Native NVMe-oF™ Drive Specification

Version 1.1

ABSTRACT: This document describes the features and functions of a storage device class known as Native NVMe-oF Drives. It includes a taxonomy covering the scope of involved device capabilities.

This document has been released and approved by the SNIA. The SNIA believes that the ideas, methodologies and technologies described in this document accurately represent the SNIA goals and are appropriate for widespread distribution. Suggestions for revisions should be directed to https://www.snia.org/feedback/.

SNIA Standard

June 17, 2021

USAGE

The SNIA hereby grants permission for individuals to use this document for personal use only, and for corporations and other business entities to use this document for internal use only (including internal copying, distribution, and display) provided that:

1. Any text, diagram, chart, table or definition reproduced shall be reproduced in its entirety with no alteration, and,

2. Any document, printed or electronic, in which material from this document (or any portion hereof) is reproduced, shall acknowledge the SNIA copyright on that material, and shall credit the SNIA for granting permission for its reuse.

Other than as explicitly provided above, you may not make any commercial use of this document or any portion thereof, or distribute this document to third parties. All rights not explicitly granted are expressly reserved to SNIA.

Permission to use this document for purposes other than those enumerated above may be requested by e-mailing tcmd@snia.org. Please include the identity of the requesting individual and/or company and a brief description of the purpose, nature, and scope of the requested use.

All code fragments, scripts, data tables, and sample code in this SNIA document are made available under the following license:

BSD 3-Clause Software License

Copyright © 2021, The Storage Networking Industry Association.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

* Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

* Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

* Neither the name of The Storage Networking Industry Association (SNIA) nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

DISCLAIMER

The information contained in this publication is subject to change without notice. The SNIA makes no warranty of any kind with regard to this specification, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. The SNIA shall not be liable for errors contained herein or for incidental or consequential damages in connection with the furnishing, performance, or use of this specification.

NVMe™ and NVMe-oF™ are trademarks of NVM Express, Inc.

Revision History

Revision	Date	Sections	Originator	Comments
1.1a	August 25, 2020	All	Mark Carlson	First draft. Adding reference to Mapping Document.
1.1b	March 9. 2021	Management	Mark Carlson	Correspond to 1.2.2 release of Swordfish
1.1c	March 9, 2021	Management	Mark Carlson	Modified in TWG meeting

Intended Audience

This document is intended for use by individuals and companies engaged in developing storage systems utilizing Native NVMe-oF Drive devices and/or related sub-systems.

Changes to the Specification

Each publication of this document is uniquely identified by a three-level identifier, comprised of a version number, a release number and an update letter. The current identifier for this document is listed on the title page of this document. Future publications of this document are subject to specific constraints on the scope of change that is permissible from one publication to the next and the degree of interoperability and backward compatibility that should be assumed between products designed to different publications of this standard. The SNIA has defined three levels of change to a specification:

- Major Revision: A major revision of the specification represents a substantial change to the underlying scope or architecture of the specification. A major revision results in an increase in the version number of the version identifier (e.g., from version 1.x.x to version 2.x x). There is no assurance of interoperability or backward compatibility between releases with different version numbers.
- Minor Revision: A minor revision of the specification represents a technical change to existing content or an adjustment to the scope of the specification. A minor revision results in an increase in the release number of the specification's identifier (e.g., from x.1.x to x.2.x). Minor revisions with the same version number preserve interoperability and backward compatibility.
- Update: An update to the specification is limited to minor corrections or clarifications of existing specification content. An update will result in an increase in the third component of the release identifier (e.g., from x.x.a to x.x.b). Updates with the same version and minor release levels preserve interoperability and backward compatibility.

About SNIA

The Storage Networking Industry Association (SNIA) is a non-profit organization made up of member companies spanning information technology. A globally recognized and trusted authority, SNIA's mission is to lead the storage industry in developing and promoting vendor-neutral architectures, standards and educational services that facilitate the efficient management, movement and security of information. For additional information, visit the SNIA web site at www.snia.org.

Acknowledgements

The SNIA Object Storage Technical Working Group (TWG), which developed and reviewed this document, would like to recognize the significant contributions made by the following members:

Mark Carlson ..... Kioxia
John Geldman ..... Kioxia

Contents

List of Tables and Figures

1 Introduction

The growing popularity of NVMe and NVMe-oF storage has resulted in the development of Ethernet-connected storage devices, herein referred to as Native NVMe-oF Drives and subsystems supporting this standard interface. This document specifies the interfaces to these devices.

