

Enterprise and Data Center SSD Form Factor – *the end of the 2.5in disk era?*

A SNIA Webcast Panel Discussion

August 4, 2020 11:00 am PT/2:00 pm ET

Moderated by Jonmichael Hands, SNIA SSD SIG Co-Chair

Today's Speakers



Jonmichael Hands
Strategic Planner, Intel
Co-Chair, CMSI SSD
Special Interest Group



Paul Kaler
Storage Architect, HPE



Jason Adrian
Storage Hardware Architect
Microsoft Azure



Jonathan Hinkle
Executive Dir and Distinguished
Researcher - System Architecture
Lenovo



Bill Lynn
Server Advanced Engineering
Architect, Dell

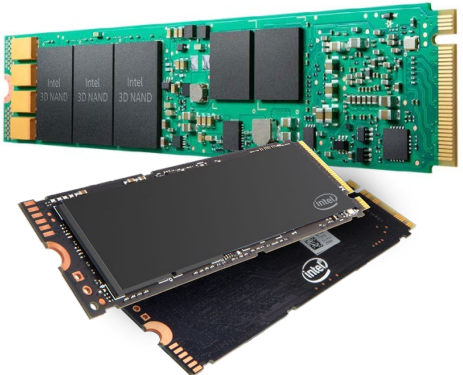


Ross Stenfort, Storage
Hardware Engineer
Facebook



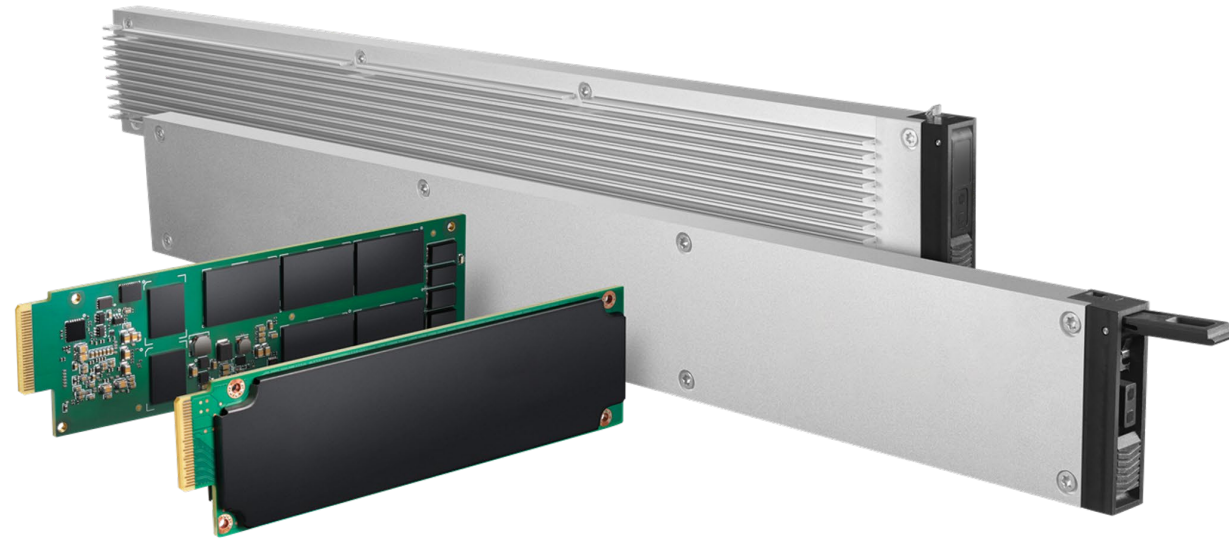
John Geldman
Director, SSD Industry Standards
Kioxia

