zones: The controller does not allow additional active zones.
BEh	Too Many Open Zones: The controller does not allow additional open zones.
BFh	Invalid Zone State Transition: The request is not a valid zone state transition.

Dword 0 of the completion queue entry indicates if the zone capacity of the specified zone has been changed. The definition of Dword 0 of the completion queue entry is in Figure 41.

Figure 41: Zone Management Send – Completion Queue Entry Dword 0

Bits	Description
31:01	Reserved
00	Zone Capacity Changed: This bit indicates if the zone capacity has changed for one or more of the zones specified by the command. If this bit is set to '1', then a zone capacity has changed due to this command. If this bit is cleared to '0', then a zone capacity has not changed due to this command.

If a zone capacity change occurred, then the host may use the Zone Management Receive command as defined in section 3.4.1 to determine what has changed.

4 Admin Commands for the Zoned Namespace Command Set

4.1 Admin Command behavior for the Zoned Namespace Command Set

The Admin Commands are as defined in the NVM Express Base Specification. The Zoned Namespace Command Set specific behavior for Admin Commands is defined in this section.

4.1.1 Asynchronous Event Request Command

The Asynchronous Event Request command operates as defined in the NVM Express Base Specification. In addition to the Asynchronous Events defined in the NVM Express Base Specification and in the NVM Command Set Specification, the Zoned Namespace Command Set defines the Asynchronous Events in this section.

Figure 42 defines the Zoned Namespace Command Set specific Asynchronous Event Information – Notice data structure.

Figure 42: Asynchronous Event Information – Notice, Zoned Namespace Command Set

Value	Description
EFh	<p>Zone Descriptor Changed: The Zone Descriptor data structure for a zone in a specific zoned namespace (refer to Figure 43) has changed. The Zone Descriptor of the zone that changed is indicated in the Changed Zone List log page. To clear this event, host software reads the Zone Changed List log page using the Get Log Page command with the Retain Asynchronous Event bit cleared to '0'.</p> <p>For a specific zone, a Zone Descriptor data structure change that occurred, for any of the following reasons, shall not generate a Zone Descriptor Changed event and shall not cause modifications to the Changed Zones List log page:</p> <ul style="list-style-type: none"> a) a Zone Management Send command that specified that zone; b) a Zone Management Send command that specified all zones; c) a write operation that transitioned that zone: <ul style="list-style-type: none"> i. from the ZSE:Empty state to the ZSIO:Implicitly Opened state; ii. from the ZSIO:Implicitly Opened state to the ZSF:Full state; iii. from the ZSEO:Explicitly Opened state to the ZSF:Full state; or iv. from the ZSC:Closed state to the ZSIO:Implicitly Opened state; d) the controller transitioning that zone to the ZSF:Full state due to an NVM Subsystem Reset; and e) the controller transitioning that zone to the ZSC:Closed state (refer to section 2.1.1.4.1).

4.1.1.1 Command Completion

A completion queue entry (CQE) is posted to the Admin Completion Queue if there is an asynchronous event to report to the host. A controller implements the same logic defined for the Asynchronous Event Request command in the NVM Express Base Specification, with the following addition:

Figure 43: Asynchronous Event Request – Completion Queue Entry Dword 1

Bytes	Description
3:0	<p>Namespace Identifier (NSID): If the Asynchronous Event Request command returned a Zoned Namespace Command Set Notice event (refer to Figure 42), then this field indicates the namespace identifier that the asynchronous event occurred on. If the Asynchronous Event Request command returned an event other than a Zoned Namespace Command Set Notice event, then this field is reserved.</p> <p>If the Asynchronous Event Request command did not complete successfully, then this field is reserved.</p>

4.1.2 Format NVM command

The Format NVM command operates as defined in the NVM Express Base Specification and in the NVM Command Set Specification. The Format Index indicates:

- a) a valid User Data Format from the LBA Format field in the NVM Command Set Identify Namespace data structure;
- b) a valid Extended LBA Format in the I/O Command Set specific Identify Namespace Data Structure for the NVM Command Set; and
- c) a valid LBA Format Extension in the I/O Command Set specific Identify Namespace data structure for the Zoned Namespace Command Set.

4.1.3 Get Features and Set Features Commands

Features support requirements for I/O controllers supporting the Zoned Namespace Command Set are as defined in the NVM Command Set Specification and in this section.