Some of the aspects in this document may be met by various designs that fall outside of a typical drive form factor and yet may still interoperate at some level, such as an interposer for PCIe drives. Examples are an interposer that has two different interfaces (e.g. converting RoCE to PCIe) or an entire enclosure comprised of standard PCIe storage devices fronted by an NVMe-oF interface processing unit that proxies the individual drives.

2 Normative References

Table 1 lists the standards, specifications and other documents related to this document.

Table 1 - References

Author/Owner	Title	Revision	URL
SNIA	TLS Specification for Storage Systems	1.0.1	http://www.snia.org/sites/default/files/TLSspec-v1.1_Technical_Position.pdf
SNIA	Swordfish NVMeMappingGuide.pdf	1.2.2	Swordfish v1.2.2_NVMeMappingGuide
SNIA	Swordfish_Profiles.zip	1.2.2	Swordfish_v1.2.2_Profiles.zip
IEEE	IEEE 802.3-2018		http://www.ieee802.org
IETF	RFC 4171 - iSNS		https://www.ietf.org/rfc.html
IETF	RFC 791 - IPV4		https://www.ietf.org/rfc.html
IETF	RFC 2460 - IPV6		https://www.ietf.org/rfc.html
IETF	RFC 2131 – DHCP		https://www.ietf.org/rfc.html
IETF	RFC 2132 - DHCP Options and BOOTP		https://www.ietf.org/rfc.html
IETF	RFC 1035 - DNS		https://www.ietf.org/rfc.html
IETF	RFC 6762 Multi-case Domain name Service		https://www.ietf.org/rfc.html
IETF	RFC 6763 – mDNS/DNS-SD		https://www.ietf.org/rfc.html
DMTF	Redfish		http://www.dmtf.org/standards/redfish
SFF Committee	SFF-9639 – Multifunction 6X Unshielded Connector Pinouts	2.1	https://www.snia.org/sff
SFF Committee	REF-TA-1012 – Pin Assignment Reference for SFF-TA-1002 Connectors	1.0	https://www.snia.org/sff
SFF Committee	SFF-8201 – 2.5” Drive Form Factor Dimensions		https://www.snia.org/sff
SFF Committee	SFF-8223 – 2.5” Drive Form Factor with Serial Connector		https://www.snia.org/sff
SFF Committee	SFF-8301 – 3.5” Drive Form Factor Dimensions		https://www.snia.org/sff
SFF Committee	SFF-8323 – 3.5” Drive Form Factor with Serial Connector		https://www.snia.org/sff
SFF Committee	SFF-8639 - Multifunction 6X Unshielded Connector		https://www.snia.org/sff
SFF Committee	SFF-TA-1001- Universal x4 Link Definition for SFF-8639		https://www.snia.org/sff
SFF Committee	SFF-TA-1009 - Enterprise and Datacenter SSD Pin and Signal Specification		https://www.snia.org/sff

3 Scope

This document focuses on interfaces for drive (e.g. SSD) form factors. Larger and smaller form factors and varying physical connectivity are not precluded. Additionally, there is no bias or advocacy toward any particular data persistence technology.

This document is divided into several sections. The first lists key definitions, conventions, etc. The second describes fundamental attributes of Native NVMe-oF Drives. The third details management functionality.

4 Definitions, Symbols, Abbreviations, and Conventions

4.1 Overview

For the purposes of this document, the terms and definitions given in the SNIA 2019 Dictionary (www.snia.org/education/dictionary) apply. In cases where the current definitions in the SNIA dictionary conflict with those presented in this document, the definitions in this document have precedence.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- IEC Electropedia: available at http://www.electropedia.org/
- ISO Online browsing platform: available at http://www.iso.org/obp

4.2 Definitions

4.2.1

key value storage

Type of object storage interface where a key is used to address the associated object

4.2.2

object

Encapsulation of data and associated meta-data

4.2.3

Native NVMe-oF Drive

Storage element that directly implements the NVMe-oF interface standard

4.3 Acronyms and Abbreviated Terms

API	Application Programming Interface
ATA	Advanced Technology Attachment
CDMI	Cloud Data Management Interface
DAS	Direct Attached Storage
DHCP	Dynamic Host Configuration Protocol
DMTF	Distributed Management Task Force
DNS	Domain Name Service
DSaaS	Data Storage as a Service
GbE	Gigabit Ethernet
HDFS	Hadoop Distributed File System
HTTP	HyperText Transfer Protocol
I2C	Inter-Integrated Circuit
IEC	International Electrotechnical Commission
INCITS	International Committee on Information Technology Standards
IP	Internet Protocol (Alt. Intellectual Property)