Data Center SSDs: Previous and Current Options

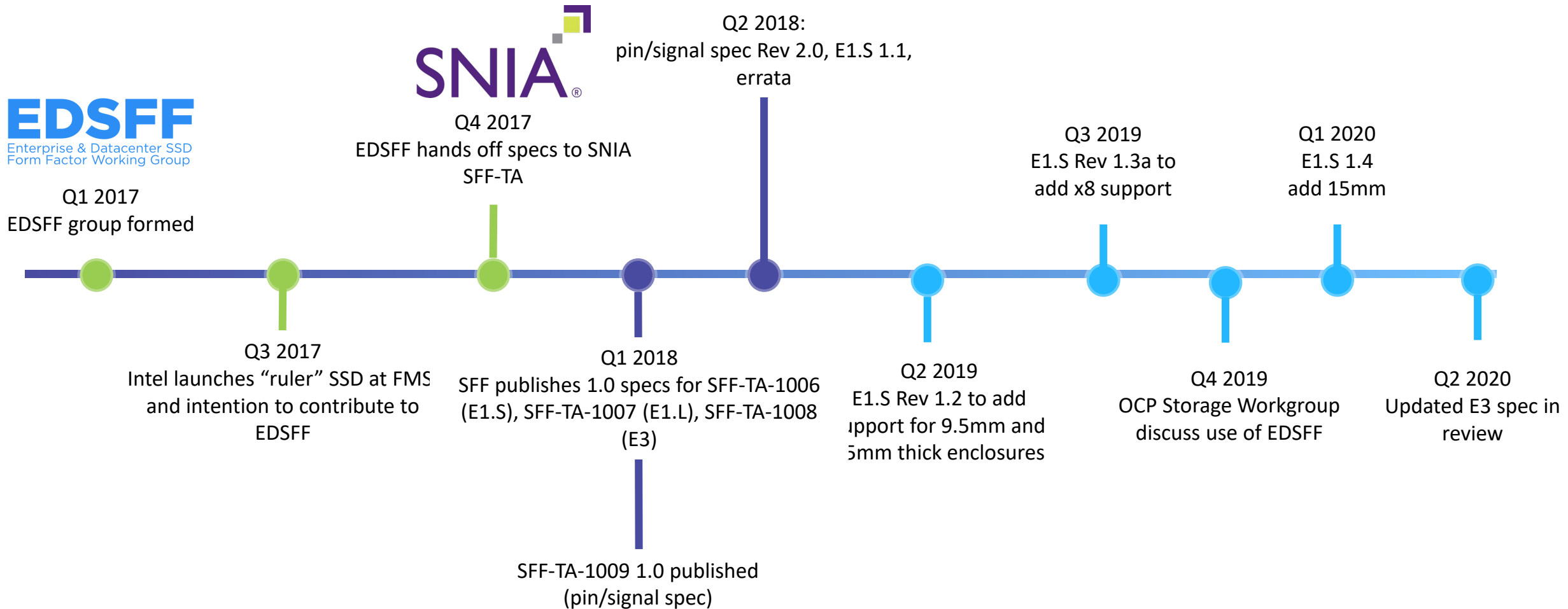
AIC / CEM - Generic	M.2 – Consumer	2.5in Form Factor
		
Good: High-performance, general compatibility	Good: Small and Modular	Good: Hot-plug, Storage features
Bad: need PCIe® AIC slots for other devices, limited hot-plug	Bad: Low capacity, no hot-plug	Bad: Mechanical design descended from HDD
Ugly: consumes lots of space	Ugly: limited power and thermal scaling for data center use	Ugly: Blocks airflow to the hottest components in server

What is EDSFF?

- Enterprise and Data Center SSD Form Factor
- Improved thermals, power, and scalability
- High-speed common connector, pinout – scalable to faster speed PCIe
- Integrated serviceability, hot-plug support
- Built in LEDs, carrier-less design
- Customizable latch for toolless serviceability



EDSFF History



EDSFF Family

- Family of form factors and standards for data center NVMe SSDs
- E1.S for scalable & flexible performance storage
- E1.L for high capacity storage (e.g. QLC)
- E3 high performance SSD for 2U server / storage