Figure 44 defines the Zoned Namespace Command Set specific Set Features Command Asynchronous Event Configuration – Command Dword 11 data structure (refer to the NVM Express Base Specification).

Figure 44: Asynchronous Event Configuration – Command Dword 11

Bit	Description
27	Zone Descriptor Changed Notices: This bit specifies whether an asynchronous event notification is sent to the host for a Zone Descriptor Changed event. If this bit is set to '1', then the Zone Descriptor Changed event is sent as defined in Figure 42 when this condition occurs. If this bit is cleared to '0', then the controller shall not send the Zone Descriptor Changed event to the host.

4.1.4 Get Log Page Command

The Get Log Page command operates as defined in the NVM Express Base Specification and in the NVM Command Set 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 status Invalid Field in Command.

In addition to the log pages defined in the NVM Express Base Specification and in the NVM Command Set Specification, the Zoned Namespace Command Set defines the log pages in this section.

Log page scope is as defined in the NVM Express Base Specification, except as modified by this specification.

The rules for namespace identifier usage are defined in the NVM Express Base Specification.

Figure 45: Log Page Identifiers

Log Identifier	Scope	Log Page Name	Reference
BFh	Namespace ¹	Changed Zone List	4.1.4.1
Key: Namespace = The log page contains information about a specific zoned namespace.			
Notes: 1. For namespace identifiers other than 0h or FFFFFFFFh.			

4.1.4.1 Changed Zone List (Log Identifier BFh)

This log page indicates if a Zone Descriptor Changed event has occurred for one or more zones. If there is an enabled Zone Descriptor Changed Event pending for a specific zone, then the Changed Zone List includes an entry for that zone.

The Changed Zone List log page is 4,096 bytes in size. This log page contains a Zone Identifier List with up to 511 Zone Identifier entries. The Changed Zone List log page is defined in Figure 46.

A Zone Identifier entry contains the ZSLBA of the zone associated with the changed Zone Descriptor. Each ZSLBA shall appear not more than once in the Zone Identifier List. Unused entries shall be zero-filled.

Figure 46: Changed Zone List Log Page

Bytes	Description
Header	
01:00	Number of Zone Identifiers: This field indicates the number of Zone Identifiers entries in the list.
07:02	Reserved
Zone Identifier List	
15:08	Zone Identifier 0: This field contains the ZSLBA of the first zone in the list, if present.
23:16	Zone Identifier 1: This field contains the ZSLBA of the second zone in the list, if present.
...	...
(N*8+15):(N*8+8)	Zone Identifier N: This field contains the ZSLBA of the N+1 zone in the list, if present.

The controller shall set the Number of Zone Identifiers field in the log page to the number of valid entries contained in the Zone Identifier List data structure.

To determine changes, the host reads the Changed Zone List log page, and for each ZSLBA specified in a Zone Identifier entry, issues a Zone Management Receive command with the Zone Receive Action set to Report Zones or Extended Report Zones, and specifying the ZSLBA from that Zone Identifier entry.

The host should read the entire page with the RAE bit cleared to '0'. This log page is dynamic and when reading with offset without the header then the data may be changed from a previous read of the page.

For a given zoned namespace, if the controller is unable to populate the Changed Zone List log page with any valid entries and there is at least one Zone Descriptor Changed event since the last time the log page was read, then the controller shall set the Number of Zone Identifiers in that log page to FFFFh, and the remainder of the list shall be zero filled. The host may read the entire Report Zones data structure or the entire Extended Report Zones data structure to discover which Zone Descriptors have changed since the last time the information was read.

The log page content should not be retained after an NVM Subsystem Reset.

4.1.5 Identify Command

This specification implements the Identify Command and associated Identify data structures defined in the NVM Express Base Specification and in the NVM Command Set Specification. Additionally, the Zoned Namespace Command Set defines the Identify data structures defined in this section.

Each I/O Command Set is assigned a specific Command Set Identifier (CSI) value by the NVM Express Base Specification. The Zoned Namespace Command Set is assigned a CSI value of 02h.

