WHITE PAPER

April 2023

Introducing the EDSFF E3 Family of Form Factors

Co-Authors

Bill Lynn, Dell®

Paul Kaler, HPE®

John Geldman, KIOXIA

Industry Landscape

In today's world, server designers are faced with a myriad of new requirements poised by an ever-changing set of use cases. Driven by new technologies, classical servers are finding their way into a broad spectrum of environments. These servers can be located anywhere from high density data centers to smaller distributed edge deployments. Advancements in CPU, memory, accelerator and networking technologies are pushing the limits of the server storage architecture.

Several years ago, the industry realized that the 2.5-inch[1] disk drive form factor was nearing its limits. The Enterprise and Datacenter Standard Form Factor (EDSFF) working group was formed to explore the needs of the industry and architect a new set of form factors that address future enterprise architectural requirements. From this effort, the E3 Family of form factors was created.

What's Wrong with Today's Form Factors?

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 IoT and IIoT technologies (e.g., edge computing, machine to machine communications, continuous real-time video/audio capture), coupled with faster communication protocols and more robust interfaces (5G, WiFi and GPON), data storage media based on 2.5-inch drive formats has become limited. Storage media based on these legacy hard drive formats is challenged to keep technological pace with new server demands, as well as future servers based on PCIe® 5.0 and 6.0 technologies.

The 2.5-inch form factor originated with hard disk drives and is not optimal for flash memory packaging or optimized for flash memory channels. As performance scales to exercise all of the flash memory and activate all of the dies, the power of both the flash memory and interface increase. A form factor is needed that can scale power, PCIe speed increases, and wider PCIe link widths to enable full-throttle input/output operations per second (IOPS) performance per terabyte[2] (TB) capacity.

The scaling of SSD performance and capacity also has a domino effect on SSD power. The existing 2.5-inch SSD format typically caps out at 25W based on a practical limit as to what can be economically cooled. Efficient SSD airflow is critical for ensuring that the overall airflow within the data center is sufficient as it is very hard to provide efficient heat dissipation beyond ~20W.

The connectors used on the 2.5-inch form factor were not designed for extended signal integrity challenges, such as those found in the upcoming PCIe 5.0 and PCIe 6.0 specifications. Signal integrity challenges can manifest resulting in reduced drive count, higher solution costs, Quality of Service (QoS) issues or even system failure.

Requirements

When architecting a new form factor, many different aspects must be considered. The resulting architecture is a balance of these needs that often require trade-offs to achieve an optimal system design. Some of these requirements include:

Signal Integrity (SI)	Multiple Device Types	Link Width	Form Factor Size	Power Envelopes	Thermal Environments
Support for next generation high frequency interfaces. At a minimum the connector system must support PCIe Gen5 and PCIe Gen6 interfaces, and ideally, support interfaces beyond PCIe Gen6.	Support for multiple device types and include NVMe™ SSDs, CXL storage class memory devices, computational storage devices, low end accelerators, and front-facing I/O devices such as NICs.	Support for multiple host connection link widths. Different device types will require different link widths including PCIe x2, PCIe x4, PCIe x8 and PCIe x16 connections	Support for different size requirements to work optimally in both 1U and 2U platforms, to be large enough to accommodate multiple device types, and large enough to accommodate high performance NAND controllers.	Support for reasonable power envelopes in the future that scale to higher power devices. For NVMe SSDs, 25W is required to saturate a PCIe Gen4 x4 link. For Gen5, 30W is the expected max power. For low end accelerators, 70W max power is sufficient.	Support that enables operations in reasonable server thermal environments.

The 2.5-inch format was an outstanding form factor that served the industry for almost 30 years. Ideally, the next generation form factor should meet the requirements of the industry for at least the next 10 to 15 years.