ISO	International Standards Organization
LLDP	Link Level Discovery Protocol
KV	Key Value
NTP	Network Time Protocol
NVM	Non-Volatile Memory
NVMe	Non-Volatile Memory Express
NVMe-oF	NVMe over Fabrics
OS	Operating System
PCIe	Peripheral Component Interface Express
RFC	Request For Comment
SAS	Serial Attached SCSI
SATA	Serial ATA Interface
SCSI	Small Computer Systems Interface
SFF	Small Form Factor
SNIA	Storage Network Industry Association
SSD	Solid State Drive
SSDP	Simple Service Discovery Protocol
SW	Software
T10	INCITS SCSI Technical Committee
T13	INCITS AT Attachment Technical Committee
TCP	Transmission Control Protocol
TLS	Transport Level Specification
TWG	Technical Working Group
URL	Uniform Resource Locator

4.4 Keywords

4.4.1

expected

A keyword used to describe the behavior of the hardware or software in the design models presumed by this standard. Other hardware and software design models may also be implemented.

mandatory

A keyword indicating an item that is required to be implemented as defined in this document to claim compliance with this document.

may

A keyword that indicates flexibility of choice with no implied preference.

obsolete

A keyword indicating that an item was defined in prior revisions to this document but has been removed from this revision.

optional
A keyword that describes features that are not required to be operational. However, if any optional feature is operational, it shall be implemented as defined in this document.

prohibited
A keyword used to describe a feature or behavior that is not allowed to be present.

required
A keyword used to describe a behavior that shall be implemented.

reserved
A keyword referring to bits, bytes, words, fields, and code values that are set aside for future standardization.

Note 1 to entry: A reserved bit, byte, word or field shall be set to zero, or in accordance with a future extension to this standard.

Note 2 to entry: Recipients are not required to check reserved bits, bytes, words or fields for zero values; receipt of reserved code values in defined fields shall be reported as an error.

shall
A keyword indicating a mandatory requirement.

should
A keyword indicating flexibility of choice with a preferred alternative; equivalent to the phrase “it is recommended”.

4.5 Conventions

Certain words and terms used in this document have a specific meaning beyond their normal English meaning. These words and terms are defined either in 4.2 or in the text where they first appear.

Numbers that are not immediately followed by lower-case b or h are decimal values.

Numbers immediately followed by lower-case b (xxb) are binary values.

Numbers immediately followed by lower-case h (xxh) are hexadecimal values.

Hexadecimal digits that are alphabetic characters are upper case (i.e., ABCDEF, not abcdef).

Hexadecimal numbers may be separated into groups of four digits by spaces. If the number is not a multiple of four digits, the first group may have fewer than four digits (e.g., AB CDEF 1234 5678h)

Storage capacities shall be reported in base-10. IO transfer sizes and offsets shall be reported in base-2. The associated units and abbreviations used in this document are:

- A kilobyte (KB) is equal to 1,000 (10^3) bytes.
- A megabyte (MB) is equal to 1,000,000 (10^6) bytes.
- A gigabyte (GB) is equal to 1,000,000,000 (10^9) bytes.
- A terabyte (TB) is equal to 1,000,000,000,000 (10^{12}) bytes.
- A petabyte (PB) is equal to 1,000,000,000,000,000 (10^{15}) bytes.
- An exabyte (EB) is equal to 1,000,000,000,000,000,000 (10^{18}) bytes.
- A kibibyte (KiB) is equal to 2^{10} bytes.

- A mebibyte (MiB) is equal to 2^{20} bytes.
- A gibibyte (GiB) is equal to 2^{30} bytes.
- A tebibyte (TiB) is equal to 2^{40} bytes.
- A pebibyte (PiB) is equal to 2^{50} bytes.
- An exibyte (EiB) is equal to 2^{60} bytes.

5 Native NVMe-oF Drive Characteristics and Requirements

5.1 Physical Layer – Form Factor

Drives may be delivered in different standard physical dimensions. Example form factors as defined in the references noted in Table 1 include:

- SFF-8201 2.5” Drive Form Factor Dimensions;
- SFF-8223 2.5” Drive Form Factor with Serial Connector;
- SFF-8301 3.5” Drive Form Factor Dimensions; or
- SFF-8323 3.5” Drive Form Factor with Serial Connector.
- SFF-100x A series of form factors all using the SFF-1002 Connector

5.2 Electrical

This document is for Ethernet connected media. Other media are possible and would not be outside the definition of a Native NVMe-oF Drive.

A standard for the modified pin-out of SFF-8639 (Serial Attachment 12 Gb/s 6X Unshielded Connector) is documented here and referenced in SFF-9639.

