NVM Express® Technical Errata

Errata ID	114
Revision Date	6/26/2023
Affected Spec Ver.	NVM Express® Base Specification Revision 2.0c
Corrected Spec Ver.	

Errata Author(s)

Name	Company
Randy Jennings	Pure Storage
Mike Allison	Samsung
David Black	Dell EMC
Erik Smith	Dell EMC
Fred Knight	NetApp

Errata Overview

This ECN updates and clarifies text within the NVM Express Base.

Revision History

Revision Date	Change Description
10/25/2022	Initial creation
11/3/2022	Revised for section/figures to match published specs. Responded to Mike Allison's comments Added DH-HMAC-CHAP reference correction Added AERL correction from Erik Smith
11/8/2022	Fixed bad references From Administrative Controller – Feature Support, restored Note 3, and added a note about renumbering notes after 3. Added figure reference about using NSID field in Get Log Page Referred to TP4109 as basis for change in 5.16.1.14.1 Persistent Event Log Page LID Specific Parameter Field Resolved comments from Mike Allison
1/11/2023	Updated editorial text for "LID Supported and Effects Data Structure – LID Specific Parameter Field" to match to NVM Express Base Specification 2.0 2023.01.10 NEXT Fixed AERL correction Fixed ordering in "Description of Changes" and added AERL correction
1/18/2023	Added note to "Figure 31 Discovery Controller Initialization process flow" for the editor to increase superscript font to 14pt. Added context for 3.1.1 change, "In the conditions terminating an association, the IO Queue was listed in the third & forth bullet; removed the third bullet reference."

	Add change, for controller supported features, feature "Rotational Media" corrected to "Spinup Control".
	Highlighted in yellow the new references in new footnote for "Figure 139: Opcodes for Admin Commands".
1/23/2023	Left text about features supported by I/O Controllers in the Admin Controllers section; more needs to be done about Admin controllers and instructions in the Command specifications. Now Bug 36.
	Added text that other fields in LID Supported and Effects Data Structure shall be ignored if LSUPP is cleared to '0'.
2/1/2023	Remove Power Management change.
	Moved copyright dates to include 2023
	Use '®' in footer
2/6/2023	Remove column width/justification changes from 8.18.2 RPMB Operations, leaving '.' fixes.
5/1/2023	Integrated
	Made blues consistent (light blue)
	"Section 5.16.1" -> "section 5.16.1"
	Combined two Figure 34 Discovery Controller – Feature Support
6/25/2023	Editorial updates per Randy Jennings and Mike Allison
6/26/2023	Fixed spelling
	Corrected Features Support section number for Discovery Controllers.

Description of Changes

NVM Express Base specification:

Editorial Changes:

- For I/O, Admin, and Discovery controller supported features, feature "Rotational Media" corrected to "Spinup Control".
- In the conditions terminating an association, the IO Queue was listed in the third & forth bullet; removed the third bullet reference.
- Reservations are not supported by Admin Controllers. Therefore, reservation related features are prohibited. (Host Identifier is okay, but reservations are not a motivation to have it.)
- Swapped footnote content on Figure 31: Discovery Controller Initialization process flow. Also, added note for superscripts (footnote references) to change to 14pt font in this table.
- NVMe-MI log page and feature should not be mandatory for Discovery Controllers; NVMe-MI Send/Receive are prohibited.
- Property E1Ch is reserved; it should not be mandatory for Admin Controllers.
- Removed footnotes from some tables that are not used
- Footnote clarifications about Namespace use in Figure 135: Opcodes for Admin Commands
- Clarified conditions to clear Asynchronous Events.
- Clarified that the rest of the fields in the LID Supported and Effects Data Structure (Figure 204) should be ignored if LSUPP is cleared to '0'.
- Implementations that do not support Persistent Events do not need to indicate support for Establish Context and Read 512 Bytes of Header.
- AERL is mandatory for Discovery Controllers that implement explicit persistent connections, not reserved.
- Filled out section references to 8.13.5.9 Generated PSK for TLS from 8.13.5.4 DH-HMAC-CHAP_Reply Message. Previous reference was to 8.13, which is not helpful.
- Clarified that Replay Protected Memory Block Message Types are sent and received with Security Send/Receive commands.
- Fix '.'s in a couple of tables from 8.18.2 RPMB Operations.
- In the Host Considerations about Fuse commands, clarified that the SLBA in the write is the first LBA to write. (It is also the first LBA to compare, as the text currently states, but this change is easier to process).

