WHITE PAPER

July 2023

Overview of EDSFF E1.S Form Factors

Co-Authors



Ross Stenfort, Meta™

Jason Adrian, Microsoft® Azure®

Maulik Sompura, KIOXIA

Industry Landscape

The main building block of cloud computing is the hyperscale data center - an infrastructure built with simplicity in mind that includes in its design: management, hot swapping, SKU reductions, front panel serviceability, and more. Hyperscale data centers enable cloud providers to deliver efficient, scalable and cost-effective platforms for addressing today's big data needs. When compared to traditional data centers, they differ in how they manage storage drives to meet customer and application requirements, reduce customer storage acquisition costs and enable customer scalability on an as needed basis. Hyperscale architectures are expanding in use and size, and more companies are adopting them.

E1.S form factors provide many benefits that improve upon the capabilities of the 2.5-inch and M.2 form factors.

Solid state drives (SSDs) are an integral component of hyperscale data center storage and used primarily for fast system boot-ups and highly performant data storage. Drives based on 2.5-inch[1] and M.2 form factors are widely used in hyperscale architectures but are nearing their functional capabilities relating to drive performance and latency, power loss protection[2], drive serviceability, signal integrity, power envelop delivery, thermal management, security, Quality of Service (QoS), and others. The Enterprise and Datacenter Standard Form Factor (EDSFF) specifications were developed to address the industry needs for simplicity and architect a new set of form factors for future data centers. From this effort, the SFF Technical Affiliate (SFF-TA) working group of the Storage Networking Industry Association (SNIA) developed E1.S form factor specifications. These new E1.S form factor SSDs have been adopted by the Open Compute Project (OCP), who in turn authored the OCP Datacenter NVMe™ SSD specification.

E1.S SSDs are designed for flash memory chips and leverage PCIe® interface and NVMe protocol technologies.

E1.S form factors provide many benefits that improve upon the capabilities of the 2.5-inch and M.2 form factors. E1.S SSDs are designed for flash memory chips and leverage PCIe® interface and NVMe protocol technologies. Design benefits include power/performance, storage density, thermal management, and more. Initial E1.S SSDs are based on the PCIe 4.0 interface, but EDSFF SSDs are designed for PCIe 5.0 and 6.0 generations, and possibly beyond. As such, SSDs with E1.S form factors are projected to grow from 7.2% of the total exabyte share of combined PCIe form factors in 2022 to 25.9% market share in 2027 (Figure 1). Additionally, E1.S SSDs are projected to grow from 8% of the PCIe units shipped in 2022 to 40.4% in 2027 (Figure 2). In both market projections, there is downward growth of 2.5-inch (U.2) and M.2 SSDs, while E1.S SSDs show significant growth.

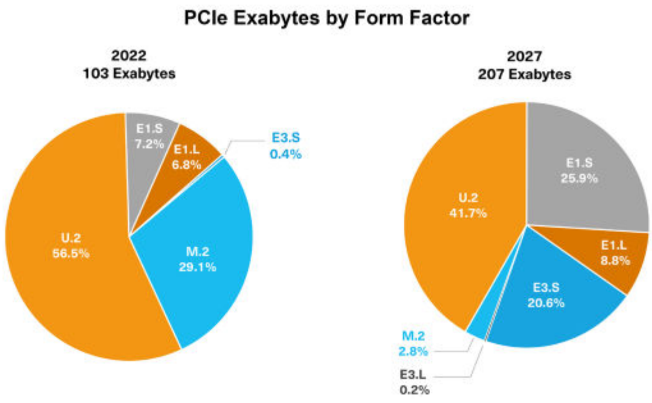

Figure 1: PCIe Form Factor Comparison, by Exabytes, from 2022 to 2027.
Source: TrendFocus, March 2023

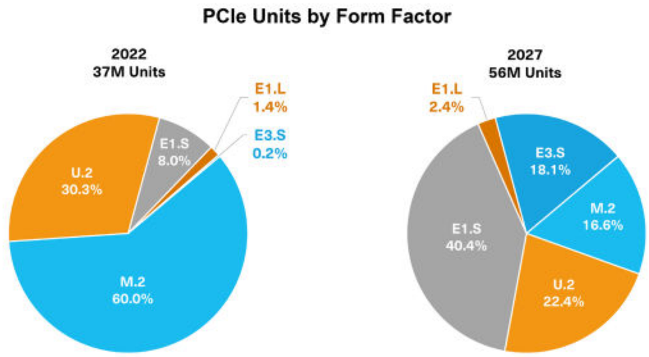

Figure 2: PCIe Form Factor Comparison, by Units Shipped, from 2022 to 2027.
Source: TrendFocus, March 2023.