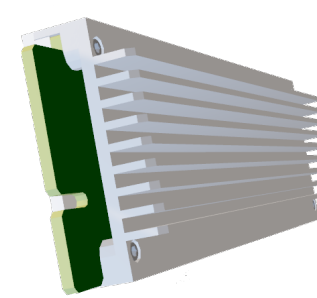
E1.S



5.9mm



9.5mm

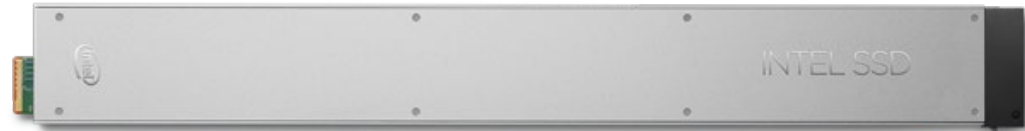


15mm



25mm

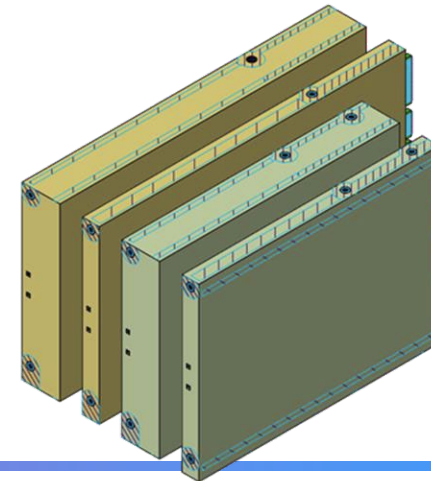
E1.L 9.5mm



E1.L 18mm



E3 v2



Connector Ecosystem

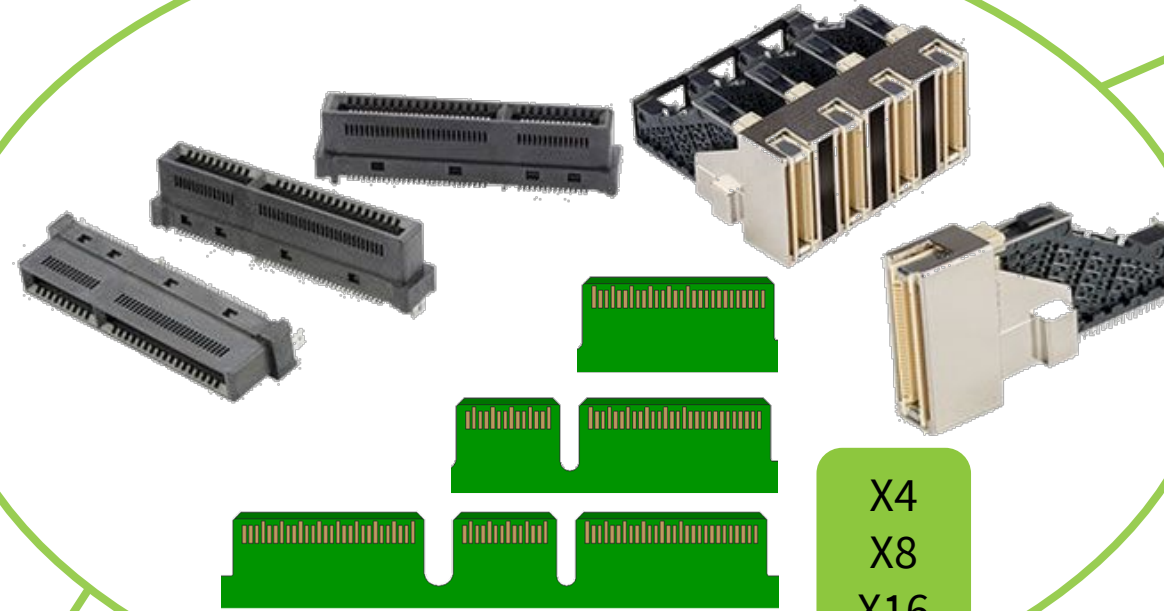


EDSFF

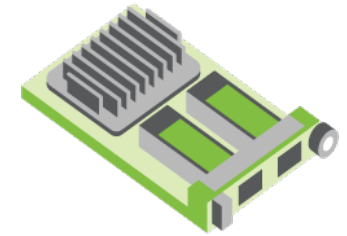
Native NVMe-oF Drive



High Speed Cables