The E3 Family

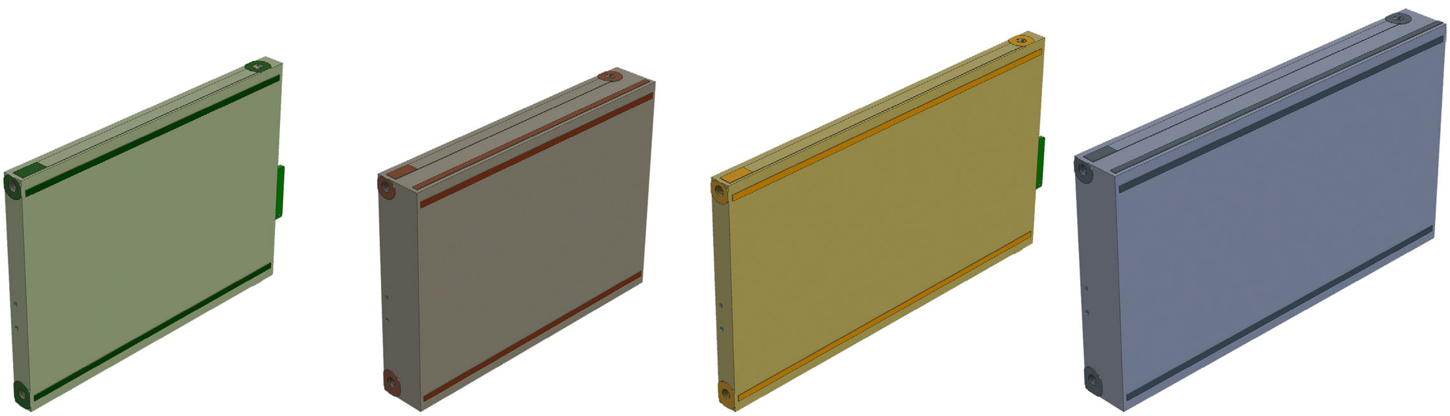
The E3 family of devices currently consists of four different form factors that are defined by a group of SNIA SFF specifications. The SNIA SFF specifications that define the E3 family include:

Specification	Description
SNIA-SFF-TA-1008 Rev. 2.0.2[3]	Enterprise and Datacenter Device Form Factor (E3)
SNIA-SFF-TA-1002 Rev. 1.3[3]	Protocol Agnostic Multi-Lane High Speed Connector
SNIA-SFF-TA-1009 Rev. 3.0a[3]	Enterprise and Datacenter Standard Pin and Signal Specification (EDSFF)
SNIA SFF-TA-1023 Rev. 1.0a[3]	Thermal Characterization Specification for EDSFF Devices

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

The E3 Family of form factors pictured below include the following descriptions and the primary device it is intended for:

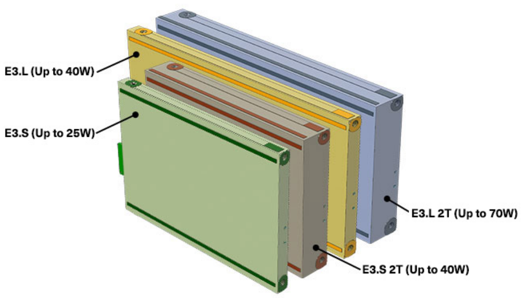
E3 Family Description



E3 Short Thin (E3.S)	E3 Short Thick (E3.S 2T)	E3 Long Thin (E3.L)	E3 Long Thick (E3.L 2T)
Targeted to NVMe SSDs with x4 PCIe link widths though it can mechanically fit an x16 card edge. It supports power profiles up to 25W and positioned to be a primary form factor for mainstream NVMe server storage subsystems as it can be used across a wide variety of platforms including modular and short depth chassis. The thickness designator of '1T' is not required.	Targeted to higher performance NVMe SSDs (PCIe Gen5 x4 saturation), CXL, SCM, computational storage or front I/O implementations. It supports x4, x8 or x16 PCIe link widths and power profiles up to 40W.	Targeted to be a primary form factor for storage subsystems and server platforms requiring maximum capacity for each 'U' configuration that utilize deeper chassis, and for high-capacity NVMe SSDs or SCM devices with support for x4, x8 or x16 PCIe link widths and power profiles up to 40W.	Targeted to FPGAs or accelerators, with support for x4, x8 or x16 PCIe link widths and power profiles up to 70W.

