



LEGAL NOTICE:

© Copyright 2008 to 2023 NVM Express®, Inc. ALL RIGHTS RESERVED.

This Technical Proposal is proprietary to the NVM Express, Inc. (also referred to as "Company") and/or its successors and assigns.

NOTICE TO USERS WHO ARE NVM EXPRESS, INC. MEMBERS: Members of NVM Express, Inc. have the right to use and implement this Technical Proposal subject, however, to the Member's continued compliance with the Company's Intellectual Property Policy and Bylaws and the Member's Participation Agreement.

NOTICE TO NON-MEMBERS OF NVM EXPRESS, INC.: If you are not a Member of NVM Express, Inc. and you have obtained a copy of this document, you only have a right to review this document or make reference to or cite this document. Any such references or citations to this document must acknowledge NVM Express, Inc. copyright ownership of this document. The proper copyright citation or reference is as follows: "© 2008 to 2023 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.

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

NVM Express® Technical Proposal for New Feature

Technical Proposal ID	TP4171 Subsystem Local Memory Command Set Specification changes to other specifications
Change Date	2023-10-04
Builds on Specification	NVM Express Base Specification 2.0c NVM Command Set Specification 1.0c
References Specification	TP4130 Cross Namespace Copy TP4131 Subsystem Local Memory

Technical Proposal Author(s)

Name	Company
Kim Malone	Intel
Bill Martin	Samsung
Fred Knight	NetApp
David Black	Dell EMC

This proposal provides changes to the NVM Command Set Specification and NVM Express Base Specification necessary for the Subsystem Local Memory Command Set.

Revision History

Revision Date	Change Description
2023-05-17	<ul style="list-style-type: none">Initial version, moved content from Subsystem Local Memory TP4131
2023-07-05	<ul style="list-style-type: none">Updated NVMe logo
2023-08-03	<ul style="list-style-type: none">Integrated
2023-08-21	<ul style="list-style-type: none">Added Command Set Identifier value
2023-08-24	<ul style="list-style-type: none">Deleted “h” in copy descriptor format bit column. Reformatting.
2023-09-14	<ul style="list-style-type: none">Integrated
2023-10-04	<ul style="list-style-type: none">Updated per Kim Malone review

Summary of changes:

- NVM Command Set Specification
 - Modifications to Copy command
- NVM Express Base Specification
 - Modifications to Identify Controller Data Structure (CNS 01h)
 - Modifications to Host Behavior Support (Feature Identifier 16h)
 - Modifications to Command Set Identifiers

Markup Conventions:

Black:	Unchanged (however, hot links are removed)
Red Strikethrough:	Deleted
Orange Text	Unchanged text from TP4130
Orange Text	Deleted text from TP4130
Blue:	New
Blue Highlighted:	TBD values, anchors, and links to be inserted in new text
Purple	Text moved without change.
Purple Text	Content that was removed and put in a new location
Green Text:	Editor's notes

Modify portions of NVMe Base Specification 2.0c as follows:

5 Admin Command Set

...

5.17 Identify command

5.17.1 Identify command overview

...

Figure 274: Command Set Identifiers

Command Set Identifier Value	Description	Reference Section
00h	NVM Command Set	Refer to the NVM Command Set Specification
01h	Key Value Command Set	Refer to the Key Value Command Set Specification
02h	Zoned Namespace Command Set	Refer to the Zoned Namespace Command Set Specification
03h	Subsystem Local Memory Command Set	Refer to the Subsystem Local Memory Command Set Specification
04h to 2Fh	Reserved	
30h to 3Fh	Vendor specific	
40h to FFh	Reserved	

...

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	Admi n 1	Dis c 1	Description					
NVM Command Set Attributes									
521:520	M	M	R	Optional NVM Command Support (ONCS): This field indicates the optional I/O commands and features supported by the controller. Refer to section 3.1.2.					
				<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								