For many companies with traditional data centers who are planning to move to a hyperscale architecture, OCP systems and architectures can provide benefits and economies of scale as required by the largest hyperscalers. OCP takes an open source, industry collaborative approach to the data center landscape, addressing such storage challenges as drive performance and latency, power loss protection, drive serviceability, signal integrity, efficient power use, thermal management, security, and Quality of Service (QoS). OCP evaluated and standardized on E1.S form factors developed by the SNIA SFF-TA working group. The latest generation of OCP servers are designed with SSDs that comply with the OCP Datacenter NVMe SSD Specification.

Leading hyperscalers prefer E1.S designs on new and upcoming platforms setting the precedence for industry-wide adoption of the E1.S form factor, particularly for 1U deployments. This paper introduces E1.S form factors and the benefits they provide over 2.5-inch and M.2 drives.

How Can SSD Form Factors Evolve?

New server and storage system designs are becoming more use case-specific and require data storage that can easily adapt to optimizations in performance or capacity (or a hybrid of both). With advancements in faster networking protocols, memory bandwidth and interconnect technologies, coupled with more robust interfaces, such as PCIe 4.0, 5.0, 6.0 and beyond, data storage in 2.5-inch and M.2 drive formats are limiting future capabilities of flash storage in data centers. These currently predominant formats are challenged to keep technological pace with new server and storage demands setting the criteria for a new form factor architecture as outlined below:

2.5-inch Form Factors

Power / Performance	The 2.5-inch form factor originated with hard disk drives and is not optimal for flash memory as it cannot scale beyond four PCIe lanes (x8 and x16).	WHAT'S NEEDED	A form factor that can scale as both power and PCIe speed increases - and provide wider PCIe link widths to enable full-throttle IOPS performance per terabyte[3] (TB) of capacity.
Thermal Capabilities	SSD performance and capacity scaling has a domino effect on SSD power. The existing 2.5-inch form factor design does not provide airflow benefits for flash-based SSDs.	WHAT'S NEEDED	A form factor with better heat dissipation that enables a system to operate at lower airflow levels.
Signal Integrity Issues	Connectors used with 2.5-inch form factors are typically not designed to meet the signal integrity challenges that may be apparent in the upcoming PCIe 6.0 interface specifications.	WHAT'S NEEDED	A form factor that is future-proofed for the next generation of high-frequency interfaces such as PCIe 6.0 and beyond.

M.2 Form Factors

The M.2 form factor brings SSD storage to the size of a stick of gum and is deployed in servers and client notebooks/laptops as primary storage and boot drives. M.2 connectors plug directly into the system's motherboard, so an M.2 connector slot is required for interfacing these SSDs, but no extra or special M.2 cabling is required. This slim form factor offers clean, uncluttered server and PC interiors that occupy less real estate and require only a few cables versus 2.5-inch drives. However, M.2 SSDs also have limitations that need to be addressed:

Power / Performance	Performance on a typical M.2 SSD is usually limited at 8.25W (as of this publication), and is especially seen with higher capacities such as 4 TB and beyond. This poses a problem in support of PCIe 4.0 as M.2 SSDs cannot achieve full PCIe 4.0 16 (gigatransfers per second (GT/s) x4) bandwidth with an 8.25W power envelope.	WHAT'S NEEDED	A form factor with supported SSDs that can achieve full PCIe 4.0 (16 GT/s x4) bandwidth and 'soon-to-be-available' PCIe 5.0 (32 GT/s x4) bandwidth and even PCIe 6.0 (64 GT/s x4) bandwidth with support for up to a 25W power envelope depending on the drive's capacity.
Thermal Capabilities	SSDs need to operate in temperature ranges that require continued operations despite unreasonable thermal conditions*. *Note: Thermal tolerances may vary.	WHAT'S NEEDED	A form factor with heatsink options built into the SSD, enabling selection of a thermal profile best suited to cool a system. Also, thermal throttling mechanisms that detect if an SSD has reached an operating temperature beyond its specification, throttling performance and reducing heat generated by flash memory die activation.
Physical Serviceability	To replace an M.2 SSD, an entire server would normally need to be powered down - a time-consuming process that requires removing the system from the rack, removing the M.2 SSD, inserting a new M.2 SSD and reconfiguring the assembly.	WHAT'S NEEDED	A form factor that enables supported SSDs to be hot-swapped from the front of the server enabling time-saving and efficient physical serviceability that reduces TCO.
Capacity	An M.2 SSD uses a 22 mm wide enclosure that constrains flash memory placements and limits high drive capacities.	WHAT'S NEEDED	A form factor that provides supported SSDs with wider enclosure widths resulting in more space on the device for additional flash memory chips that yield high capacity SSDs that provide more capacity per allowable space.
Security	An M.2 SSD does not support a 'presence detect' capability so if the SSD is powered off, the system does not know if the device is in the system or not.	WHAT'S NEEDED	A form factor that has a 'presence detect' capability so if the device is powered off, it can be detected if it is in the system or removed from the system.