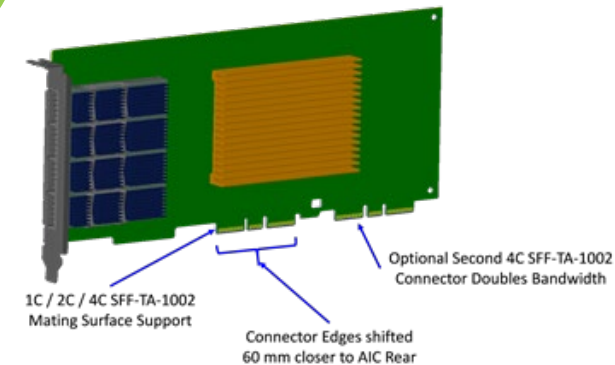


X4
X8
X16



OCP NIC 3.0

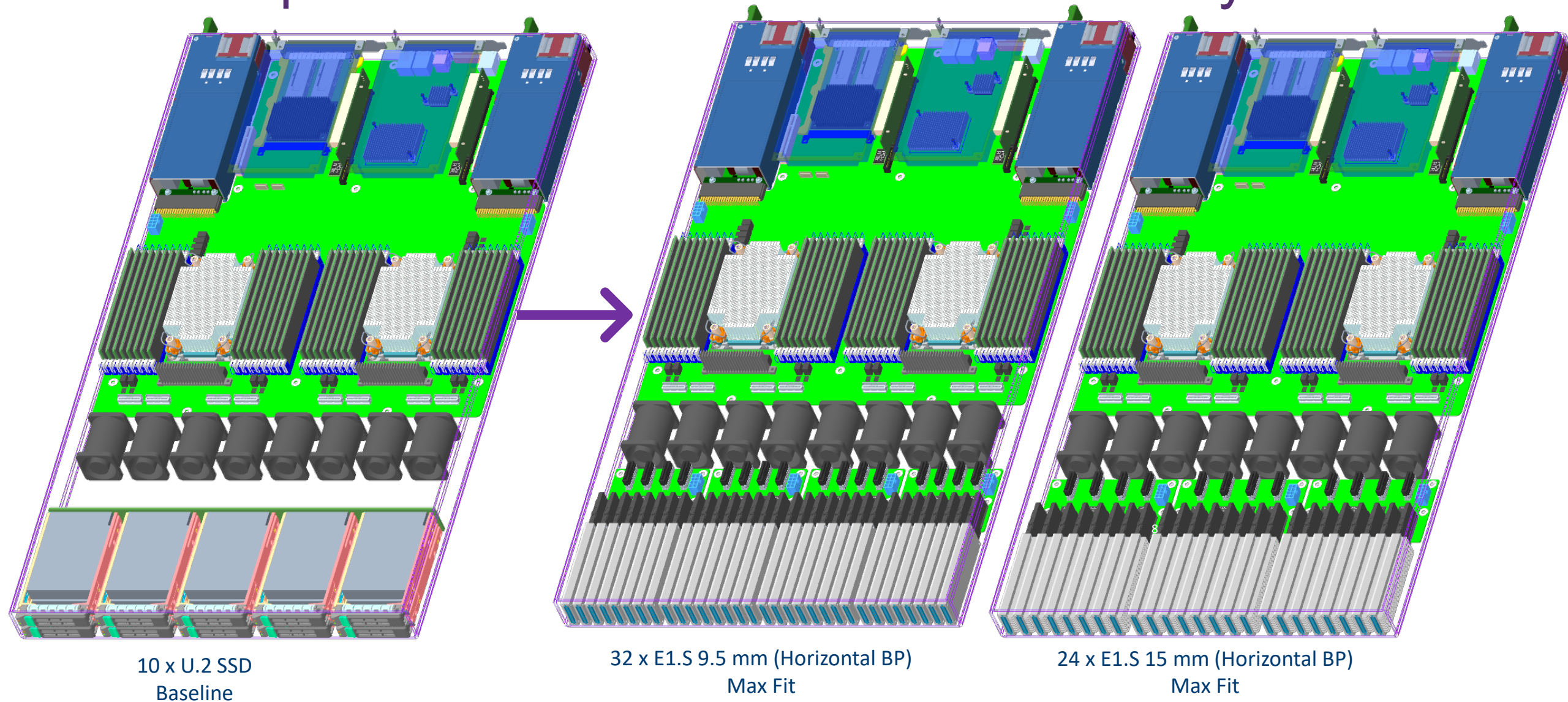
PECFF / SFF-TA-1021



SFF-TA-1002 Connector versions: Vertical (like PCIe CEM), Right-Angle (Like PCIe M.2), Orthogonal, Straddle, Cable. High speed up to 112 GT/s PAM4

SFF-TA-1012 shows pinout differences between EDSFF, OCP, Gen Z, etc.