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

Bytes	I/O 1	Admi n 1	Dis c 1	Description	
				10	NVM All Fast Copy (NVMAFC): If set to '1', then within this NVM subsystem, all copy operations performed by the controller for NVM Command Set Copy commands are fast copy operations (refer to the Fast copy operations section of the NVM Command Set Specification). If cleared to '0', then within this NVM subsystem, some copy operations performed by the controller for NVM Command Set Copy commands may not be fast copy operations. This bit is ignored by the host if the controller does not support the NVM Command Set Copy command.
				...	
				NOTE: This field applies to all I/O Command Sets. The original name has been retained for historical continuity.	
...					
535:534	M	R	R	Copy Descriptor Formats Supported:	
				Bits	Description
				15:45	Reserved
				4	If set to '1', then the controller supports Copy Descriptor Format 4h. If cleared to '0', then the controller does not support Copy Descriptor Format 4h.
				3	If set to '1', then the controller supports Copy Descriptor Format 3h. If cleared to '0', then the controller does not support Copy Descriptor Format 3h.
				2	If set to '1', then the controller supports Copy Descriptor Format 2h. If cleared to '0', then the controller does not support Copy Descriptor Format 2h.
				1	If set to '1', then the controller supports Copy Descriptor Format 1h. If cleared to '0', then the controller does not support Copy Descriptor Format 1h.
				0	If set to '1', then the controller supports Copy Descriptor Format 0h. If cleared to '0', then the controller does not support Copy Descriptor Format 0h.

...

5.17.2.21 Identify I/O Command Set data structure (CNS 1Ch)

...

Figure 1: I/O Command Set Vector

Bit	Description
63:43	Reserved

Figure 1: I/O Command Set Vector

Bit	Description
3	Subsystem Local Memory Command Set: This bit is set to '1' if the Subsystem Local Memory Command Set is selected. This bit is cleared to '0' if the Subsystem Local Memory Command Set is not selected.
2	Zoned Namespace Command Set: This bit is set to '1' if the Zoned Namespace Command Set is selected. This bit is cleared to '0' if the Zoned Namespace Command Set is not selected.
1	Key Value Command Set: This bit is set to '1' if the Key Value Command Set is selected. This bit is cleared to '0' if the Key Value Command Set is not selected.
0	NVM Command Set: This bit is set to '1' if the NVM Command Set is selected. This bit is cleared to '0' if the NVM Command Set is not selected.

...

5.27 Set Features command

...

5.27.1 Feature Specific Information

...

5.27.1.18 Host Behavior Support (Feature Identifier 16h)

...

Figure 351: Host Behavior Support – Data Structure

Bytes	Description												
...													
05:04	Copy Descriptor Formats Enable (CDFE): The bits in this field that are used (i.e., are neither reserved nor unused) enable the corresponding Copy Descriptor Formats (e.g., bit 2 enables Copy Descriptor Format 2h). Bits in this field that this specification defines as unused (i.e., bits 1:0) correspond to Copy Descriptor Formats (refer to the NVM Command Set Specification) that are always enabled if they are supported.												
	<table><tr><th>Bits</th><th>Description</th></tr><tr><td>15:45</td><td>Reserved</td></tr><tr><td>4</td><td>Copy Descriptor Format 4h Enable (CDF4E): If set to '1' and the controller supports Copy Descriptor Format 4h, then Copy Descriptor Format 4h is enabled. If cleared to '0', or the controller does not support Copy Descriptor Format 4h, then Copy Descriptor Format 4h is disabled.</td></tr><tr><td>3</td><td>Copy Descriptor Format 3h Enable (CDF3E): If set to '1' and the controller supports Copy Descriptor Format 3h, then Copy Descriptor Format 3h is enabled. If cleared to '0', or the controller does not support Copy Descriptor Format 3h, then Copy Descriptor Format 3h is disabled.</td></tr><tr><td>2</td><td>Copy Descriptor Format 2h Enable (CDF2E): If set to '1' and the controller supports Copy Descriptor Format 2h, then Copy Descriptor Format 2h is enabled. If cleared to '0', or the controller does not support Copy Descriptor Format 2h, then Copy Descriptor Format 2h is disabled.</td></tr><tr><td>1:0</td><td>Unused. The controller shall ignore these bits when processing a Set Features command that specifies this Feature and shall clear these bits to '0' in the attributes returned in the data buffer for a Get Features command that specifies this Feature.</td></tr></table>	Bits	Description	15:45	Reserved	4	Copy Descriptor Format 4h Enable (CDF4E): If set to '1' and the controller supports Copy Descriptor Format 4h, then Copy Descriptor Format 4h is enabled. If cleared to '0', or the controller does not support Copy Descriptor Format 4h, then Copy Descriptor Format 4h is disabled.	3	Copy Descriptor Format 3h Enable (CDF3E): If set to '1' and the controller supports Copy Descriptor Format 3h, then Copy Descriptor Format 3h is enabled. If cleared to '0', or the controller does not support Copy Descriptor Format 3h, then Copy Descriptor Format 3h is disabled.	2	Copy Descriptor Format 2h Enable (CDF2E): If set to '1' and the controller supports Copy Descriptor Format 2h, then Copy Descriptor Format 2h is enabled. If cleared to '0', or the controller does not support Copy Descriptor Format 2h, then Copy Descriptor Format 2h is disabled.	1:0	Unused. The controller shall ignore these bits when processing a Set Features command that specifies this Feature and shall clear these bits to '0' in the attributes returned in the data buffer for a Get Features command that specifies this Feature.
	Bits	Description											
	15:45	Reserved											
	4	Copy Descriptor Format 4h Enable (CDF4E): If set to '1' and the controller supports Copy Descriptor Format 4h, then Copy Descriptor Format 4h is enabled. If cleared to '0', or the controller does not support Copy Descriptor Format 4h, then Copy Descriptor Format 4h is disabled.											
	3	Copy Descriptor Format 3h Enable (CDF3E): If set to '1' and the controller supports Copy Descriptor Format 3h, then Copy Descriptor Format 3h is enabled. If cleared to '0', or the controller does not support Copy Descriptor Format 3h, then Copy Descriptor Format 3h is disabled.											
	2	Copy Descriptor Format 2h Enable (CDF2E): If set to '1' and the controller supports Copy Descriptor Format 2h, then Copy Descriptor Format 2h is enabled. If cleared to '0', or the controller does not support Copy Descriptor Format 2h, then Copy Descriptor Format 2h is disabled.											
1:0	Unused. The controller shall ignore these bits when processing a Set Features command that specifies this Feature and shall clear these bits to '0' in the attributes returned in the data buffer for a Get Features command that specifies this Feature.												
511:06	Reserved												