A standard for the modified pin-out of SFF-1002 is documented here and referenced in REF-TA-1012.

Native NVMe-oF Drives shall implement the pin out in the SFF-9639 (overlay of SFF-8639) Native NVMe-oF Ethernet Drive.

The utilization of SMBus and Power Disable between chassis and drives based on type is shown in Table 2.

Table 2 - Utilization of I2C/SMBus and Power Disable between chassis and drives based on type

	SNIA Native NVMe-oF Drive
SNIA Native NVMe-oF Drive chassis	Drive and/or chassis may support Power Disable and/or SMBus

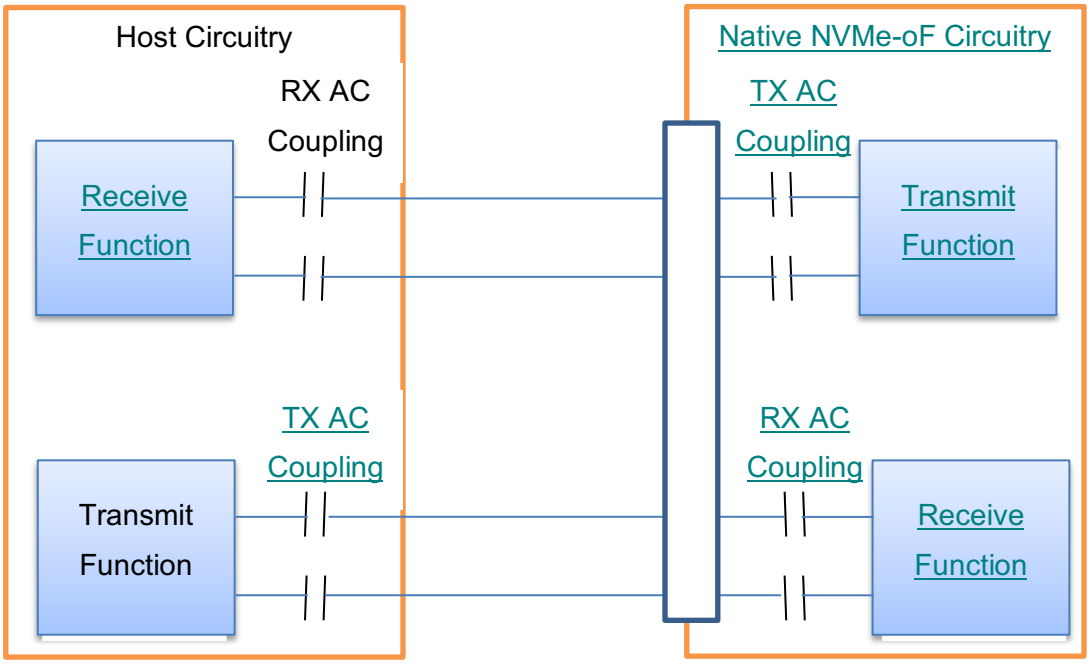
IEEE 802.3-2018 defines AC Coupling for the RX lines associated with the receive function of both the host circuitry and the Native NVMe-oF device circuitry. See the IEEE 802.3-2018 specification for further definition of the AC coupling on the RX lines of a receive function.

For Native NVMe-oF Drives, there shall be AC coupling capacitors on the TX lines associated with transmit functions of Native NVMe-oF devices. These AC coupling capacitors should be sized so that the effective capacitance of the combination of the TX line AC coupling capacitance and the RX line AC coupling capacitance is 90% or greater than the RX line AC Coupling capacitance.

For Native NVMe-oF host circuitry, there should be AC coupling capacitors on the TX lines of the host’s Native NVMe-oF transmit functions. These AC coupling capacitors should be sized so that the effective capacitance of the combination of the TX line AC coupling capacitance and the RX line AC coupling capacitance is 90% or greater than the RX line AC Coupling capacitance.

The AC coupling capacitors on the TX lines provide protection against mis-insertion (e.g., insertion of a Native NVMe-oF SSD into a non-Ethernet host port, or a non-Native NVM-oF into a Native NVMe-oF host).

Figure 1 – Example of AC Coupling for one NVMe-oF lane



5.3 Link Layer

A Native NVMe-oF Drive shall support the NVM Express Base specification, the NVMe-oF specification, and the IEEE 802.3-2018 specification.

A Native NVMe-oF Drive shall support operation at 25 Gb/s, and may support higher speeds in any combination. If more than one speed is supported, then a Native NVMe-oF Drive shall perform speed auto-negotiation at power-on as defined by the applicable transport committee (e.g., IEEE 802.3). The goal of speed negotiation is to achieve the highest speed possible by both sides of a connection.