E1.S Optimal for 1U Performance Scalability



Flash Optimized Form Factors



Jason Adrian
Storage Hardware Architect
Microsoft Azure

Jason is a storage hardware architect at Microsoft Azure, defining the technologies and system designs that power one of the largest storage fleets. His experience ranges from high performance flash systems to archival storage, and everything in between. Jason was previously a hardware systems engineer at Facebook, storage hardware architect at Dell, and a design engineer at EMC. Jason is the co-chair of the Open Compute Storage workgroup, and has 36 granted patents.

How We Got Here

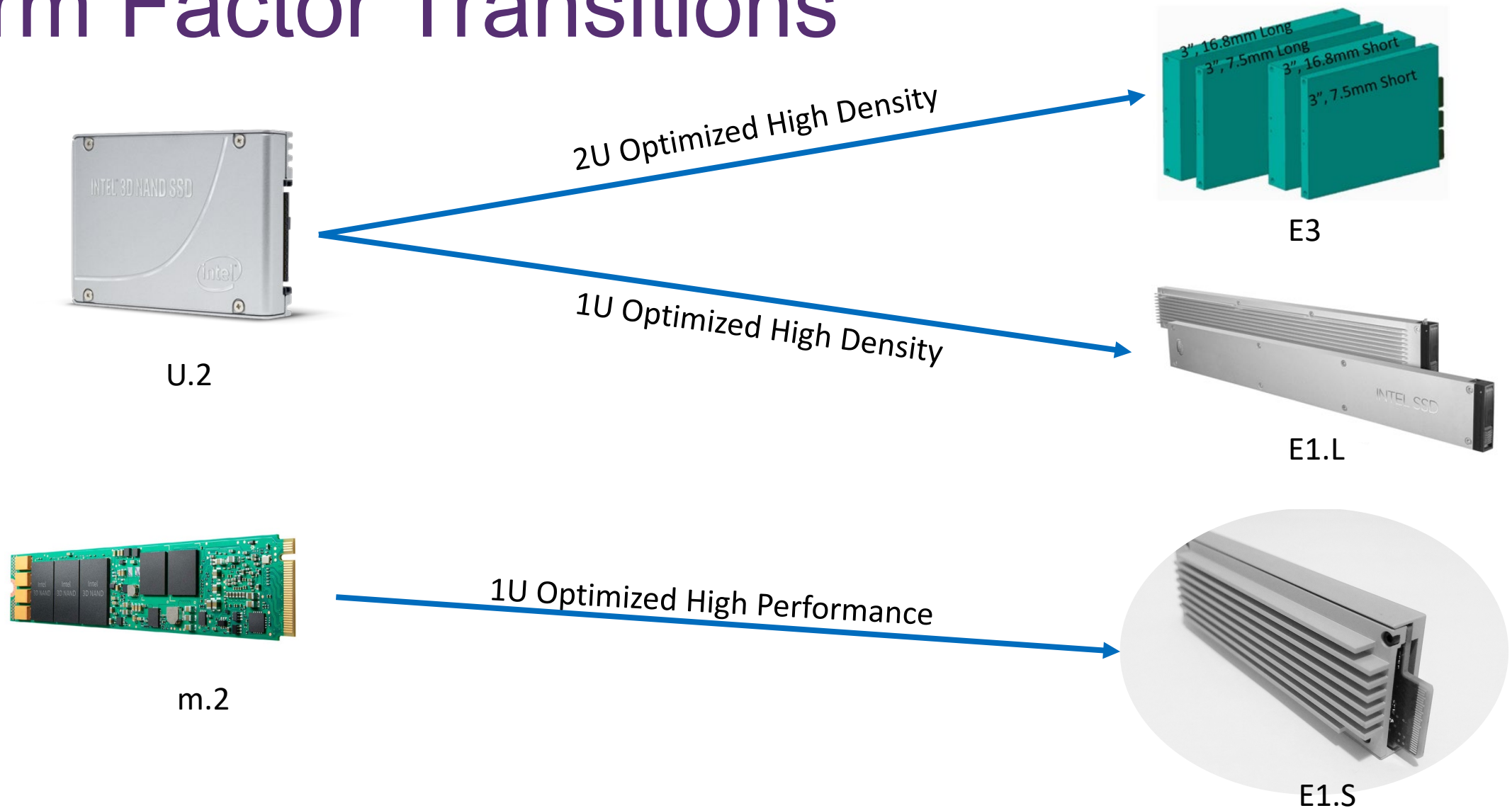
- Easy Adoption
- PCIe backplanes
- SAS/SATA backplanes