Editor's Note:
BLACK text indicates unchanged text; **BLUE** text indicates newly inserted text, <mark>**HIGHLIGHED YELLOW**</mark>, if it is a reference to add, **RED** text indicates deleted text; **PURPLE** text indicates moved text, **GREEN** text indicates editor notes.

Description of NVM Express® Base Specification changes

3 NVM Express Architecture
3.1 NVM Controller Architecture
3.1.1 Controller Model

…

An association between a host and controller is terminated if:

- the controller is shutdown as described in section 3.6.2;
- a Controller Level Reset occurs;
- the NVMe Transport connection is lost between the host and controller for the Admin Queue ~~or any I/O Queue~~; or
- an NVMe Transport connection is lost between the host and controller for any I/O Queue and the host or controller does not support individual I/O Queue deletion (refer to section 3.3.2.4).

…

3.1.2 Controller Types
3.1.2.1 I/O Controller
3.1.2.1.3 Features Support

…

Figure 25: I/O Controller – Feature Support

Feature Name	Feature Support Requirements [1]	Logged in Persistent Event Log [1]
…		
~~Rotational Media~~ Spinup Control	O	O
Notes:		
…		

3.1.2.2 Administrative Controller
3.1.2.2.3 Features Support

…

Figure 30: Administrative Controller – Feature Support

Feature Name	Feature Support Requirements [1]	Logged in Persistent Event Log [1]
…		
Host Identifier	O[4]	O
Reservation Notification Mask	~~O[5]~~P	~~O~~P
Reservation Persistence	~~O[5]~~P	~~O~~P
…		
~~Rotational Media~~ Spinup Control	P	P

Figure 30: Administrative Controller – Feature Support

Feature Name	Feature Support Requirements [1]	Logged in Persistent Event Log [1]
Notes:		

Notes:
1. O/M/P/NR definition: O = Optional, M = Mandatory, P = Prohibited, NR = Not Recommended.
2. The feature is mandatory for NVMe over PCIe. This feature is not supported for NVMe over Fabrics.
3. Mandatory if Telemetry Log, Firmware Commit or SMART/Health Critical Warnings are supported.
4. ~~Mandatory if reservations are supported as indicated in the Identify Controller data structure.~~
   <NOTE: please renumber the notes following 3, and update the references>
5. ~~Mandatory if reservations are supported by the namespace as indicated by a non-zero value in the Reservation Capabilities (RESCAP) field in the Identify Namespace data structure.~~
6. This feature is optional for NVM subsystems that do not implement a Management Endpoint. For NVM subsystems that implement any Management Endpoint refer to the NVM Express Management Interface Specification.