Taking into account 'what's needed' for a new form factor architecture given 2.5-inch and M.2 drive limitations, the resulting architecture needs to be a balance of these requirements to achieve an optimal design:

Requirements to Address	Description
Signal Integrity Issues	*May be apparent in next-generation high-frequency interfaces such as the upcoming PCIe 6.0 interface specification.*
Link Widths	*Multiple host connection link widths should support device types with link widths for PCIe x4 and x8 connections.*
Form Factor Sizes	*Different sizes should be supported for 1U and 2U platforms, and be large enough to accommodate multiple device types and high performance flash memory controllers.*
Power Envelopes	*Options should be available that scale power envelopes to higher power devices. For PCIe 4.0 NVMe SSDs, 25W is required to saturate a PCIe 4.0 (16 GT/s x4) link.*
Thermal Capabilities	*Capabilities that enable continued operations in extreme server temperature environments. Improved airflow can be very beneficial to a system and is based on Cubic Feet of air moved per Minute (CFM) versus system temperature. A system that operates with lower CFM helps to reduce power and cooling when larger heatsinks are used.*
Physical Serviceability	*Hot-swapping drives without having to power down an entire server.*
Capacity	*Enclosure widths that provide optimal space for flash memory chips, which in turn can enable higher capacity SSDs and more capacity per allowable space.*
Security	*'Presence detect' capability so if the device is powered off, it can be detected if it is in the system or removed from the system.*
PCIe Generations	*The connector system at a minimum should support PCIe 5.0 and PCIe 6.0 interfaces, and ideally, PCIe interfaces beyond PCIe 6.0.*

The 2.5-inch format was an outstanding form factor that served the storage industry for almost 30 years. The M.2 format is a more recent design. Ideally, the next generation form factor should meet the above industry requirements for at least the next 10 to 15 years. Hence, the advent of E1.S form factors.

E1.S Form Factors

E1.S is a flexible, power-efficient building block for hyperscale and enterprise compute nodes and storage. The basic M.2 form factor was popular in hyperscale data centers due to its low-cost structure, flexibility and scalability of multiple drives per server – but has challenges as noted. E1.S form factors solve these problems while also maintaining a small package design. At 33.75 mm wide and 118.75 mm long, an E1.S SSD enclosure is wider and longer than a basic M.2 SSD (22 mm wide and 110 mm long), to accommodate more flash memory chips that fit vertically in a 1U chassis.

E1.S variants offer flexibility for power, performance, scalability and thermal efficiency. The 9.5 mm symmetrical enclosure allows scalability up to 20W and a PCIe x8 connection if required. The 15 mm and 25 mm asymmetrical enclosures offer a tradeoff of fewer drives per rack unit but improved power and performance per drive, as well as improved cooling and thermal performance, with the ability to decrease the required airflow.

E1.S Variants	Device Enclosure	Sustained* Power	Enclosure Width	Enclosure Length	Enclosure Thickness
E1.S 9.5 mm	*Symmetric*	*20W*	*33.75 mm*	*118.75 mm*	*9.5 mm*
E1.S 15 mm	*Asymmetric*	*25W+*	*33.75 mm*	*118.75 mm*	*15 mm*
E1.S 25 mm	*Asymmetric*	*25W+*	*33.75 mm*	*118.75 mm*	*25 mm*

The three E1.S SSD types are represented by KIOXIA XD7P Series PCIe 4.0 Data Center NVMe E1.S SSDs:

9.5 mm[4]

15 mm[4]

25 mm[4]

The E1.S form factor is defined by the following SNIA SFF-TA specifications[5]:

