NVM Express® Technical Proposal (TP)

Technical Proposal ID	TP8026 Clarifying HPDA and CPDA usage with NVMeTCP PDUs
Revision Date	2023.06.28
Builds on Specification(s)	NVM Express TCP Transport Specification 1.0c TP8010a NVMe-oF Centralized Discovery Controller
References	TP8010a

Technical Proposal Author(s)

Name	Company
Erik Smith	Dell

Technical Proposal Overview

This TP resolves an NVMe/TCP interoperability problem caused by different interpretations of the HPDA and CPDA fields as defined in the NVMe/TCP Transport Specification.

To address the interoperability problem, this TP will update the definition of PDU data alignment to indicate that PDU data alignment is not used on "kickstart discovery NVMe/TCP connections"

This TP will also update two NVMe/TCP PDUs added by TP8010a; "Kickstart Discovery Request" (KDReq) and "KickStart Discovery Response" (KDResp). The definitions for both KDReq and KDResp will be modified to indicate that the HPDA and CPDA fields are now "Obsolete".

Revision History

Revision Date	Change Description
2022.12.16	Initial draft

2023.01.20	Updated proposal to indicate that HPDA and CPDA fields are not used with kickstart discovery NVMe/TCP connections
2023.02.21	Updated the "Description for Changes Document for NVM Express TCP Transport Specification 1.0c" to correctly reflect the changes being requested for the TCP Transport specification.
2023.06.28	Integrated

Description for Changes Document for NVM Express TCP Transport Specification 1.0c

New Features/Feature Enhancements/Required Changes:

- Update the definition of HPDA and CPDA fields (refer to 3.3 Data Transfer Model) and specify they only apply to non-kickstart discovery NVMe/TCP connections

Description for Changes Document for TP8010a

New Features/Feature Enhancements/Required Changes:

- Update the PDO field in the KDReq and KDResp PDUs to "Obsolete".

Markup Conventions:

Black:	Unchanged (however, hot links are removed)
~~Red Strikethrough~~:	Deleted
Blue:	New
Blue Highlighted:	TBD values, anchors, and links to be inserted in new text.
\<Green Bracketed\>:	Notes to editor

Description of Specification Changes for NVM Express TCP Transport Specification 1.0c

3.3 Data Transfer Model

…

Host and controller PDU Data is optionally aligned for non-kickstart discovery NVMe/TCP connections (refer to the Kickstart Discovery Pull Registration Requests section in the NVM Express Base Specification). PDU Data alignment is designed to allow the host or controller to guarantee that the data (and data digest) starting offset be aligned to some value (usually a cache line). The alignment of data in a PDU is specified by the host and the controller when a connection is established. The Host PDU Data Alignment (HPDA) field in the ICReq PDU specifies the required alignment of PDU Data (DATA) from the start of the PDU for PDUs that are transferred from the controller to the host. The Controller PDU Data Alignment (CPDA) field in the ICResp PDU specifies the required alignment of PDU Data (DATA) from the start of the PDU for PDUs that are transferred from the host to the controller. An appropriate number of padding bytes shall be inserted by the controller or host in the PAD field to achieve the required alignment. The number of PAD bytes is a function of the required alignment and the size of the PDU Header. PDU PAD bytes are considered as reserved bytes and are not protected by HDGST nor by DDGST. Neither the Host PDU Data Alignment field (HPDA) nor the Controller PDU Data Alignment field shall exceed 128 bytes.

Host and controller PDU Data alignment is not used for kickstart discovery NVMe/TCP connections.

Description of Specification Changes for TP8010a

3.6.2.NEW1 Kickstart Discovery Request PDU (KDReq)

Figure NEW.A: Kickstart Discovery Request PDU (KDReq)

Bytes	PDU Section	Description
00	CH	**PDU-Type:** 0Ah
01	CH	**FLAGS:** <table><tr><th>Bits</th><th>Description</th></tr><tr><td>7:2</td><td>Reserved</td></tr><tr><td>1</td><td>**DDGSTF:** If set to '1', then the DDGST field follows the PDU Data and contains a valid value. If cleared to '0', then the DDGST field is not present.</td></tr><tr><td>0</td><td>**HDGSTF:** If set to '1', then the HDGST field follows the PDU Header and contains a valid value. If cleared to '0', then the HDGST field is not present.</td></tr></table>
02	CH	**HLEN:** Fixed length of 12 bytes (i.e., Ch).
03	CH	**PDO:** ~~Data offset within PDU. This value shall be a multiple of the data alignment specified by the CPDA field in the ICResp PDU (refer to section 3.6.2.3) that was previously sent by the CDC on this TCP connection.~~ Obsolete

Figure NEW.A: Kickstart Discovery Request PDU (KDReq)