3.1.2.3 Discovery Controller

3.1.2.3.1 Discovery Controller Initialization

…

Figure 31 Discovery Controller Initialization process flow

<Editor: please increase the font size of superscripts to 14pt>
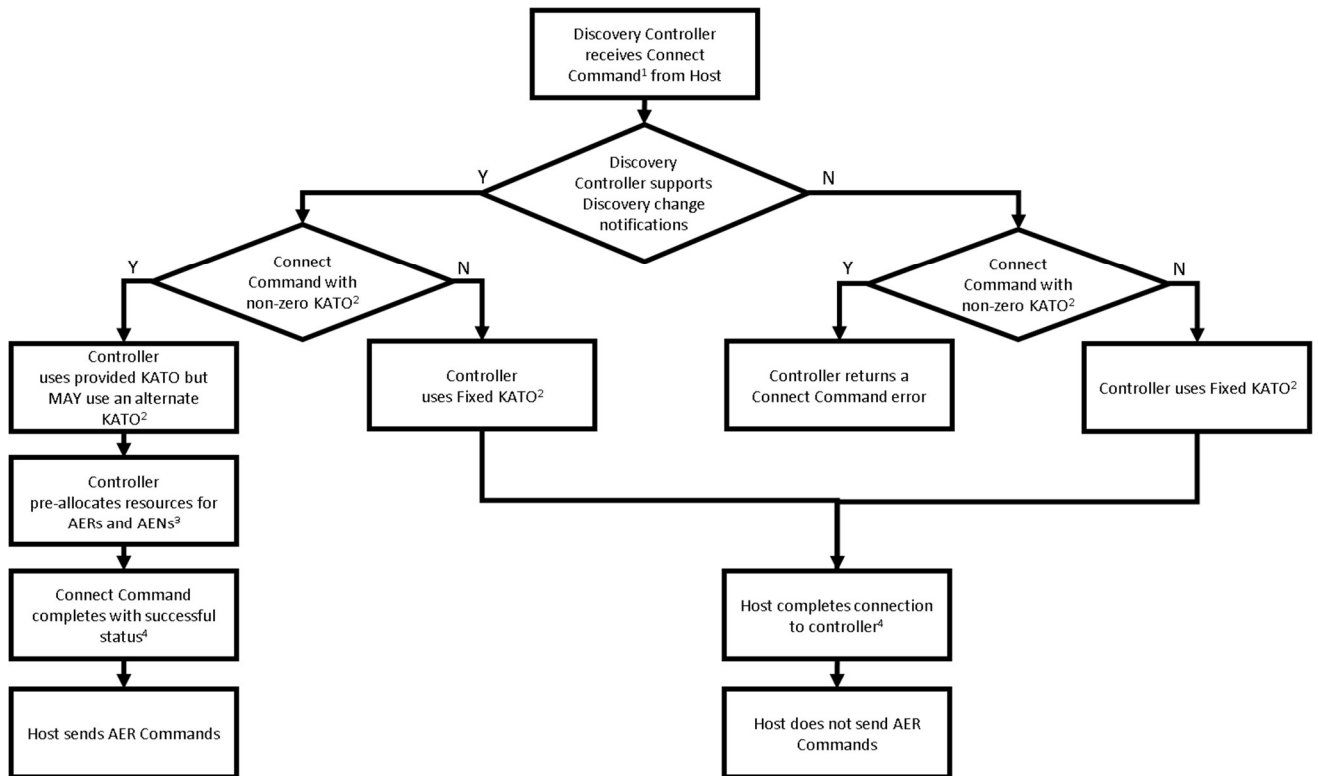
Notes:
1. Refer to section 6.3~~;~~.
2. <Editor: Note swapped footnote content>Refer to the Keep Alive command in section 5.18~~;~~.
3. Refer to the Asynchronous Event Request command in section 5.2~~; and~~.
4. Refer to the following steps in this section.

…

3.1.2.3.3 Log Page Support

…

Figure 33: Discovery Controller – Log Page Support

Log Page Name	Command Support Requirements [1]
…	
NVMe-MI Commands Supported and Effects	~~M~~[3] O
…	

Notes
1. O/M/P definition: O = Optional, M = Mandatory, P = Prohibited
2. Optional if Set Features command is not supported (refer to Figure 32).
3. ~~Optional if NVMe-MI Send command and NVMe-MI Receive command is not supported (refer to Figure 32).~~
   <NOTE: please renumber the notes following 3, and update the references>
4. Optional for versions 1.1 and earlier of the NVMe over Fabrics specification.

3.1.2.3.4 Features Support

…

Figure 34: Discovery Controller – Feature Support

Feature Name	Feature Support Requirements [1]	Logged in Persistent Event Log [1]
…		
Enhanced Controller Metadata	O[2]	O
Controller Metadata	O[2]	O
Namespace Metadata	O[2]	O
~~Rotational Media~~ Spinup Control	P	P

Notes:
1. O/M/P/NR definition: O = Optional, M = Mandatory, P = Prohibited, NR = Not Recommended.
2. ~~This feature is optional for NVM subsystems that do not implement a Management Endpoint. For NVM subsystems that implement any Management Endpoint refer to the NVM Express Management Interface Specification.~~