Specification	Description
SNIA-SFF-TA-1006 Rev. 1.5	*Enterprise and Datacenter 1U Short Device Form Factor (E1.S)*
SNIA-SFF-TA-1009 Rev. 3.1	*Enterprise and Datacenter Standard Form Factor Pin and Signal Specification (EDSFF)*

The E1.S form factor supports the following SSD capacities:

Expected E1.S SSD Capacities
960 gigabytes[3] (GB), 1.92 TB, 3.84 TB, 7.68 TB, 15.36 TB

E1.S SSDs are targeted for the following server configurations:

E1.S Variant	Recommended Server Configuration
E1.S 9.5 mm	1U Edge Servers
E1.S 15 mm	1U Rack Servers
E1.S 25 mm	1U, 2U and Larger Rack Servers

Addressing Form Factor Requirements

E1.S form factor address the requirements for a new form factor architecture given the limitations of 2.5-inch and M.2 form factors as follows:

Status and Fault/Locate LEDs

E1.S-compliant devices have two LEDs on its front face which are mandatory for all device variations. One is a green 'Status LED' that indicates overall device status and is controlled by the device firmware. The second is an amber host-driven 'Fault/Locate LED' that indicates when a device is in a fault condition or when the host needs to identify the device in a chassis. The amber LED is controlled by the host via the LED pin on the device connector.

E1.S Use Cases

E1.S form factors are well-suited for these use cases:

E1.S 9.5 mm	Use Cases
Target System	Small footprint systems
Examples	Blade servers; Edge compute and storage systems; Dense, scaled servers
Primary Need(s)	Performance; Excellent thermal capabilities
Use(s)	U.2 and M.2 replacement
Typical Configuration	6 to 12 E1.S SSDs

Source: SNIA[6]

E1.S 15 mm	Use Cases
Target System	1U/2U/4U and special-purpose systems
Examples	Compute blades and systems; Performance and capacity-optimized storage systems; Artificial Intelligence/Machine Learning (AI/ML) and High Performance Computing (HPC) systems; Edge systems
Primary Need(s)	Scalable performance and capacity; Optimized thermal performance
Use(s)	U.2 and M.2 replacement for boot and primary storage
Typical Configuration	2 to 32 E1.S SSDs

Source: SNIA[7]

E1.S 25 mm	Use Cases
Target System	2U and larger systems
Examples	Storage appliances; Input/Output (I/O) and storage-rich servers / databases; Performance-oriented storage systems
Primary Need(s)	Scalability of performance with capacity; Lower cost per performance
Use(s)	Mainstream U.2 replacement
Typical Configuration	36 to 64 E1.S SSDs

Source: SNIA[8]

E1.S Key System Benefits

The E1.S form factor delivers data center-optimized PCIe 4.0 NVMe drives that provide system benefits designed to the OCP Datacenter NVMe SSD Specification:

- Enables a much smaller system with high density storage
- Delivers significantly improved system airflow and thermal capabilities compared to 2.5-inch drives
- Provides very efficient modular scaling of NVMe SSD capacity and performance

One of the main differentials that set E1.S SSDs apart from 2.5-inch drives is a significantly improved airflow system supported by thermal capabilities. The airflow of a drive is based on CFM versus system temperature. Figure 3 (referenced in SNIA webcast[6]) is an example of the airflow within both an E1.S SSD at a width of 33.75 mm versus a standard 2.5-inch (U.2) drive[6]. The results indicate that as temperatures rise within a system, an E1.S SSD can move air must faster than a 2.5-inch drive.

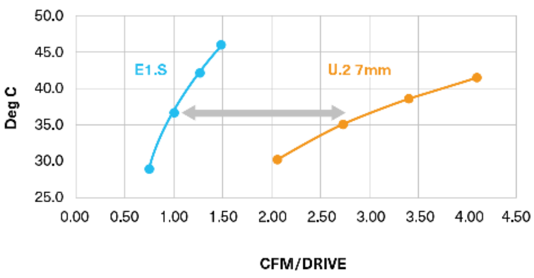

Figure 3: Airflow comparison between an E1.S SSD and a 2.5-inch drive
Source: SNIA CMSI webcast[6]

Open Compute Project E1.S Systems

Here are three OCP E1.S systems including the: (1) Yosemite V3 Flash Platform: SV7100G4; (2) Grand Canyon Storage System; and (3) Olympus Concept Server. A brief description of each now follows:

OCP Yosemite V3 Flash Platform: SV7100G4

The OCP Yosemite V3 flash platform (SV7100G4) is a highly modularized system built with E1.S drives in mind, and includes server boards, expansion boards, baseboard, power modules and cooling fans that fit into a 4OU cubby chassis. The cubby can carry 3 sleds, with up to 4 blades per sled, totaling 12 1OU blades. The system allows up to 48 E1.S SSDs (25 mm), resulting in up to 768 TB capacity per chassis.