Native NVMe-oF Drives:

- May implement manual link speed assignment (run at other than negotiated speed), and
- Shall implement MDNS (Multicast DNS)

5.4 Environmental

This document specifies a discovery and reporting mechanism to provide information on device environmental characteristics (e.g. power and temperature).

5.5 Taxonomy

A Native NVMe-oF Drive may be defined in many forms. This taxonomy lists several forms with the idea that innovation could expand this list over time.

A tabular taxonomy is shown in Table 3.

Table 3 - Tabular taxonomy

Drive Type	Host Connect	Protocol Examples	Management Defined In…	Comments
Block Device	NVMe	NVMe	NVM Express	Low level interface. Complex host stack. Host sees block interface (PMR, …). Typically single host/client.
Key-Value (KV)	NVMe	NVMe-KV	NVM Express	High level KV interface. Simplified host stack. Multi-host/client. Failover for availability.
Interposer (addition to drive)	NVMe-oF	NVMe-oF	NVM Express	May be short term solution prior to embedded functionality.

6 Pin-outs

6.1 SFF-TA-1002 based pinouts

The following table 4 lists the signal to pin-out relationships for several standards that share the SFF-TA-1002 connector. The signals labeled SFF-TA-ETH are defined by this specification.

Table 4 - EDSFF X4 SFF-TA-ETH (and SFF-TA-1009) DEVICE EDGE PINOUT

Pin	Contact Sequence	Signal SFF-TA-1009 *1	Signal SFF-TA-ETH *2	Signal SFF-TA-ETH *2	Signal SFF-TA-1009 *1	Contact Sequence	Pin
A28	1st mate	GND	GND	GND	GND	1st mate	B28
A27	2nd mate	PERp3	ENRp3	ENTp3	PETp3	2nd mate	B27
A26	2nd mate	PERn3	ENRn3	ENTn3	PETn3	2nd mate	B26
A25	1st mate	GND	GND	GND	GND	1st mate	B25
A24	2nd mate	PERp2	ENRp2	ENTp2	PETp2	2nd mate	B24
A23	2nd mate	PERn2	ENRn2	ENTn2	PETn2	2nd mate	B23
A22	1st mate	GND	GND	GND	GND	1st mate	B22
A21	2nd mate	PERp1	ENRp1	ENTp1	PETp1	2nd mate	B21
A20	2nd mate	PERn1	ENRn1	ENTn1	PETn1	2nd mate	B20
A19	1st mate	GND	GND	GND	GND	1st mate	B19
A18	2nd mate	PERp0	ENRp0	ENTp0	PETp0	2nd mate	B18
A17	2nd mate	PERn0	ENRn0	ENTn0	PETn0	2nd mate	B17
A16	1st mate	GND	GND	GND	GND	1st mate	B16
A15	2nd mate	REFCLKp1			REFCLKp0	2nd mate	B15
A14	2nd mate	REFCLKn1			REFCLKn0	2nd mate	B14
A13	1st mate	GND	GND	GND	GND	1st mate	B13
A12	2nd mate	PRSNT0#	PRSNT0#	PWRDIS	PWRDIS	2nd mate	B12
A11	2nd mate	PERST1#/ CLKREQ#	ENRST1#	3.3 Vaux	3.3 Vaux	2nd mate	B11
A10	2nd mate	LED/ ACTIVITY	LED/ ACTIVITY	ENRST0#	PERST0#	2nd mate	B10
A9	2nd mate	SMBRST#	SMBRST#	DUALPORT EN#	DUALPORT EN#	2nd mate	B9
A8	2nd mate	SMBDAT	SMBDAT	RFU	RFU	2nd mate	B8
A7	2nd mate	SMBCLK	SMBCLK	MFG	MFG	2nd mate	B7
A6	1st mate	GND	GND	12 V	12 V	2nd mate	B6
A5	1st mate	GND	GND	12 V	12 V	2nd mate	B5
A4	1st mate	GND	GND	12 V	12 V	2nd mate	B4
A3	1st mate	GND	GND	12 V	12 V	2nd mate	B3
A2	1st mate	GND	GND	12 V	12 V	2nd mate	B2
A1	1st mate	GND	GND	12 V	12 V	2nd mate	B1

*1 For Reference only.

*2 The TXx- and the TXx+ pins shall be connected to the Ethernet Receiver differential pairs on the EDSFF X4 SFF-TA -ETH device. The RXx- and the RXx+ pins shall be connected to the Ethernet Transmitter differential pairs on the EDSFF X4 SFF-TA -ETH device.