3.1.3 Controller Properties

…

Figure 35: Property Definition

Offset (OFST)	Size (in bytes)	I/O Controller[1]	Admin. Controller[1]	Discovery Controller[1]	Name
…					
E1Ch		R	~~M~~R	R	Reserved
…					

…

3.3 NVM Queue Models

3.3.3 Queueing Data Structures

3.3.3.2   Common Completion Queue Entry

3.3.3.2.1   Status Field Definition

3.3.3.2.1.2 Command Specific Status Definition

…

Figure 96: Status Code – Command Specific Status Values, I/O Commands

Value	Description
80h	Conflicting Attributes
81h	Invalid Protection Information
82h	Attempted Write to Read Only Range
83h	Command Size Limit Exceeded
84h	Invalid Command ID
85h to B7h	Reserved
B8h	Zoned Boundary Error
B9h	Zone Is Full
BAh	Zone Is Read Only
BBh	Zone Is Offline
BCh	Zone Invalid Write
BDh	Too Many Active Zones
BEh	Too Many Open Zones
BFh	Invalid Zone State Transition
Notes:	
1. A = All I/O Command Sets, C = Command Set Specific.	

…

5 Admin Command Set

…

Figure 139: Opcodes for Admin Commands

Opcode by Field			Combined Opcode [1]	Namespace Identifier Used [2]	Command	Command Set Specific [8]
(07) Generic Command	(06:02) Function	(01:00) Data Transfer[3]				
…						
0b	000 00b	10b	02h	Yes[10] <note footnote addition>	Get Log Page	No
…						
0b	111 11b	11b	7Fh	NOTE 9 No	Fabrics Commands[9]	No
…						
1b	000 01b	10b	86h	NOTE 4 Yes[4]	Get LBA Status	NVM, ZNS

Figure 139: Opcodes for Admin Commands

Opcode by Field			Combined Opcode [1]	Namespace Identifier Used [2]	Command	Command Set Specific [8]
(07) Generic Command	(06:02) Function	(01:00) Data Transfer[3]				
Vendor Specific						
1b	n/a	NOTE 3	C0h to FFh		Vendor specific	

Notes:
1. Opcodes not listed are reserved.
2. A subset of commands use the Namespace Identifier (NSID) field. If the Namespace Identifier field is used, then the value FFFFFFFFh is supported in this field unless otherwise indicated in footnotes in this figure that a specific command does not support that value or supports that value only under specific conditions. When this field is not used, the field is cleared to 0h as described in Figure 87.
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 does not support the use of the Namespace Identifier (NSID) field set to FFFFFFFFh.
5. Support for the Namespace Identifier field set to FFFFFFFFh depends on the Directive Operation (refer to section 8.7).
6. Use of the Namespace Identifier field depends on the CNS value in the Identify Command as described in Figure 273.
7. The use of the Namespace Identifier is Security Protocol specific.
8. No = Not I/O Command Set specific, A = All I/O Command Sets, NVM = NVM Command Set specific, ZNS = Zoned Namespace Command Set.
9. All Fabrics commands use the opcode 7Fh. Refer to section 6 for details.
10. Use of the Namespace Identifier field is specified further in section 5.16.1 and Figure 202. <NOTE: editor, please insert link.> <NOTE: NVM Express Base Specification 2.0 2023.01.10 NEXT already has a footnote 10. This should come after.>

5.2 Asynchronous Event Request command
5.2.1 Command Completion

…

Figure 147: Asynchronous Event Information – Notice

Value	Description
...	
01h	**Firmware Activation Starting:** The controller is starting a firmware activation process during which command processing is paused. The host ~~software~~ may use CSTS.PP to determine when command processing has resumed. To clear this event, the host ~~software~~ reads the Firmware Slot Information log page with the Retain Asynchronous Event bit cleared to '0'.
...	
04h	**Predictable Latency Event Aggregate Log Change:** Indicates that event pending entries for one or more NVM Sets (refer to section 5.22.1.12) have been added to the Predictable Latency Event Aggregate log. To clear this event, the host reads the Predictable Latency Event Aggregate log page with the Retain Asynchronous Event bit cleared to '0'.
...	
F0h	**Discovery Log Page Change:** A change has occurred to one or more of the Discovery log pages. ~~The host or Discovery controller should submit a Get Log Page command to receive updated Discovery log pages.~~ To clear this event, the host or Discovery controller reads the Discovery log page with the Retain Asynchronous Event bit cleared to '0'.
...	