The following table lists key characteristics of the E3 form factors:

Device Variation	Height	Length	Width	Power
E3.S	76mm	112.75mm	7.5mm	Up to 25W
E3.S 2T	76mm	112.75mm	16.8mm	Up to 40W
E3.L	76mm	142.2mm	7.5mm	Up to 40W
E3.L 2T	76mm	142.2mm	16.8mm	Up to 70W



E3 devices also support two LED locations on its front face that have the following functions[4]:

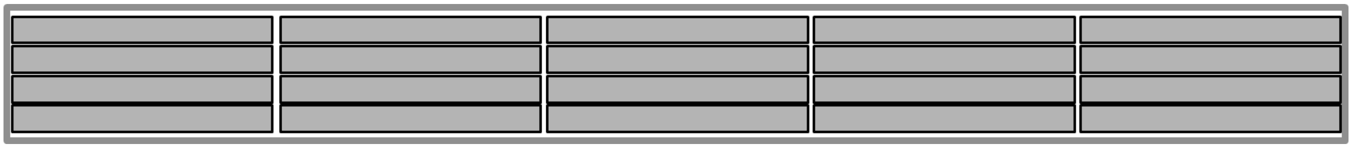
Status LED: The Status LED is green or green/white bi-color and indicates the overall status of the device. The green element is mandatory for all device variations and is controlled by the device firmware. The white element is optional and may be implemented by devices that require indication when it is safe to remove the device from the host. If implemented, the white element is controlled by the device firmware. The white element should not be implemented for NVMe SSD devices.

Fault/Locate LED: The Fault/Locate LED is amber/blue bi-color and indicates when a device is in a fault condition or when the host needs to identify the device in a chassis. The amber and blue elements are mandatory for all device variations and are controlled by the host via the LED pin on the device connector. Specific properties for the various LEDs (such as wavelength and luminosity), as well as the LED control pin definition, are defined in the SFF-TA-1009 specification.

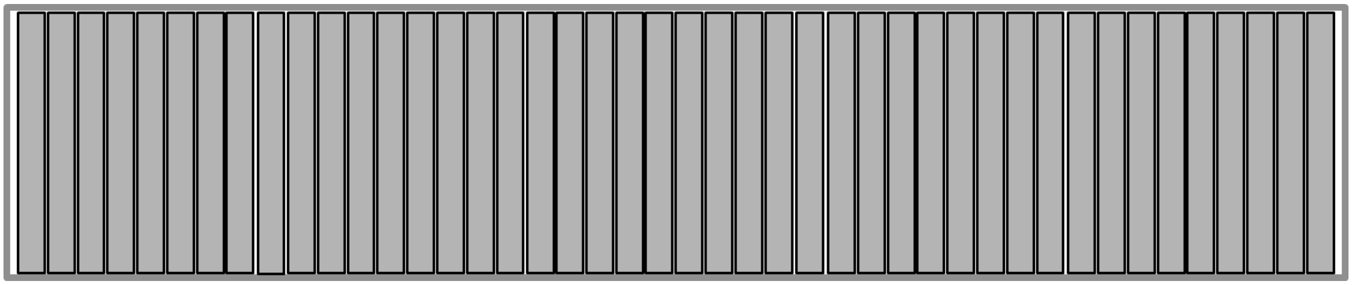
System Design with E3

E3 devices were designed to make efficient use of flash memory chips for SSD storage density, and efficient use of the number of devices in server and storage systems. An E3.S 2T is designed to enable 25% to 50% more flash packages compared to today's U.2 form factor SSD (2.5-inch, 15mm Z-height). E3.L variants provide roughly the same and ~2x the capacity of today's 2.5" SSDs with E3.L (7.5mm Z-height) and E3.L 2T form factors, respectively.