Table 5 - EDSFF X8 SFF-TA-ETH (and SFF-TA-1009) DEVICE EDGE PINOUT

Pin	Contact Sequence	Signal SFF-TA-1009 *1	Signal SFF-TA-ETH *2	Signal SFF-TA-ETH *2	Signal SFF-TA-1009 *1	Contact Sequence	Pin
A42	2nd mate	RFU	RFU	PRSNT1#	PRSNT1#	2nd mate	B42
A41	1st mate	GND	GND	GND	GND	1st mate	B41
A40	2nd mate	PERp7	ENRp7	ENTp7	PETp7	2nd mate	B40
A39	2nd mate	PERn7	ENRn7	ENTn7	PETn7	2nd mate	B39
A38	1st mate	GND	GND	GND	GND	1st mate	B38
A37	2nd mate	PERp6	ENRp6	ENTp6	PETp6	2nd mate	B37
A36	2nd mate	PERn6	ENRn6	ENTn6	PETn6	2nd mate	B36
A35	1st mate	GND	GND	GND	GND	1st mate	B35
A34	2nd mate	PERp5	ENRp5	ENTp5	PETp5	2nd mate	B34
A33	2nd mate	PERn5	ENRn5	ENTn5	PETn5	2nd mate	B33
A32	1st mate	GND	GND	GND	GND	1st mate	B32
A31	2nd mate	PERp4	ENRp4	ENTp4	PETp4	2nd mate	B31
A30	2nd mate	PERn4	ENRn4	ENTn4	PETn4	2nd mate	B30
A29	1st mate	GND	GND	GND	GND	1st mate	B29
A28	1st mate	GND	GND	GND	GND	1st mate	B28
A27	2nd mate	PERp3	ENRp3	ENTp3	PETp3	2nd mate	B27
A26	2nd mate	PERn3	ENRn3	ENTn3	PETn3	2nd mate	B26
A25	1st mate	GND	GND	GND	GND	1st mate	B25
A24	2nd mate	PERp2	ENRp2	ENTp2	PETp2	2nd mate	B24
A23	2nd mate	PERn2	ENRn2	ENTn2	PETn2	2nd mate	B23
A22	1st mate	GND	GND	GND	GND	1st mate	B22
A21	2nd mate	PERp1	ENRp1	ENTp1	PETp1	2nd mate	B21
A20	2nd mate	PERn1	ENRn1	ENTn1	PETn1	2nd mate	B20
A19	1st mate	GND	GND	GND	GND	1st mate	B19
A18	2nd mate	PERp0	ENRp0	ENTp0	PETp0	2nd mate	B18
A17	2nd mate	PERn0	ENRn0	ENTn0	PETn0	2nd mate	B17
A16	1st mate	GND	GND	GND	GND	1st mate	B16
A15	2nd mate	REFCLKp1			REFCLKp0	2nd mate	B15
A14	2nd mate	REFCLKn1			REFCLKn0	2nd mate	B14
A13	1st mate	GND	GND	GND	GND	1st mate	B13
A12	2nd mate	PRSNT0#	PRSNT0#	PWRDIS	PWRDIS	2nd mate	B12
A11	2nd mate	PERST1#/ CLKREQ#	ENRST1#	3.3 Vaux	3.3 Vaux	2nd mate	B11
A10	2nd mate	LED/ ACTIVITY	LED/ ACTIVITY	ENRST0#	PERST0#	2nd mate	B10
A9	2nd mate	SMBRST#	SMBRST#	DUALPORTEN#	DUALPORTEN#	2nd mate	B9
A8	2nd mate	SMBDAT	SMBDAT	RFU	RFU	2nd mate	B8
A7	2nd mate	SMBCLK	SMBCLK	MFG	MFG	2nd mate	B7
A6	1st mate	GND	GND	12 V	12 V	2nd mate	B6
A5	1st mate	GND	GND	12 V	12 V	2nd mate	B5
A4	1st mate	GND	GND	12 V	12 V	2nd mate	B4
A3	1st mate	GND	GND	12 V	12 V	2nd mate	B3
A2	1st mate	GND	GND	12 V	12 V	2nd mate	B2
A1	1st mate	GND	GND	12 V	12 V	2nd mate	B1

*1 For Reference Only

*2 The TXx- and the TXx+ pins shall be connected to the Ethernet Receiver differential pair on the EDSFF X8 SFF-TA-ETH device. The RXx- and the RXx+ pins shall be connected to the Ethernet Transmitter differential pair on the EDSFF X8 SFF-TA-ETH device.