Figure 47: CNS Values

CNS Value	O/M ¹	Definition	NSID ²	CNTID ³	CSI ⁴	Reference Section
Active Namespace Management						
00h	M	Identify Namespace data structure for the specified NSID or the common namespace capabilities for the NVM Command Set. ⁵	Y	N	N ⁶	NVM Command Set Specification
01h	M	Identify Controller Data Structure, I/O Command Set Independent. ⁵	N	N	N	NVM Express Base Specification
05h	M	I/O Command Set specific Identify Namespace data structure for the specified NSID for the I/O Command Set specified in the CSI field. ⁵	Y	N	Y	CSI 00h: NVM Command Set Specification
						CSI 02h: 4.1.5.1

Figure 47: CNS Values

CNS Value	O/M ¹	Definition	NSID ²	CNTID ³	CSI ⁴	Reference Section
06h	M	I/O Command Set specific Identify Controller data structure for the controller processing the command. 5	Y	N	Y	4.1.5.2
16h	O	A Namespace Granularity List is returned to the host for up to sixteen Namespace Granularity Entries.	N	N	N ⁶	NVM Command Set Specification

Notes:

- O/M definition: O = Optional, M = Mandatory.
- The NSID field is used: Y = Yes, N = No.
- The CDW10.CNTID field is used: Y = Yes, N = No.
- The CDW11.CSI field is used: Y = Yes, N = No.
- Selection of a UUID may be supported. Refer to the UUIDs for Vendor Specific Information section of the NVM Command Set Specification.
- This Identify data structure applies to namespaces that are associated with command sets that specify logical blocks (i.e., Command Set Identifier 0h or Command Set Identifier 02h).

4.1.5.1 I/O Command Set Specific Identify Namespace Data Structure for the Zoned Namespace Command Set (CNS 05h, CSI 02h)

Figure 48 defines the I/O Command Set specific Identify Namespace data structure for the Zoned Namespace Command Set.

Figure 48: I/O Command Set Specific Identify Namespace Data Structure for the Zoned Namespace Command Set

Bytes	O/M ¹	Description				
01:00	O	Zone Operation Characteristics (ZOC): This field indicates the zone operation characteristics of the zoned namespace.				
		<table><tr><th>Bits</th><th>Description</th></tr><tr><td>15:2</td><td>Reserved</td></tr></table>	Bits	Description	15:2	Reserved
		Bits	Description			
		15:2	Reserved			
1	Zone Active Excursions: If set to '1', then a controller may transition a zone in the ZSIO:Implicitly Opened state, the ZSEO:Explicitly Opened state, or the ZSC:Closed state to the ZSF:Full state due to a vendor specific excursion event. If cleared to '0', then a controller shall not transition a zone due to a vendor specific excursion event. Refer to section 5.6.					
0	Variable Zone Capacity: if set to '1', then the capacity for a zone may change without a change to the format of the zoned namespace. If cleared to '0', then the capacity for a zone does not change without a change to the format of the zoned namespace. Refer to Figure 37.					
03:02	O	Optional Zoned Command Support (OZCS): This field defines optional features of the zoned namespace.				
		<table><tr><th>Bits</th><th>Description</th></tr><tr><td>15:1</td><td>Reserved</td></tr></table>	Bits	Description	15:1	Reserved
		Bits	Description			
15:1	Reserved					
0	Read Across Zone Boundaries: If set to '1', then any User Data Read Access Command is allowed to perform read operations that specify an LBA range containing logical blocks in more than one zone. If cleared to '0', then any command that performs a read operation that specifies an LBA range containing logical blocks in more than one zone is aborted as defined in section 2.1.1.2.1.2.					
07:04	M	Maximum Active Resources (MAR): This field defines the maximum number of active zones in the zoned namespace. A value of FFFFFFFFh indicates that there is no limit. This is a 0's based value				

Figure 48: I/O Command Set Specific Identify Namespace Data Structure for the Zoned Namespace Command Set