Where We Go Next

- Hot pluggable devices
- Thermally optimized
- 1U + 2U
- Performance and Capacity Optimized Options

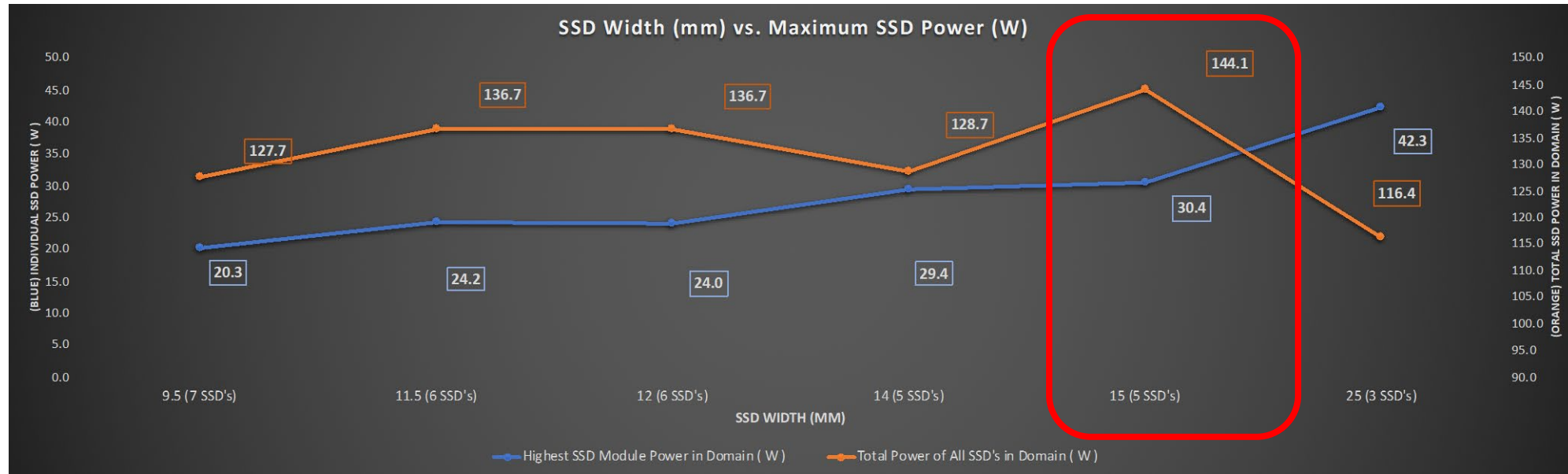
Form Factor Transitions



E1.S Options

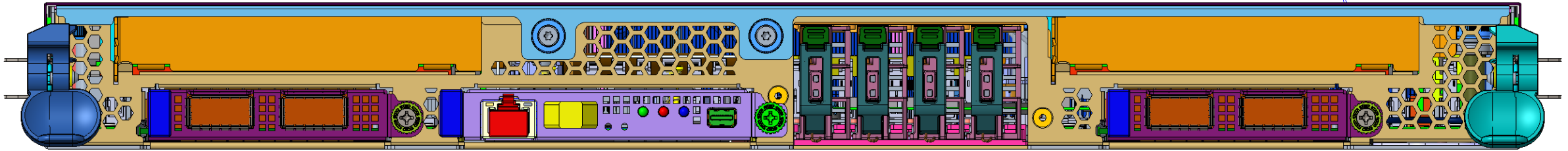
- 5 thicknesses in the TA-1006 specification
- We expect market will drive 2 use cases – boot drives + data drives
- Boot drive – 5.9mm seems like a good fit
- Data drives
 - 9.5mm, 15mm, 25mm all offer different thermal capabilities
 - Microsoft Azure believes 15mm is the sweet spot for our needs
- Regardless of the thickness, all use the same PCB!

E1.S Options