6.2 SFF-8639 based pinout

The following table 4 lists the signal to pin-out relationships for several standards that share the SFF-TA-8639 connector. The signals labeled SFF-TA-ETH are defined by this specification. The SFF-TA-9639 columns for SNIA Ethernet and OCP Kinetic are for historical reference only and are not discussed in this standard.

Table 6 - SFF-9639 P Plug with ETH

		Quad PCIe *1	**SFF-TA-1001 *1**	**SFF-9639-ETH** *2
P1	2nd Mate	WAKE#	WAKE#	
P2	2nd Mate		Reserved	
P3	1st Mate	PWRDIS	PWRDIS	PWRDIS
P4*	1st Mate	IfDet#	IfDet#	IfDet#
P5	1st Mate	Ground	Ground	Ground
P6	1st Mate	Ground	Ground	Ground
P7	1st Mate			
P8	2nd Mate			
P9	2nd Mate			
P10	1st Mate	PRSNT#	PRSNT#	PRSNT#
P11	2nd Mate	ACTIVITY#	ACTIVITY#	ACTIVITY#
P12	1st Mate	Ground	Ground	Ground
P13	1st Mate	+12 V Precharge	+12 V Precharge	+12 Precharge
P14	2nd Mate	+12 V	+12 V	+12 V
P15	2nd Mate	+12 V	+12 V	+12 V

*1 For Reference Only

*2 The TXx- and the TXx+ pins shall be connected to the Ethernet Receiver differential pair on the SFF9639-ETH device. The RXx- and the RXx+ pins shall be connected to the Ethernet Transmitter differential pair on the SFF9639-ETH device.

Table 7 - SFF-9639 S Plug with ETH

		Quad PCIe *1	SFF-TA-1001 *1	SFF-9639-ETH *2
S1	1st Mate	Ground	Ground	Ground
S2	2nd Mate		PETp0	
S3	2nd Mate		PETn0	
S4	1st Mate	Ground	Ground	Ground
S5	2nd Mate		PERn0	
S6	2nd Mate		PERp0	
S7	1st Mate	Ground	Ground	Ground
S8	1st Mate	Ground	Ground	Ground
S9	2nd Mate		PETp1	
S10	2nd Mate		PETn1	
S11	1st Mate	Ground	Ground	Ground
S12	2nd Mate		PERn1	
S13	2nd Mate		PERp1	
S14	1st Mate	Ground	Ground	Ground
S15	2nd Mate	Reserved	HPT0	HPT0
S16	1st Mate	Ground	Ground	Ground
S17	2nd Mate	PETp1	PETp2	TX0+
S18	2nd Mate	PETn1	PETn2	TX0-
S19	1st Mate	Ground	Ground	Ground
S20	2nd Mate	PERn1	PERn2	
S21	2nd Mate	PERp1	PERp2	
S22	1st Mate	Ground	Ground	Ground
S23	2nd Mate	PETp2	PETp3	TX1+
S24	2nd Mate	PETn2	PETn3	TX1-
S25	1st Mate	Ground	Ground	Ground
S26	2nd Mate	PERn2	PERn3	
S27	2nd Mate	PERp2	PERp3	
S28	1st Mate	Ground	Ground	Ground

*1 For Reference Only

*2 The TXx- and the TXx+ pins shall be connected to the Ethernet Receiver differential pair on the SFF9639-ETH device. The RXx- and the RXx+ pins shall be connected to the Ethernet Transmitter differential pair on the SFF9639-ETH device.

Table 8 - SFF-9639 E Plug with ETH

		Quad PCIe *1	SFF-TA-1001 *1	SFF-9639-ETH *2
E1	2nd Mate	REFCLKB+	REFCLKB+	
E2	2nd Mate	REFCLKB-	REFCLKB-	
E3	2nd Mate	+3.3 Vaux	+3.3V aux	+3.3V aux
E4	2nd Mate	CLKREQ#/ PERSTB#	PERSTB#	ENRST1#
E5	2nd Mate	PERST#	PERST#	ENRST#
E6	2nd Mate	Reserved	IFDET2#	IFDET2#
E7	2nd Mate	REFCLK+	REFCLK+	
E8	2nd Mate	REFCLK-	REFCLK-	
E9	1st Mate	Ground	Ground	Ground
E10	2nd Mate	PETp0		
E11	2nd Mate	PETn0		
E12	1st Mate	Ground	Ground	Ground
E13	2nd Mate	PERn0		RX0-
E14	2nd Mate	PERp0		RX0+
E15	1st Mate	Ground	Ground	Ground
E16	2nd Mate	Reserved	HPT1	HPT1
E17	2nd Mate	PETp3		
E18	2nd Mate	PETn3		
E19	1st Mate	Ground	Ground	Ground
E20	2nd Mate	PERn3		RX1-
E21	2nd Mate	PERp3		RX1+
E22	1st Mate	Ground	Ground	Ground
E23	2nd Mate	SMBCLK	SMBCLK	SMBCLK
E24	2nd Mate	SMBDAT	SMBDAT	SMBDAT
E25	2nd Mate	DualPort En#	DualPort En#	DualPort En#