Bytes	O/M ¹	Description
11:08	M	Maximum Open Resources (MOR): This field defines the maximum number of open zones in the zoned namespace. This field shall be less than or equal to the Maximum Active Resources field. A value of FFFFFFFFh indicates that there is no limit. This is a 0's based value
15:12	O	Reset Recommended Limit (RRL): If the zone attribute Reset Zone Recommended Time Limit field is set to 00b, then this field indicates the number of seconds before the NVM subsystem may perform a vendor specific action on a zone after the Reset Zone Recommended zone attribute is set to '1' for that zone. If this field is cleared to 0h, then no Reset Recommended Limit is reported. Refer to section 5.4.
19:16	O	Finish Recommended Limit (FRL): If the zone attribute Finish Zone Recommended Time Limit field is set to 00b, then this field indicates the number of seconds before the NVM subsystem may perform the vendor specific action on a zone after the Finish Zone Recommended zone attribute is set to '1' for that zone. If this field is cleared to 0h, then no Finish Recommended Limit is reported. Refer to section 5.5.
23:20	O	Reset Recommended Limit 1 (RRL 1): If the zone attribute Reset Zone Recommended Time Limit field is set to 01b, then this field indicates the number of seconds before the NVM subsystem may perform a vendor specific action on a zone after the Reset Zone Recommended zone attribute is set to '1' for that zone. If this field is cleared to 0h, then no Reset Recommended Limit is reported. Refer to section 5.4.
27:24	O	Reset Recommended Limit 2 (RRL 2): If the zone attribute Reset Zone Recommended Time Limit field is set to 10b, then this field indicates the number of seconds before the NVM subsystem may perform a vendor specific action on a zone after the Reset Zone Recommended zone attribute is set to '1' for that zone. If this field is cleared to 0h, then no Reset Recommended Limit is reported. Refer to section 5.4.
31:28	O	Reset Recommended Limit 3 (RRL 3): If the zone attribute Reset Zone Recommended Time Limit field is set to 11b, then this field indicates the number of seconds before the NVM subsystem may perform a vendor specific action on a zone after the Reset Zone Recommended zone attribute is set to '1' for that zone. If this field is cleared to 0h, then no Reset Recommended Limit is reported. Refer to section 5.4.
35:32	O	Finish Recommended Limit 1 (FRL1): If the zone attribute Finish Zone Recommended Time Limit field is set to 01b, then this field indicates the number of seconds before the NVM subsystem may perform the vendor specific action on a zone after the Finish Zone Recommended zone attribute is set to '1' for that zone. If this field is cleared to 0h, then no Finish Recommended Limit is reported. Refer to section 5.5.
39:36	O	Finish Recommended Limit 2 (FRL2): If the zone attribute Finish Zone Recommended Time Limit field is set to 10b, then this field indicates the number of seconds before the NVM subsystem may perform the vendor specific action on a zone after the Finish Zone Recommended zone attribute is set to '1' for that zone. If this field is cleared to 0h, then no Finish Recommended Limit is reported. Refer to section 5.5.
43:40	O	Finish Recommended Limit 3 (FRL3): If the zone attribute Finish Zone Recommended Time Limit field is set to 11b, then this field indicates the number of seconds before the NVM subsystem may perform the vendor specific action on a zone after the Finish Zone Recommended zone attribute is set to '1' for that zone. If this field is cleared to 0h, then no Finish Recommended Limit is reported. Refer to section 5.5.
2815:44		Reserved
LBA Format Extensions List		
2831:2816	M	LBA Format 0 Extension (LBAFE0): This field indicates the LBA format Extension 0 that is supported by the controller. The LBA Format Extension field is defined in Figure 49.

Figure 48: I/O Command Set Specific Identify Namespace Data Structure for the Zoned Namespace Command Set

Bytes	O/M ¹	Description
2847:2832	O	LBA Format 1 Extension (LBAFE1): This field indicates the LBA format Extension 1 that is supported by the controller. The LBA Format Extension field is defined in Figure 49.
...		
3839:3824	O	LBA Format 63 Extension (LBAFE63): This field indicates the LBA format Extension 63 that is supported by the controller. The LBA Format Extension field is defined in Figure 49.
4095:3840	O	Vendor Specific
Notes:		
1. O/M definition: O = Optional, M = Mandatory.		

Figure 49 defines the Identify LBA Format Extension data structure.

Figure 49: LBA Format Extension Data Structure

Bits	Description
127:72	Reserved
71:64	Zone Descriptor Extension Size (ZDES): This field indicates the Zone Descriptor Extension size for each zone. The value is reported in 64B units (e.g., 1h corresponds to 64B, 2h corresponds to 128B). A value of 0h indicates that Zone Descriptor Extensions are not supported. Refer to section 5.3.
63:00	Zone Size (ZSZE): This field contains the size of each zone in the zoned namespace. The value is reported as a number of logical blocks. The value of the field shall not be cleared to 0h.

The LBA Format data structure, NVM Command Set Specific data structure is as defined in the NVM Command Set Specification. The data structure is extended through the LBA Format Extension data structure.

Commands (e.g., Format NVM and Namespace Management) that use an index to refer to the LBA Format data structure in the NVM Command Set specific Identify Namespace data structure, shall use that same index to refer to the LBA Format Extension data structure and shall process both the LBA Format data structure and LBA Format Extension data structure when an LBA Format is specified.