For those systems designed for use with fault-tolerant RAID controllers, drives with a smaller capacity or 'blast radius' results in lower rebuild times when the drive fails. As stated in the Requirements section, all E3 devices are usable in both 1U and 2U platforms. Using a device pitch of 9.3mm, a 1U platform can support up to twenty (20) E3 (7.5mm) or ten (10) E3 2T (16.8mm) devices. This provides similar capacities as 2.5"-based systems using E3.S devices, and doubles the capacity when using E3.L drives.



The figure above shows a 1U system with twenty (20) front loading E3 thin devices. Using a device pitch of 9.3mm, a 2U platform can support a maximum of between forty-four (44) and forty-six (46) front loading E3 devices depending on the 2U chassis mechanical structure. Actual implementations will vary based on thermal and workload targets as described below.

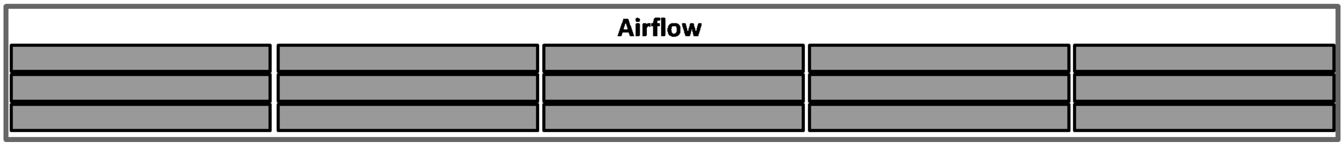


The figure above shows a 2U chassis with forty-four (44) E3 thin devices.

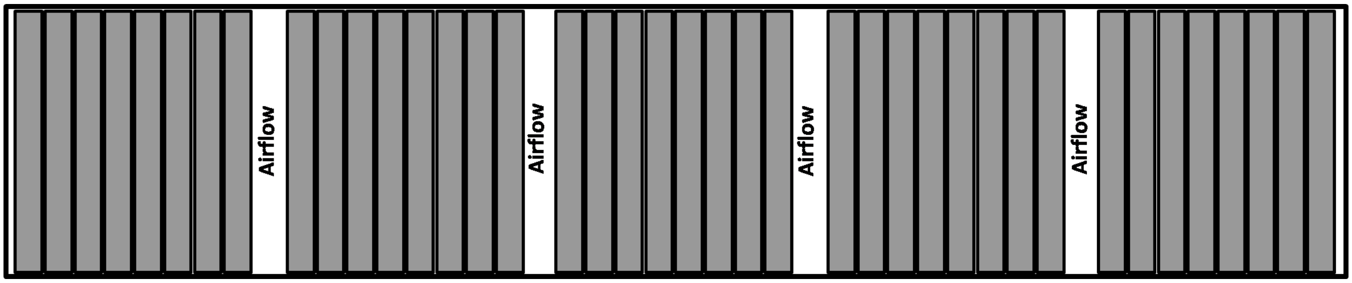
Looking ahead, platform architects will face many new challenges. One of the biggest is that platform power will increase dramatically. Factors that are driving this trend include:

- *CPU power envelopes are currently pushing past 300W*
- *DDR5 doubles DIMM power from 7.5W to 15W per DIMM*
- *GPGPU power is increasing*
- *Storage power is increasing (46 x E3.L @ 40W = 1.84KW)*

To accommodate rising system power, platforms need to increase the amount of airflow through the system. This can be achieved by depopulating some number of devices in the center of the chassis. The figure below shows a 1U chassis with the top row of devices depopulated.