Figure 148: Asynchronous Event Information – I/O Command Specific Status

Value	Description
00h	**Reservation Log Page Available:** Indicates that one or more Reservation Notification log pages (refer to section 5.16.1.24) have been added to the Reservation Notification log. To clear this event, the host reads the Reservation log page with the Retain Asynchronous Event bit cleared to '0'.

Figure 148: Asynchronous Event Information – I/O Command Specific Status

Value	Description
01h	**Sanitize Operation Completed:** Indicates that a sanitize operation has completed (including any associated additional media modification, refer to the No-Deallocate Modifies Media After Sanitize field in Figure 275) without unexpected deallocation of all user data (refer to section 5.27.1.19) and status is available in the Sanitize Status log page (refer to section 5.16.1.25). To clear this event, the host reads the Sanitize Status log page with the Retain Asynchronous Event bit cleared to '0'.
02h	**Sanitize Operation Completed With Unexpected Deallocation:** Indicates that a sanitize operation for which No-Deallocate After Sanitize (refer to Figure 303) was requested has completed with the unexpected deallocation of all user data (refer to section 5.27.1.19) and status is available in the Sanitize Status log page (refer to section 5.16.1.25). To clear this event, the host reads the Sanitize Status log page with the Retain Asynchronous Event bit cleared to '0'.
03h to FFh	Reserved

…

5.16 Get Log Page command

5.16.1 Log Specific Information

5.16.1.1 Supported Log Pages (Log Identifier 00h)

…

Figure 204: LID Supported and Effects Data Structure

Bits	Description
…	
0	**LID Supported (LSUPP):** If this bit is set to '1', then the controller supports this LID for a Get Log Page command. If this bit is cleared to '0', then the controller does not support this LID for a Get Log Page command, and the host should ignore other fields in this data structure. Refer to section 3.1.2 for the LID support requirements for each controller type.

Figure 205: LID Supported and Effects Data Structure – LID Specific Parameter Field

<Note: TP4109 moved this field to 5.22.1.14.1 Persistent Event Log Page LID Specific Parameter Field, Figure 247: Persistent Event Log LID Specific Parameter Field. Showing text from NVM Express Base Specification 2.0 2023.01.10 NEXT.)>

Bits	Description
15:1	Reserved
0	**Establish Context and Read 512 Bytes of Header Supported (ECRH):** If this bit is cleared to '0', then the controller does not support the Establish Context and Read 512 Bytes of Header action (refer to Figure 250). If this bit is set to '1', then the controller supports: • the Establish Context and Read 512 Bytes of Header action; and • the Generation Number field in the Persistent Event log page. If the Persistent Event log page is supported by the controller, then implementations compliant with NVM Express Base Specification, Revision 2.0 and later shall set this bit to '1'.

5.17   Identify command

5.17.2 Identify Data Structures

5.17.2.1 Identify Controller Data Structure (CNS 01h)

…

Figure 275: Identify – Identify Controller Data Structure, I/O Command Set Independent

Bytes	I/O[1]	Admin[1]	Disc[1]	Description
…				
259	M	M	~~R~~M[3]	**Asynchronous Event Request Limit (AERL):** This field is used to convey the maximum number of concurrently outstanding Asynchronous Event Request commands supported by the controller (refer to section 5.2). This is a 0's based value. It is recommended that implementations support a minimum of four Asynchronous Event Request ~~Limit~~ commands outstanding simultaneously.
…				

…
1. O/M/R definition: O = Optional, M = Mandatory, R = Reserved.
2. Mandatory for I/O controllers using a message-based transport. Reserved for I/O controllers using a memory-based transport.
3. Mandatory for Discovery controllers that support explicit persistent connections. Reserved for Discovery controllers that do not support explicit persistent connections.   <Note: this is opposite order from Figure 32, but it matches the order in 2 (Reserved last)>

8 Extended Capabilities

8.13 NVMe over Fabrics Secure Channel and In-band Authentication

8.13.5 DH-HMAC-CHAP Protocol

8.13.5.4 DH-HMAC-CHAP_Reply Message

…

Challenge Valid: If the host does not require bidirectional authentication or no establishment of a secure channel after unidirectional authentication is sought (refer to section 8.13.5.9<Note: reference update>), this field shall be cleared to 0h. Otherwise, this field shall be set to 01h.