The host specified namespace management fields are as defined for the NVM Command Set.

4.1.5.2 I/O Command Set Specific Identify Controller Data Structure (CNS 06h, CSI 02h)

Figure 50 defines the I/O Command Set specific Identify Controller data structure for the Zoned Namespace Command Set.

Figure 50: I/O Command Set Specific Identify Controller Data Structure for the Zoned Namespace Command Set

Bytes	O/M ¹	Description
00	O ²	<p>Zone Append Size Limit (ZASL): If the Zone Append command is supported then:</p> <ul style="list-style-type: none"> a) a non-zero value in this field indicates the maximum data transfer size for the Zone Append command (refer to section 3.4.1); and b) a value of 0h in this field indicates that the maximum data transfer size for the Zone Append command is indicated by the Maximum Data Transfer Size (MDTS) field (refer to the NVM Express Base Specification). <p>The value is in units of the minimum memory page size (CAP.MPSMIN) and is reported as a power of two (2^n). This field includes the length of metadata if metadata is interleaved with the stored logical block data.</p> <p>The value of this field shall be less than or equal to the Maximum Data Transfer Size (MDTS).</p>
4095:01		Reserved

Figure 50: I/O Command Set Specific Identify Controller Data Structure for the Zoned Namespace Command Set

Bytes	O/M ¹	Description
Notes:		
1. O/M definition: O = Optional, M = Mandatory.		
2. Mandatory for controllers that support the Zone Append command.		

4.1.6 Namespace Management command

The Namespace Management command operates as defined in the NVM Command Set Specification with the addition that the Format Index (refer to section 1.4.2.2) also specifies the entry in the LBA Format Extensions list (refer to Figure 48) for zone attributes used to format a created namespace.

4.1.7 Sanitize command

The Sanitize command operates as defined in the NVM Express Base Specification, with additional requirements for controllers implementing the Zoned Namespace Command Set defined in this section.

The Sanitize Status (SSTAT) (refer to Sanitize Status (Log Identifier 81h) section in the NVM Express Base Specification) and resultant zone state as a result of a successful sanitize operation of a zoned namespace that supports the Zoned Namespace Command Set are defined in Figure 51.

All fields of the Sanitize command are as defined in the NVM Express Base Specification.

The resulting zone state and event to report as a result of a successful sanitize operation is dependent on the setting of the:

- No Deallocate After Sanitize bit in the Sanitize command that requested the sanitize operation;
- No-Deallocate Modifies Media After Sanitize field;
- No-Deallocate Inhibited bit;
- No-Deallocate Response Mode bit; and
- Sanitize Action field.

Figure 51: Sanitize Behavior for the Zoned Namespace Command Set

No Deallocate After Sanitize	No-Deallocate Modifies Media After Sanitize (NODMMAS)	No-Deallocate Inhibited (NDI)	No-Deallocate Response Mode (NODRM)	Results of a successful sanitize operation		
				Zone State ¹	Logical Block Content ²	Sanitize Status ⁴
0b	n/a ³	n/a ³	n/a ³	ZSE:Empty	Refer to section 2.1.1.2.1.2	001b
1b	00b	0b	n/a ³	ZSE:Empty	Refer to section 2.1.1.2.1.2	001b
1b	00b	1b	0b			n/a ⁵
1b	00b	1b	1b	ZSE:Empty	Refer to section 2.1.1.2.1.2	001b
1b	01b (does not modify)	0b (not inhibited)	n/a ³	ZSF:Full	Block Erase: not defined in this specification Crypto Erase: not defined in this specification Overwrite: Shall be the overwrite pattern specified in the Sanitize command that requested the sanitize operation	001b