Image used with permission from SNIA

FEATURE	4OU CHASSIS	
Number of blades supported	*12 for 1OU blades*	*6 for 2OU blades*
Number of SSDs supported	*48 for 25 mm E1.S SSDs*	*36 for 25 mm E1.S SSDs*
Total capacity supported per chassis	*up to 768 TB*	*up to 576 TB*

The SV7100G4 platform is power-efficient, supporting up to 88W and 26 processing cores. It can host four 1U blades in a sled that fits into the OCP cubby chassis. To fill the Open Rack with 1OU blades, a 4OU system can carry 12 blades enabling the entire rack to accommodate 8 systems. Up to 96 servers in a rack is supported making the system an ideal high-density, 2,496-core computing rack for hyperscale data centers.

OCP Grand Canyon Storage System

The Grand Canyon storage system is Meta's latest HDD storage server chassis that follows the previous generation Bryce Canyon storage platform, retaining its high-level system architecture. The system includes a 4OU tall, single drawer, dense storage cubby chassis that contains 4 EDSFF E1.S SSDs, up to 72 HDDs (3.5-inch), 2 compute modules, 1 network interface controller (NIC) and 2 SSDs per compute module. This design is Open Rack v2 (ORv2) compliant and includes an upgraded 1S server card, user interface component (UIC) and a storage controller card (SCC).

Image used with permission from Meta

Architecturally, Grand Canyon maintains many similarities to Bryce Canyon. It is a 4OU system with most of the components installed and serviced from the top of the chassis. An inner drawer design fully extends the unit out of the rack, even when operational. There are three Grand Canyon configurations: (1) Dual storage server; (2) Single storage server; and (3) Open Compute Project.

Olympus Concept Server

(not publicly introduced as of this publication)

The Olympus Concept Server adds E1.S form factors as a common building block to Project Olympus. The addition of E1.S in the 15 mm form factor delivers a common module suitable for deployment across workloads and platform form factors, and enables the reuse of common components such as backplane interconnects and mechanical drive cages. A single common device also allows for rapid development and deployment of density/capacity-optimized server and storage platforms to meet growing demands and ever-expanding hyperscale workload needs. The reduced airflow requirements of ES.1 form factors result in a platform that is optimized for cooling performance in the modern hyperscale data center while next-generation PCIe 5.0 interconnects provide scalable performance.

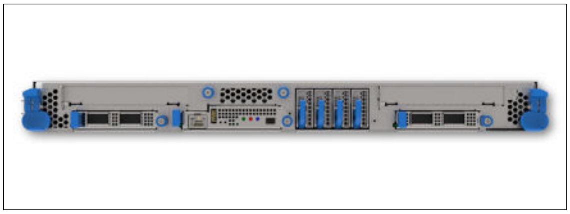

Image used with permission from Microsoft

Microsoft can deploy platforms that consume E1.S 15 mm form factors in support of multiple workloads ranging from compute, high performance storage, density-optimized storage, and artificial intelligence/machine learning in multiple platform form factors including 1U, 2U and 4U.

Who's On Board with E1.S?

Leading authors of the OCP Datacenter NVMe SSD specification, Meta and Microsoft decided on E1.S designs for new and upcoming platforms. Other data center companies took note, and today OCP has over 300 members, including the largest hyperscalers. Additionally, the University of New Hampshire InterOperability Lab (UNH-IOL) has conducted testing for OCP NVMe SSD compliant drives. Also, publications such as Storage Review, Serve the Home, AnandTech, and others, have conducted reviews on E1.S systems and SSDs.

Availability of E1.S Solutions

Many SSD suppliers have shipped E1.S SSDs in all form factor variants (9.5 mm, 15 mm, and 25 mm) in 2022. Available E1.S servers include the Supermicro® SSG-1029P-NES32R E1.S server, the Supermicro SSG-121E-NES24R E1.S server and the most recent SSG-221E-NES24R E1.S server.