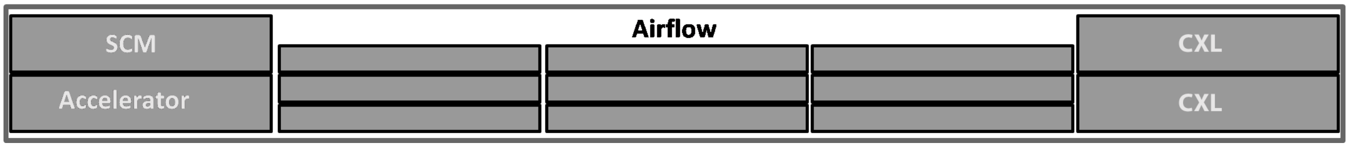
By using the E3 form factor platform, designers can significantly increase overall system airflow while maintaining a relatively large number of devices (in this case 15) to increase storage subsystem performance. The figure below shows a similar approach with a 2U chassis.



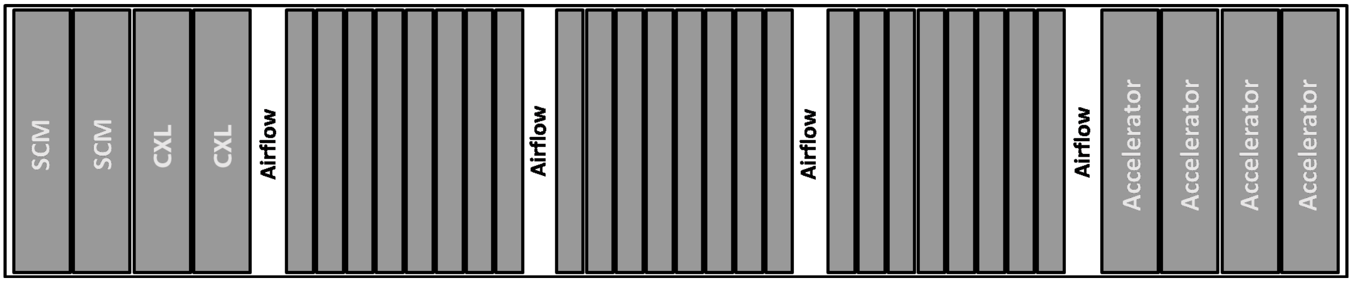
In the example above, the system supports forty (40) E3 devices.

Another challenge facing platform architects is the changing role of the server storage subsystem. Historically, the front of a server has been dedicated to traditional storage devices, but with new E3 designs, this will no longer be true. Future server architectures will share the front bulkhead space of a server with a multitude of device types including more powerful NVMe SSDs, CXL devices, computational storage devices, accelerators and front-facing I/O devices, such as NICs.

The fact that E3 supports multiple mechanical sizes, host link widths and power profiles with a family of interchangeable form factors makes it an ideal choice for system designers. The E3 mechanical architecture supports both carrier based and non-carrier based implementations. Using a wrap-around carrier, system designers can design a single drive bay mechanics set that will allow interchangeability of E3 (thin) and E3 2T (thick) devices. The figure below shows an example of a 1U system supporting four (4) alternate device types while still maintaining nine (9) SSD slots.



The figure below shows an example of a 2U system supporting up to eight (8) alternate device types while still maintaining twenty-four (24) SSD slots.



The flexibility of the E3 form factor gives platform architects a wide range of options when it comes to supporting multiple system use cases. The ability to optimize around either density, host bandwidth, system power or device type makes the E3 form factor an ideal choice for platform architects and system designers.

Key Benefits

The new E3 Family provides a number of key benefits to both system architects/designers and end-users alike, and fall into six categories as follows:

Flash-optimized Form Factors Maximize Density and Performance

Designed specifically to enable flash memory capacity to maximize SSD density

Wide Range of Power Profiles for Higher Capacity/Higher Performance SSDs

Features power profiles up to 70W with the ability to support more NAND chips, increased SSD capacities and higher performance. Higher power also enables additional device type support.