*1 For Reference Only

*2 The TXx- and the TXx+ pins shall be connected to the Ethernet Receiver differential pair on the SFF9639-ETH device. The RXx- and the RXx+ pins shall be connected to the Ethernet Transmitter differential pair on the SFF9639-ETH device.

6.3 SNIA Ethernet Specific Signal Definitions

6.3.1 DUALPORTEN# (Host Output)

This signal is reserved for Native NVMe-oF devices. See section 7.3 for link aggregation requirements.

6.3.2 ENRST0# (Host Output)

ENRST0# is a functional reset of the Ethernet interface on lane 0 (and any linked lanes) as well as associated controllers. ENRST0# should not affect other supported Ethernet interface lanes if the Native NVMe-oF device has multiple MAC addresses active.

6.3.3 ENRST1# (Host Output)

ENRST1# is a functional reset of the Ethernet interface on lane 1 (and any linked lanes) as well as associated controllers. ENRST1# should not affect other supported Ethernet interface lanes if the Native NVMe-oF device has multiple MAC addresses active.

6.3.4 HPT0 (Host Output)

If left Open by the host, then this indicates support for SFF-8639-ETH or PCI Quad (see SFF-TA-1001). If set to Ground by the host, then a SFF-8639-ETH device shall not enable it's Ethernet interface.

6.3.5 HPT1 (Host Output)

If left Open by the host, then this indicates support for SFF-8639-ETH or PCI Quad (see SFF-TA-1001). If set to Ground by the host, then a SFF-8639-ETH device shall not enable it's Ethernet interface.

6.3.6 IFDET# (Host Input)

Shall be set to Ground by a Native NVMe-oF device.

6.3.7 IFDET2# (Host Input)

Shall be set to Ground by a Native NVMe-oF device.

6.3.8 PRSTNT#, PRSNT0# (Host Input)

These signals shall be set to Ground by a Native NVMe-oF device.

7 Native NVMe-oF Drive Management

This section describes management functionality at the device level. This can be expanded to include subsystem enclosures and multi-rack configurations.

7.1 Overview

There are two aspects to device discovery and management:

1. Management Discovery;
2. Service Discovery.

Management Discovery is described in this section. Service Discovery is the responsibility of the service (application) operating on the device and is beyond the scope of this document.

Management Discovery has the following structure and sequence further described in this section:

- Establish Physical link (Ethernet, LACP, LLDP);
- Assign IP Address (DHCP);
- Discover Basic Services (DNS-Name, NTP-Time);
- Provide Management Services (Redfish).

7.2 Establish Physical Link

It is assumed that a physical Ethernet link has been established. See section 5.3.

7.3 Assign IP Address

Drives shall support IPv4 and IPv6 for the management endpoint.

Drives shall support DHCP V4 (see RFC 2131) and DHCP V6 (see RFC 8415).

Drives may support factory configured IP addresses, IP address configuration via I2C/SMBus, or Stateless Address Auto-configuration (see RFC 4862).

If more than one port is implemented, such ports may be configured for redundancy or other purposes.

Link aggregation, if needed may be done through LACP and as such is not part of this document.

7.4 Discover Basic Services

The discovery of Native NVMe-oF Drive basic services involves finding its access point on a network and perhaps its position in whatever enclosure it is located. It also involves determining the Drive's capabilities.

Redfish references SSDP (Simple Service Discovery Protocol) for discovery, and time synchronizing configuration via a NTP Server configured via DHCP (RFC 2132 Section 8.3).

Native NVMe-oF Drives should use the Redfish discovery mechanism.

7.5 Redfish Manageability

Native NVMe-oF devices shall support manageability through Redfish. Implementations shall implement the SwordfishNVMeEthernetAttach.json, SwordfishNVMeDrive.json profiles and may

implement SwordfishNVMeDriveAdvancedFeatures.json profile as specified in Swordfish Profiles.zip or above.

Implementations should follow the guidance in Swordfish NVMeMappingGuide.pdf or above.