Supermicro SSG-1029P-NES32R SuperStorage

Supermicro SSG-121E-NES24R SuperServer

Supermicro SSG-221E-NES24R X13 Petascale Storage System

Summary

E1.S form factors provide many benefits and capabilities beyond 2.5-inch and M.2 form factors, delivering strong data throughput, IOPS and latency performance when driven by the PCIe interface and NVMe protocol. It supports up to 25W power envelopes enabling PCIe 4.0 NVMe SSDs to saturate a PCIe Gen4 link, improved heat dissipation and thermal capabilities featuring hot-swapping and thermal throttling, and large capacity options. OCP evaluated and standardized on E1.S form factors developed by the SNIA SFF-TA working group and can provide benefits and economies of scale to any company with a traditional data center that plans to move to a hyperscale model. There are three OCP E1.S systems including the Yosemite V3 flash platform, the Grand Canyon storage system and the Olympus Concept Server. Leading authors of the OCP Datacenter NVMe SSD specification, Meta and Microsoft decided on E1.S designs for new and upcoming platforms, and many SSD suppliers are shipping E1.S SSDs (in all form factor variants) today.

About the Authors:

Ross Stenfort is a Hardware System Engineer at Meta delivering scalable storage solutions. He has been involved in the development of storage systems, SSDs, RAID-on-Chips (RoCs), Host Bus Adapters (HBAs) and hard disk drives (HDDs) with many successful products launched, with over 40 patents. He is an OCP Storage co-lead and a long-time NVM Express board member. Ross' storage experience includes work at CNEX, Seagate, LSI, Sandforce, SiliconStor and Adaptec. He has a B.S. degree in Electronic Engineering from Cal Poly, San Luis Obispo.

Jason Adrian is the Sr. Director of Azure Platform Architecture at Microsoft, defining system architectures that power one of the largest server fleets in the world. Jason has been heavily involved with OCP for over 9 years, including 3 years as co-chair of the OCP Storage workgroup. He led the Bryce Canyon system design and has a passion for designing innovative solutions to difficult problems. He has over 46 granted patents spanning electrical, mechanical, thermal, and data center design.

Maulik Sompura is the Sr. Director of Product Management and currently leading KIOXIA's data center, enterprise and client SSD product strategy and product management teams. He has 15+ years of NAND flash memory and storage experience. With a strong technical background and key product insights, he has led his teams to achieve tremendous growth in data center and client segments. He has also been a close collaborator with hyperscalers spawning the OCP specification helping to pioneer new technologies with industry partners.

Footnotes

[1] 2.5-inch indicates the form factor of the SSD and not its physical size.

[2] Power Loss Protection (PLP) support helps to record data that resides in buffer memory to NAND flash memory by utilizing the backup power of the solid capacitor in case of a sudden power supply shut down.

[3] Definition of capacity - KIOXIA Corporation defines a megabyte (MB) as 1,000,000 bytes, a gigabyte (GB) as 1,000,000,000 bytes and a terabyte (TB) as 1,000,000,000,000 bytes. A computer operating system, however, reports storage capacity using powers of 2 for the definition of 1Gbit = 2^{30} bits = 1,073,741,824 bits, 1GB = 2^{30} bytes = 1,073,741,824 bytes and 1TB = 2^{40} bytes = 1,099,511,627,776 bytes and therefore shows less storage capacity. Available storage capacity (including examples of various media files) will vary based on file size, formatting, settings, software and operating system, and/or pre-installed software applications, or media content. Actual formatted capacity may vary.

[4] Each XD7P Series E1.S SSD product image may represent a design model and the millimeter '(mm)' size indicates the form factor of the SSD and not its physical size.

[5] SNIA SFF specifications are available to the public at: https://www.snia.org/technology-communities/sff/specifications.

[6] Source: SNIA, Webcast Panel Discussion, 'Enterprise and Data Center SSD Form Factors – The end of the 2.5in disk era,' Slide 32, August 4, 2020, 11:00am PT, https://www.snia.org/sites/default/files/SSSI/EDSFF%20Webcast%208-4-2020%20fnl.pdf.

[7] Source: SNIA, Webcast Panel Discussion, 'Enterprise and Data Center SSD Form Factors – The end of the 2.5in disk era,' Slide 17, August 4, 2020, 11:00am PT, https://www.snia.org/sites/default/files/SSSI/EDSFF%20Webcast%208-4-2020%20fnl.pdf.

[8] Source: SNIA, Webcast Panel Discussion, 'Enterprise and Data Center SSD Form Factors – The end of the 2.5in disk era,' Slide 24, August 4, 2020, 11:00am PT, https://www.snia.org/sites/default/files/SSSI/EDSFF%20Webcast%208-4-2020%20fnl.pdf.