Figure 51: Sanitize Behavior for the Zoned Namespace Command Set

No Deallocate After Sanitize	No-Deallocate Modifies Media After Sanitize (NODMMAS)	No-Deallocate Inhibited (NDI)	No-Deallocate Response Mode (NODRM)	Results of a successful sanitize operation		
				Zone State ¹	Logical Block Content ²	Sanitize Status ⁴
1b	01b	1b (inhibited)	1b	ZSE:Empty	Refer to section 2.1.1.2.1.2	100b
1b	01b	1b	0b			n/a ⁵
1b	10b (does modify)	0b	n/a ³	ZSF:Full	Block Erase: not defined in this specification Crypto Erase: not defined in this specification Overwrite: Shall be the overwrite pattern specified in the Sanitize command that requested the sanitize operation	001b
1b	10b	1b	1b	ZSE:Empty	Refer to section 2.1.1.2.1.2	100b
1b	10b	1b	0b			n/a ⁵
n/a ³	11b	n/a ³	n/a ³	Reserved		
Notes:						
1. ZSO:Offline state is a valid zone state as a result of a successful sanitize operation.						
2. This field describes the read value from a deallocated logical block. Refer to the Deallocated or Unwritten Logical Blocks section in the NVM Command Set Specification.						
3. This value is not relevant in the setup conditions defined in that row.						
4. Value reported in bits 2:0 of the Sanitize Status (SSTAT) field in the Sanitize Status (Log Identifier 81h) (refer to the NVM Express Base Specification).						
5. Sanitize command is aborted with a status code of Invalid Field in Command.						

4.2 I/O Command Set Specific Admin commands

The I/O command set specific Admin commands are as defined in the NVM Express Base Specification and in the NVM Command Set Specification.

5 Extended Capabilities

5.1 Reservations

Reservations operate as defined the NVM Express Base Specification and the NVM Command Set Specification, with the additional Command Behavior in the Presence of a Reservation defined in Figure 52.

Figure 52: 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
Zoned Namespace Read Command Group								
Zone Management Receive	A	A	C	C	A	A	C	A
Zoned Namespace Write Command Group								
Zone Append								
Zone Management Send	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 a status code of status Reservation Conflict								

5.2 Directives

Support for Directives is as defined the NVM Command Set Specification.

5.3 Zone Descriptor Extension

The Zone Descriptor Extension feature allows the host software to associate a small amount of data (e.g., UUID) to a zone. The data is attached to a zone as a Zone Descriptor Extension and can be accessed by issuing a Zone Management Receive command with a Zone Receive Action of Extended Report Zones (refer to section 3.4.2.2.2) and specifying that zone. The host software may read the Zone Descriptor Extension Valid zone attribute of that zone to identify if the zone has data associated.

A Zone Descriptor Extension is associated with a zone when the zone is transitioned from the ZSE:Empty state to the ZSC:Closed state by a Zone Management Send command with Zone Send Action of Set Zone Descriptor Extension that specifies the zone and the data to associate in the data buffer (refer to section 3.4.3.1.6).

Upon successful completion of the Zone Management Receive command:

- the data in the data buffer is associated with the Zone Descriptor Extension of that zone; and
- the Zone Descriptor Extension Valid zone attribute bit of that zone is set '1'.

When data has been associated with a Zone Descriptor Extension, the data is associated until:

- the zone transitions to the ZSE:Empty state or the ZSO:Offline state; and
- the Zone Descriptor Extension Valid zone attribute bit of that zone is cleared to '0'.

5.4 Reset Zone Recommended

A controller that schedules an internal operation (e.g., background operation on the non-volatile media) on a zone that is in the ZSF:Full state may notify host software to initiate a zone reset operation (refer

to section 3.4.3.1.4). If a controller schedules such an internal operation on a zone, the controller may notify the host by:

- a) setting the Reset Zone Recommended zone attribute of the specific zone to '1' (refer to Figure 37);
- b) setting the Reset Zone Recommended Time Limit zone attribute information to indicate the number of seconds before the controller intends to perform an internal operation on the specified zone; and
- c) generating a Zone Descriptor Changed event for the specific zone.

As a zone reset operation destroys data in a specific zone, it is optional for the host software to perform a zone reset operation on zones that have the Reset Zone Recommended zone attribute set to '1'. If the host does not perform a zone reset operation on the specific zone, then the internal operation, which may impact performance, may be performed.

If the controller has processed the internal operation or the internal operation is no longer scheduled, the controller may notify the host by:

- a) clearing the Reset Zone Recommended zone attribute of the specific zone to '0'; and
- b) generating a Zone Descriptor Changed event for the specific zone.

If a zone is in the ZSF:Read Only state or the ZSF:Offline state, then the Reset Zone Recommended attribute shall be cleared to '0'.