Larger Form Factors can Accommodate Larger SoCs

Supports a range of devices such as SSDs, NICs, GPUs, FPGAs, CXL SCM devices, and low-end accelerators, and a wide range of generalized and specialized applications

Scalable Connector Design, Multiple Link Widths and Future Support for PCIe

Supports device link options (up to x16) within the same chassis, accommodates multiple devices and power profiles from 25W to 70W, and is ready for future generations of PCIe devices.

Interoperable Form Factors Designed for Future Applications

Enables two "thin" slots to be used as one 2T slot and can accommodate two x4 link widths into one x8 link width from the same chassis and drive cage.

Support for a Broad Range of Chassis Configurations

Supports muliple chassis configurations with 25W to 70W power profiles, and x4, x8 and x16 link widths for tiered storage, storage acceleration, GPU applications and more.

Storage media based on these new E3 formats are better suited to keep technological pace with new server demands and a host of data-intensive applications.

Which Companies Will Offer E3-enabled Products?

While E3 form factors began with a focus on traditional enterprise server and storage use cases, hyperscalers are evaluating E3 for use in scale-out environments. Many server, storage and SSD companies with 2.5-inch storage solutions are aligned with the E3 Family. Vendors with production systems based on the SFF-TA-1002 and SFF-TA-1008 specifications include Dell (Figure 1), Supermicro® (Figure 2), HPE (Figure 3), and KIOXIA (Figure 4).

Figure 1: Dell PowerEdge™ C6620 Multi-node Rack Server

Figure 2: Supermicro SSG-121E-NE316R SuperServer

Figure 3: HPE Alletra 4000/4110/4120 Data Storage Servers

Figure 4: KIOXIA CD7 Series E3.S Data Center NVMe SSDs and KIOXIA CM7 Series E3.S Enterprise NVMe SSDs

HPE ProLiant® DL360 and DL380 servers now support E3.S form factors, with many other servers soon to be available.

Summary

The EDSFF E3 Family of form factors delivers on the promise of evolving SSDs to address future enterprise infrastructure requirements while supporting a variety of new devices and applications. It provides improved airflow and thermals, system implementation benefits, and options for larger SSD capacity points and non-SSD device types.

About the Authors:

Bill Lynn is a Server Advanced Engineering Architect at Dell Inc. with over 30 years of experience in the definition, design, marketing, and sales of storage and server system platforms. His areas of expertise include PCIe SSD device architectures, PCIe-based I/O virtualization, host-based storage architectures and RAID design, storage interconnect architectures, external storage platform definition/architecture/design, blade definition/architecture/design, ODM enablement, storage standards development and high speed circuit design.

Paul Kaler is the lead Storage Architect for the Future Server Architecture Team at Hewlett Packard Enterprise (HPE). He is responsible for researching and evaluating future storage and interconnect technologies and defining the server storage strategy for ProLiant servers to ensure HPE servers are ready to handle customers' next-gen workloads. Paul is also actively involved in multiple standards and industry organizations, and has been a key driver of standards including EDSFF E3, and the newly unified Cloud and Enterprise OCP NVMe SSD spec.

John Geldman is the Director of SSD Industry Standards at KIOXIA America, Inc. and leads the storage standards activities. He is currently involved in standards activities involving JEDEC®, NVM Express®, PCI-SIG, SATA, SFF, SNIA, T10, T13 and TCG. He has been contributing to standards activities for over three decades covering NAND flash memory, hard drive storage, Linux®, networking, security, and IC development. John has been on the board, officered, chaired or edited specifications for CompactFlash, the SD Card Association, USB, UFSA, IEEE 1667, JEDEC, T10, and T13, and currently serves as a member of the Board of Directors for NVM Express, Inc.

Footnotes

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

[2] 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.

[3] SNIA SFF-TA specification revision numbers are correct as of this date when the white paper was published.

[4] Excerpt from SNIA SFF-TA-1008 Specification for Enterprise and Datacenter Device Form Factor (E3), Rev. 2, published November 6, 2020, Section 5.5, page 24.