- Evaluated many thicknesses from 9.5 to 25mm
- 15mm offers best performance and density
- Shared findings via OCP Storage Workgroup, and collaboratively drive the 15mm addition to E1.S spec

E1.S 15mm - Higher Density + High IOPS

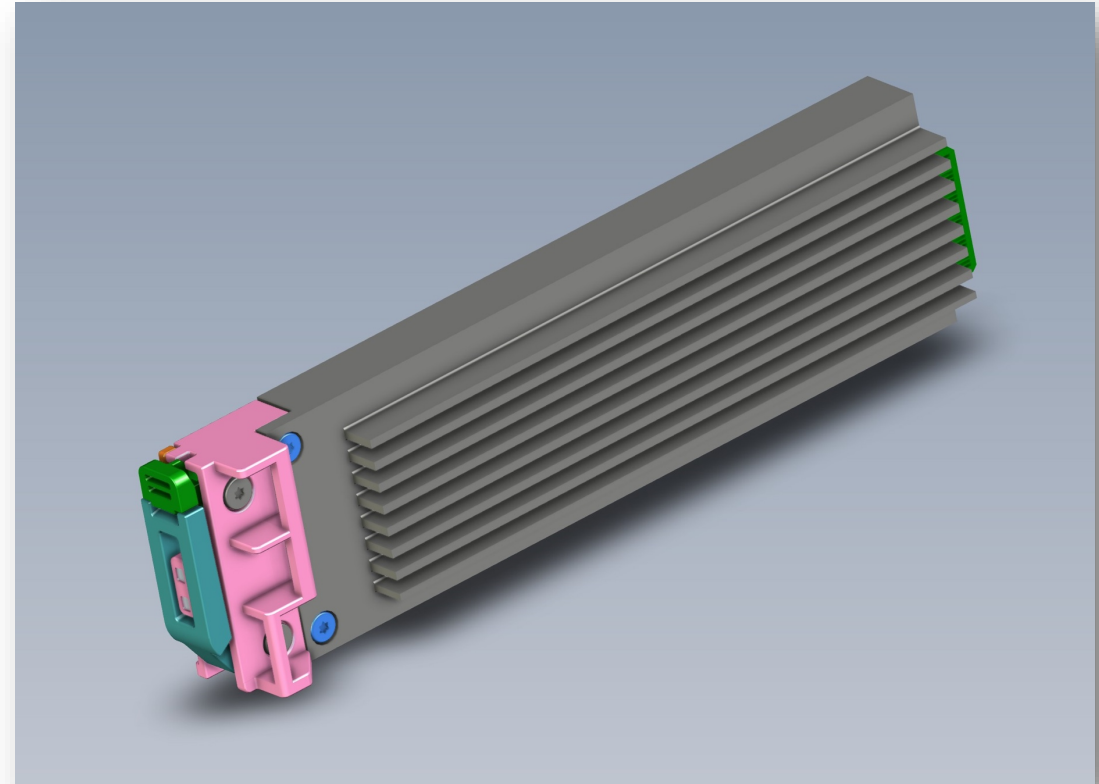


Concept Project Olympus Design

- For 1U servers 4x E1.S 15mm vs 2x U.2
- E1.S 15mm - More IOPS per inch compared to U.2 15mm

15mm E1.S

- Expecting all SSD Supplier's will support 15mm E1.S!
- Multiple options by end of 2020
- Expecting all vendors to ship E1.S 15mm by end of 2021



What about EDSFF for the highest capacity?