5.5 Finish Zone Recommended

A controller that schedules an internal operation (e.g., background operation on the non-volatile media) on a zone that is in the:

- a) ZSIO:Implicitly Opened state;
- b) ZSEO:Explicitly Opened state; or
- c) ZSC:Closed state

may notify host software to either:

- a) initiate a zone finish operation (refer to section 3.4.3.1.2) on that zone; or
- b) initiate write operations to that zone such that the zone transitions to the ZSF:Full state.

The operation initiated by the host may enable the controller to cancel the scheduled internal operation.

The controller may notify the host software by:

- a) setting the Finish Zone Recommended zone attribute of the specific zone to '1' (refer to Figure 37);
- b) setting the Finish Zone Recommended Time Limit zone attribute information to indicate the number of seconds before the controller intends to perform an internal operation on the specified zone; and
- c) generating a Zone Descriptor Changed event for the specific zone.

It is optional for the host software to process the above mitigating actions on zones that have the Finish Zone Recommended zone attribute set to '1'. If the host does not perform the mitigating actions on the specific zone, then the internal operation, which may impact performance, may be performed.

If the controller has processed the internal operation or the internal operation is no longer scheduled, the controller may notify the host by:

- a) clearing the Finish Zone Recommended zone attribute of the specific zone to '0'; and
- b) generating a Zone Descriptor Changed event for the specific zone.

If a zone is in the ZSF:Read Only state or the ZSF:Offline state, then the Finish Zone Recommended zone attribute shall be cleared to '0'.

5.6 Zone Active Excursions

A Zone Active Excursion is a vendor specific action on a zone that is in the:

- a) ZSIO:Implicitly Opened state;
- b) ZSEO:Explicitly Opened state; or
- c) ZSC:Closed state,

and transitions the zone to the ZSF:Full state. A Zone Active Excursion is orthogonal to the Finish Zone Recommended feature (refer to section 5.5), and can occur at any point when a zone has transitioned to one of the listed zone states. If the controller performs a Zone Active Excursion on such a zone, then the controller shall notify the host by:

- a) setting the Zone Finished by Controller zone attribute of that zone to '1' (refer to Figure 37);
and
- b) generating a Zone Descriptor Changed event for that zone.

If a zone is in the ZSF:Read Only state or the ZSF:Offline state, then the Zone Finished by Controller attribute shall be cleared to '0'.

Annex A. Zoned Namespaces Host Considerations (Normative)

A.1 Introduction

The Zoned Namespace Command Set defines host management requirements in addition to those defined in the NVM Command Set.

To facilitate better interactions between hosts and NVM subsystems, this annex describes the Zoned Namespaces feature from a host perspective.

A.2 Writing to Zones

In a Sequential Write Required zone, writes to a zone are required to start at the valid write pointer address.

A.3 Open Zone Considerations

A.3.1 Overview

Each zone that is in either the ZSIO:Implicitly Opened state or the ZSEO:Explicitly Opened state has both Active Resources and Open Resources associated. Limitations on the availability of Active Resources and Open Resources and the number of zones allowed in these zone states are indicated by the Maximum Open Resources field and the Maximum Active Resources field.

Each zone that is in the ZSC:Closed state has only Active Resources associated. Availability of Active Resources and the number of zones allowed in this zone state are indicated by the Maximum Active Resources field.

If a write operation is processed on LBAs in a zone in the ZSE:Empty state or the ZSC:Closed state, then that zone shall be managed as if that zone is requested to transition to the ZSIO:Implicitly Opened state and shall account for Active Resources and Open Resources as defined in section 2.1.1.2.

A zone remains in the ZSIO:Implicitly Opened state or the ZSEO:Explicitly Opened state until:

- a) a write operation is processed for this zone that causes the zone to reach its zone capacity;
- b) the zone state is ZSIO:Implicitly Opened state, and the zone is selected by the controller (which zone is selected is vendor specific) to transition to the ZSC:Closed state;
- c) a Zone Management Send command with Zone Send Action of Finish Zone is issued to the zone;
- d) a Zone Management Send command with Zone Send Action of Reset Zone is issued to the zone;
- e) a Zone Management Send command with Zone Send Action of Close Zone is issued to the zone;
- f) the controller transitions the state of the zone to the ZSRO:Read Only state or the ZSO:Offline state;
- g) the controller transitions the zone to the ZSF:Full state due to a Zone Active Excursion; and
- h) an NVM Subsystem Reset occurs (refer to the Resets section of the NVM Express Base Specification).