...

Modify portions of NVM Command Set Specification 1.0c as follows:

3 I/O Commands for the NVM Command Set

...

3.2 NVM Command Set Commands

...

3.2.2 Copy command

The Copy command is used by the host to copy user data from one or more source logical block ranges in one or more source namespaces to a single consecutive destination logical block range in a destination namespace. Each source logical block range may be in the same namespace or a different namespace with respect to any other source logical block range and with respect to the destination logical block range.

...

Figure 30: Copy – Command Dword 12

Bits	Description														
...															
11:08	Descriptor Format: Specifies the type of the Copy Descriptor Format that is used. The Copy Descriptor Format specifies the starting LBA location, number of logical blocks length, and parameters associated with the read portion of the operation.														
	<table><tr><th>Copy Descriptor Format type</th><th>Description</th></tr><tr><td>0h</td><td>Source Range Entries Copy Descriptor Format 0h is used (refer to Figure 34).</td></tr><tr><td>1h</td><td>Source Range Entries Copy Descriptor Format 1h is used (refer to Figure 35).</td></tr><tr><td>2h</td><td>Source Range Entries Copy Descriptor Format 2h is used (refer to Figure 34).</td></tr><tr><td>3h</td><td>Source Range Entries Copy Descriptor Format 3h is used (refer to Figure 35).</td></tr><tr><td>4h</td><td>Source Range Entries Copy Descriptor Format 4h is used (refer to SLM Command Set Specification).</td></tr><tr><td>All Others</td><td>Reserved</td></tr></table>	Copy Descriptor Format type	Description	0h	Source Range Entries Copy Descriptor Format 0h is used (refer to Figure 34).	1h	Source Range Entries Copy Descriptor Format 1h is used (refer to Figure 35).	2h	Source Range Entries Copy Descriptor Format 2h is used (refer to Figure 34).	3h	Source Range Entries Copy Descriptor Format 3h is used (refer to Figure 35).	4h	Source Range Entries Copy Descriptor Format 4h is used (refer to SLM Command Set Specification).	All Others	Reserved
	Copy Descriptor Format type	Description													
	0h	Source Range Entries Copy Descriptor Format 0h is used (refer to Figure 34).													
	1h	Source Range Entries Copy Descriptor Format 1h is used (refer to Figure 35).													
	2h	Source Range Entries Copy Descriptor Format 2h is used (refer to Figure 34).													
	3h	Source Range Entries Copy Descriptor Format 3h is used (refer to Figure 35).													
	4h	Source Range Entries Copy Descriptor Format 4h is used (refer to SLM Command Set Specification).													
All Others	Reserved														
...															

...