Bytes		PDU Section	Description
07:04			**PLEN:** Variable length in bytes. If HDGST and DDGST are both not present, then the length will be (NUMKR * 290 bytes) + 12 bytes. If only HDGST or only DDGST is present, then the length will be (NUMKR * 290 bytes) + 16 bytes. If HDGST and DDGST are both present, then the length will be (NUMKR * 290 bytes) + 20 bytes.
09:08		PSH	**Number of Kickstart Records (NUMKR):** This field specifies the number kickstart records included in the PDU DATA field.
11:10			**Number of Discovery Information Entries (NUMDIE):** This field specifies the maximum number of discovery information entires that the DDC is expected to return if a pull registration is requested. This field shall be cleared to 0h if a pull de-registration is being requested. Refer to the Pull Registrations and Pull De-Registrations section in the NVMe Base Specification.
HDGSTF=1	HDGSTF=0		
15:12	Not present	HDGST	**HDGST:** If HDGSTF is set to '1' in the FLAGS field, this field is valid and contains the header digest (refer to section 3.3.1.1). If the HDGSTF bit is cleared to '0', then this field is not present.
305:16	301:12	DATA	**Kickstart Record 0 (KR0):** This field specifies the first kickstart record as defined in Figure NEW.B.
…	…		…
(((N + 1) * 290) − 1) + 16:(N * 290) + 16	(((N + 1) * 290) − 1) + 12:(N * 290) + 12		**Kickstart Record N (KRN):** This field specifies the Nth kickstart record as defined in Figure NEW.B (if present).
DDGSTF=1	DDGSTF=0		
M + 3:M	Not Present	DDGST	**DDGST:** If DDGSTF is set to '1' in the FLAGS field, this field is valid and contains the data digest (refer to section 3.3.1.1). If the DDGSTF bit is cleared to '0', then this field is not present.

3.6.2.NEW2 Kickstart Discovery Response PDU (KDResp)

Figure NEW.C: Kickstart Discovery Response PDU (KDResp)

Bytes	PDU Section	Description
00	CH	**PDU-Type:** 0Bh
01	CH	**FLAGS:** <table><tr><td>**Bits**</td><td>**Description**</td></tr><tr><td>7:2</td><td>Reserved</td></tr><tr><td>1</td><td>**DDGSTF:** If set to '1', then the DDGST field follows the PDU Data and contains a valid value. If cleared to '0', then the DDGST field is not present</td></tr><tr><td>0</td><td>**HDGSTF:** If set to '1', then the HDGST field follows the PDU Header and contains a valid value. If cleared to '0', then the HDGST field is not present.</td></tr></table>
02	CH	**HLEN:** Fixed length of 10 bytes (i.e., Ah).
03	CH	**PDO:** ~~Data offset within PDU. This value shall be a multiple of the data alignment specified by the HPDA field in the ICReq PDU (refer to section 3.6.2.2) that was previously sent by the DDC on this TCP connection.~~ Obsolete
07:04	CH	**PLEN:** Fixed length of 274 bytes (i.e., 112h).
08	PSH	**Kickstart Status (KSSTAT):** <table><tr><td>**Bits**</td><td>**Description**</td></tr><tr><td>7:3</td><td>Reserved</td></tr><tr><td>1</td><td>**FAILURE:** If set to '1', then the CDC shall not perform a pull registration due to the reason indicated in the Failure Reason (FAILRSN) field.</td></tr><tr><td>0</td><td>**SUCCESS:** If set to '1', then the CDC shall perform a pull registration.</td></tr></table>
09	PSH	**Failure Reason (FAILRSN):** <table><tr><td>**Bits**</td><td>**Description**</td></tr><tr><td>7</td><td>Reserved</td></tr><tr><td>6</td><td>Insufficient Discovery Resources</td></tr><tr><td>5</td><td>TRSVCID does not match TRTYPE</td></tr><tr><td>4</td><td>TRADDR does not match ADRFAM</td></tr><tr><td>3</td><td>Invalid ADRFAM</td></tr><tr><td>2</td><td>Invalid TRTYPE</td></tr><tr><td>1</td><td>No additional information</td></tr><tr><td>0</td><td>Reserved</td></tr></table>

Figure NEW.C: Kickstart Discovery Response PDU (KDResp)

Bytes		PDU Section	Description
HDGSTF=1	HDGSTF=0		
13:10	Not Present	HDGST	**HDGST:** If HDGSTF is set to '1' in the FLAGS field, this field is valid and contains the header digest (refer to section <mark>3.3.1.1</mark>). If the HDGSTF bit is cleared to '0', then this field is not present.
269:14	265:10	DATA	**CDC NVMe Qualified Name (CDCNQN):** This field indicates the NVMe Qualified Name (NQN) that uniquely identifies the CDC. Refer to the NVMe Qualified Names section in the NVMe Base Specification for the formatting requirements of NQNs.
DDGSTF=1	DDGSTF=0		
273:270	Not Present	DDGST	**DDGST:** If DDGSTF is set to '1' in the FLAGS field, this field is valid and contains the data digest (refer to section <mark>3.3.1.1</mark>). If the DDGSTF bit is cleared to '0', then this field is not present.