A.3.2 Zones in the ZSEO:Explicitly Opened Zones and the ZSIO:Implicitly Opened states

An open zone is a zone that is in either the ZSIO:Implicitly Opened state or the ZSEO:Explicitly Opened state. For a zone that is in either the ZSE:Empty state or ZSC:Closed state, the controller transitions that zone to:

- a) ZSEO:Explicitly Opened state as a result of processing a Zone Management Send command with Zone Send Action of Open Zone; or
- b) ZSIO:Implicitly Opened state as a result of processing a write operation.

If a controller processes a Zone Management Send command with Zone Send Action of Open Zone on a zone in the ZSC:Closed state, then the controller transitions that zone to the ZSEO:Explicitly Opened state.

For resource management purposes, the controller may transition a zone in the ZSIO:Implicitly Opened state to the ZSC:Closed state in order to move the active resource associated with that zone to a different zone that is transitioning into the ZSIO:Implicitly Opened state.

A.3.3 Opening and Closing Zones

A host may use the following techniques to open a zone for writing and following that to close that zone:

- a) Issue a Zone Management Send command with Zone Send Action of Open Zone to transition a zone to the ZSEO:Explicitly Opened state, and following that the host may issue a Zone Management Send command with Zone Send Action of Close Zone to transition that zone to the ZSC:Closed state; or
- b) Issue a write operation to open a zone, and following that the host may issue a Zone Management Send command with Zone Send Action of Close Zone to transition that zone to the ZSC:Closed state.

The controller may autonomously transition any zone from the ZSIO:Implicitly Opened state to the ZSC:Closed state.

A host may use multiple Zone Management Send commands with Zone Send Action of Close Zone to close more zones than the number necessary to satisfy limitations on the number of zones that may be open.

A.3.4 Zone Send Action of Finish Zone Considerations

A zone may be opened (e.g., with a Zone Management Send command with Zone Send Action of Open Zone) prior to processing a Zone Management Send command with Zone Send Action of Finish Zone.

Regardless of how a zone is opened before processing a Zone Management Send command with Zone Send Action of Finish Zone, completion of the command transitions the zone to the ZSF:Full state, with the result that Active and Open Resources become available.

A.4 Partial Failures

A.4.1 Overview

The zone state ZSRO:Read Only provides the ability for a host to continue using a zoned namespace after part of the capacity of the zone has stopped operating (e.g., the controller transitions a zone to the ZSRO:Read Only state as a response to a media failure).

After a zone enters the ZSRO:Read Only state, the host should perform the following actions:

- 1) Transfer the data from that zone to another location (e.g., by using the Copy command); and
- 2) Transition that zone to the ZSO:Offline state by issuing a Zone Management Send command with a Zone Send Action of Offline Zone.

A.5 Capacity and Sizes

There are several data structures involved in size and capacity reporting:

Figure 53: Size and Capacity Fields

Data Structure	Field	Bytes	Description
Identify Namespace Data Structure, NVM Command Set (CNS 00h)	Namespace Size (NSZE)	7:0	(NSZE-1) indicates the highest possible LBA in the zoned namespace
	Namespace Capacity (NCAP)	15:8	The maximum number of allocatable logical blocks in the zoned namespace
	Formatted LBA Size (FLBAS)	26	Index into the list of LBA formats. Refer to the NVM Command Set Specification.
	LBA Format XX (LBAFXX)	191:128	LBADS field indicates the logical block data size (in bytes) (as a power of 2). MS field indicates the metadata size (in bytes).

Figure 53: Size and Capacity Fields

Data Structure	Field	Bytes	Description
Zoned Namespace Command Set specific Identify Namespace	LBA Format Extension XX (LBAFEXX)	3839:2816	ZDES indicates the size of the Zone Descriptor Extension (as a multiple of 64 bytes). ZSZE indicates the number of logical blocks of each zone of the zoned namespace.
Zone Descriptor	Zone Capacity (ZCAP)	15:8	Indicates the maximum number of logical blocks that may be written in the zone associated with the data descriptor.

A zone has a size (ZSZE) and a capacity (ZCAP) that are related to optional support for the variable zone capacity feature. ZCAP may change as a result of resetting the write pointer. Refer to Figure 54.

A zone transitions to the ZSF:Full after successful completion of a write operation that includes $LBA = (ZSLBA + ZCAP_{CURRENT} - 1)$, in addition to other events (e.g., successful completion of a Zone Management Send command with a Zone Send Action of Finish Zone).

Figure 54: Zone Size Relationships