The copy descriptor format types are distinguished by the supported protection information formats (refer to section 5.2.1) and whether the Copy Descriptor Format contains a Source Namespace Identifier (SNSID) field that supports a copy source in a different namespace than the copy destination, as described in Figure nnn. The source namespace specified by a Copy Descriptor Format that does not contain an SNSID field is the same as the destination namespace.

Figure nnn: Copy – Copy Descriptor Format Types

Copy Descriptor Format Type	Protection Information Formats	SNSID field present	Description
0h	16b Guard Protection Information	No	Protection Information size: 8 bytes, source namespace and destination namespace are the same.

Figure 99: Copy – Copy Descriptor Format Types

Copy Descriptor Format Type	Protection Information Formats	SNSID field present	Description
1h	32b Guard Protection Information 64b Guard Protection Information	No	Protection Information size: 16 bytes, source namespace and destination namespace are the same.
2h	16b Guard Protection Information	Yes	Protection Information size: 8 bytes, source namespace may differ from destination namespace.
3h	32b Guard Protection Information 64b Guard Protection Information	Yes	Protection Information size: 16 bytes, source namespace may differ from destination namespace.
4h	None	Yes	The source namespace supports the SLM Command Set (refer to the SLM Command Set Specification).

...

If the number of Source Range entries (i.e., the value in the NR field) is greater than the value in the MSRC field (refer to Figure 97), then the Copy command shall be aborted with a status code of Command Size Limit Exceeded.

~~If a valid Source Range Entry specifies a Number of Logical Blocks field that is greater than the value in the MSSRL field (refer to Figure 97), then the Copy command shall be aborted with a status code of Command Size Limit Exceeded.~~

~~If the sum of all Number of Logical Blocks fields in all Source Range entries is greater than the value in the MCL field (refer to Figure 97), then the Copy command shall be aborted with a status code of Command Size Limit Exceeded.~~

~~The number of logical blocks written by the Copy command is the sum of all Number of Logical Blocks fields in all Source Range entries specified in the list of Source Range entries.~~

For an LBA based Copy Descriptor Format Type (i.e., 0h, 1h, 2h, or 3h):

- the number of logical blocks written by the Copy command is the sum of all Number of Logical Blocks fields in all Source Range entries specified in the list of Source Range entries;
- if a valid Source Range Entry specifies a Number of Logical Blocks field that is greater than the value in the MSSRL field (refer to Figure 97), then the Copy command shall be aborted with a status code of Command Size Limit Exceeded; and
- if the sum of all Number of Logical Blocks fields in all Source Range Entries is greater than the value in the MCL field (refer to Figure 97), then the Copy command shall be aborted with a status code of Command Size Limit Exceeded.

For a byte based Copy Descriptor Format Type (i.e., 4h):

- the number of logical blocks written by the Copy command is determined by the sum of the Number of Bytes fields (refer to the SLM Command Set Specification) in all Source Range entries specified in the list of Source Range entries divided by the LBA Data Size field (refer to Figure 98) of the LBA Format Data Structure associated with the namespace to which the command is addressed);

- if the sum of all Number of Bytes fields in all Source Range entries does not represent a multiple of the LBA Data Size field of the namespace to which the command is addressed, then the Copy command shall be aborted with a status code of Invalid Field in Command;
- if a valid Source Range Entry specifies a Number of Bytes field that represents a number of logical blocks (i.e., as determined by the LBA Data Size field) that is greater than the value in the MSSRL field (refer to Figure 97), then the Copy command shall be aborted with a status code of Command Size Limit Exceeded; and
- if the sum of all Number of Bytes fields in all Source Range entries represents a number of logical blocks (i.e., as determined by the LBA Data Size field) that is greater than the value in the MCL field (refer to Figure 97), then the Copy command shall be aborted with a status code of Command Size Limit Exceeded.

...

3.2.2.1 Command Completion

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

If the command completes with failure (i.e., completes with a status code other than Successful Completion), then the controller may or may not have copied some of the user data and Dword 0 of the completion queue entry shall contain the number of the lowest numbered Source Range entry that was not successfully copied (e.g., if Source Range 0, Source Range 1, Source Range 2, and Source Range 5 are copied successfully and Source Range 3 and Source Range 4 are not copied successfully, then Dword 0 is set to 3). If no data was written to the destination location LBAs, then Dword 0 of the completion queue entry shall be cleared to 0h.

Copy command specific errors are defined in Figure 37.

...