DH Value Length (DHVLEN): Diffie-Hellman exponential length. This length shall be a multiple of 4. If the DH group identifier is cleared to 0h (i.e., NULL DH exchange), this field shall be cleared to 0h. Otherwise, it shall be set to the length in bytes of the DH Value.

Sequence Number (SEQNUM): 32-bit sequence number S_2. A random non-zero value shall be used as the initial value. The sequence number is incremented modulo 2^{32} after each use, except that the value 0h is skipped (i.e., incrementing the value FFFFFFFFh results in the value 00000001h). The value 0h is used to indicate that bidirectional authentication is not performed, but a challenge value C_2 is carried in order to generate a pre-shared key (PSK) for subsequent establishment of a secure channel (refer to section 8.13.5.9). <Note: reference update and added missing period>

…

8.18   Replay Protected Memory Block

…
The message types defined in Figure 461 are used by the host to communicate with an RPMB target. Request Message Types are sent from the host to the controller using Security Send commands. Response Message Types are ~~sent to~~ retrieved by the host from the controller using Security Receive commands.

Technical input submitted to the NVM Express® Workgroup is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2023 NVM Express, Inc.

…

8.18.2 RPMB Operations

8.18.2.4   Authenticated Data Read

…

Figure 468: RPMB – Authenticated Data Read Flow

Command	Bytes in Command	Field Name	Value	Objective
	Data populated by the host and sent to the controller			
	222-*N*:00	Stuff Bytes	0…00h	
	222:222-(*N*-1)	MAC/Key	0…00h<Note correction>	
	223	RPMB Target	*RPMB target to access*	
Security Send 1	239:224	Nonce	*Nonce generated by the host*	Read Data request
	243:240	Write Counter	00000000h	
	247:244	Address	*Address in RPMB*	
	251:248	Sector Count	*Number of 512B blocks*	
	253:252	Result	0000h	
	255:254	Request/Response	0004h *(Request)*	
…				

8.18.4 Authenticated Device Configuration Block Read

…

Figure 470: RPMB – Authenticated Device Configuration Block Read Flow

Command	Bytes in Command	Field Name	Value	Objective
	Data populated by the host and sent to the controller			
	222-*N*:00	Stuff Bytes	0…00h	
	222:222-(*N*-1)	MAC/Key	0…00h<Note correction>	
	223	RPMB Target	00h	
Security Send 1	239:224	Nonce	*Nonce generated by the host*	Request Device Configuration Block Read
	243:240	Write Counter	00000000h	
	247:244	Address	00000000h	
	251:248	Sector Count	00000001h	
	253:252	Result	0000h	
	255:254	Request/Response	0007h *(Request)*	
…				

Annex B. Host Considerations (Informative)

B.3 Executing a Fused Operation

…
The attributes of the Write command are:

- CMD1.CDW0.OPC is set to 01h for Write;
- CMD1.CDW0.FUSE is set to 10b indicating that this is the second command of a fused operation;
- CMD1.CDW0.CID is set to a free command identifier;
- CMD1.NSID is set to identify the appropriate namespace. This value shall be the same as CMD0.NSID;
- If metadata is being used in a separate buffer, then the location of that buffer is specified in the CMD1.MPTR field;

- The physical address of the first page of data to write is identified:
    - If the command uses PRPs, then CMD1.PRP1 is set to the physical address of the first page of the data to write and CMD1.PRP2 is set to the physical address of the PRP List. The PRP List includes three entries; or
    - If the command uses SGLs, CMD1.SGL1 is set to an appropriate SGL segment descriptor depending on whether more than one descriptor is needed;
- CMD1.CDW10.SLBA is set to the first LBA to ~~compare against~~write. Note that this field also spans Command Dword 11. This value shall be the same as CMD0.CDW10.SLBA;