- E1.L is the capacity density leader
- 0.5 PB today, 1PB soon in 1U!



E1.S + E1.L

- EDSFF family of SSDs to meet every workload



+



Hyperscale Flash Form Factors

- Ross Stenfort, Storage Hardware Engineer, Facebook



Ross Stenfort is a Hardware System Engineer at Facebook working on storage. He has been involved in development of SSDs, ROCs, HBAs and HDDs. He has over 40 granted patents. He has had extensive storage experience in both large and small companies including CNEX, Seagate, LSI, SandForce, SiliconStor and Adaptec. He has a B.S. in Electronic Engineering from Cal Poly, San Luis Obispo.

What's the problem with today's M.2s?

- Power/ Thermal Does Not Scale
- Unable to activate all dies (TLC, QLC, SCM) thus performance limited
- M.2s do not scale well past 2 TB due to performance/power limitations
- No standard thermal solution
- Poor serviceability
- No hot plug support

What Are The Benefits Of E1.S?

- E1.S with same PCB and firmware supports a diverse family of thermal options for the market:
 - High Density
 - High Performance
- Hot Plug Support
- Works for both storage and compute in 10U
- Support for Gen 4 and Gen 5 PCIe
- Fully standardized in SNIA/SFF:
 - SFF-TA-1006 Revision 1.3a



25mm High
Performance Option

9.5/ 8.01/ 5.9mm
High Density Options

Background: Hyperscale Needs

- *IOPS/ TB Performance Need To Scale Linearly*

Capacity	Performance
1 TB	1 X
2 TB	2 X
4 TB	4 X
8 TB	8 X
16 TB	16 X

- *Low Airflow is critical to ensure data center airflow is sufficient*
 - Facebook Data Center Requirement is 0.145 CFM/W
- Serviceability matters
- Solutions need to scale for the future

Background: Industry Trends

- NAND Dies are getting larger
 - 256 Gb and 512 Gb NAND are going away
 - **Result: 1 and 2 TB Data Center Performance Grade SSDs will go away**
- Dark Flash is becoming a larger issue: Flash can not be utilized due to changing compute to flash ratios
- PCIe Gen 4 and Gen 5 are coming which increase power
- Mainstream Unthrottled Power/Capacity Trends:

SSD Capacity and Type	SSD Device Power (W)	W/TB
1 TB TLC	8.25	8.25
2 TB TLC	12	6
4 TB TLC	15	3.75
8 TB TLC	20	2.5
16 TB TLC	25	1.6
Storage Class Memory	18-25	

COMING SOON

In the future:

- *Data Center SSD's capacity will increase*
- *Device power budgets will increase*
- *The W/TB ratio will decrease*
 - *Improving power efficiency*

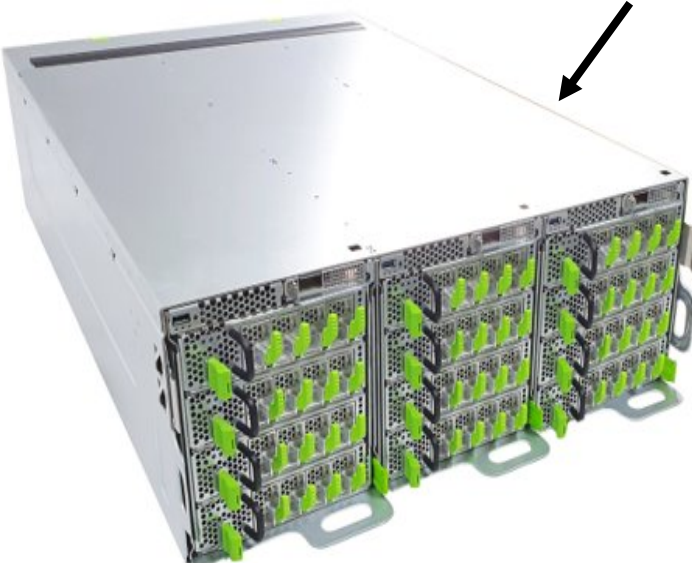
E1.S Platform: Yosemite V3

25mm E1.S Latch



1 OU Blade

Chassis with 1 OU Blades



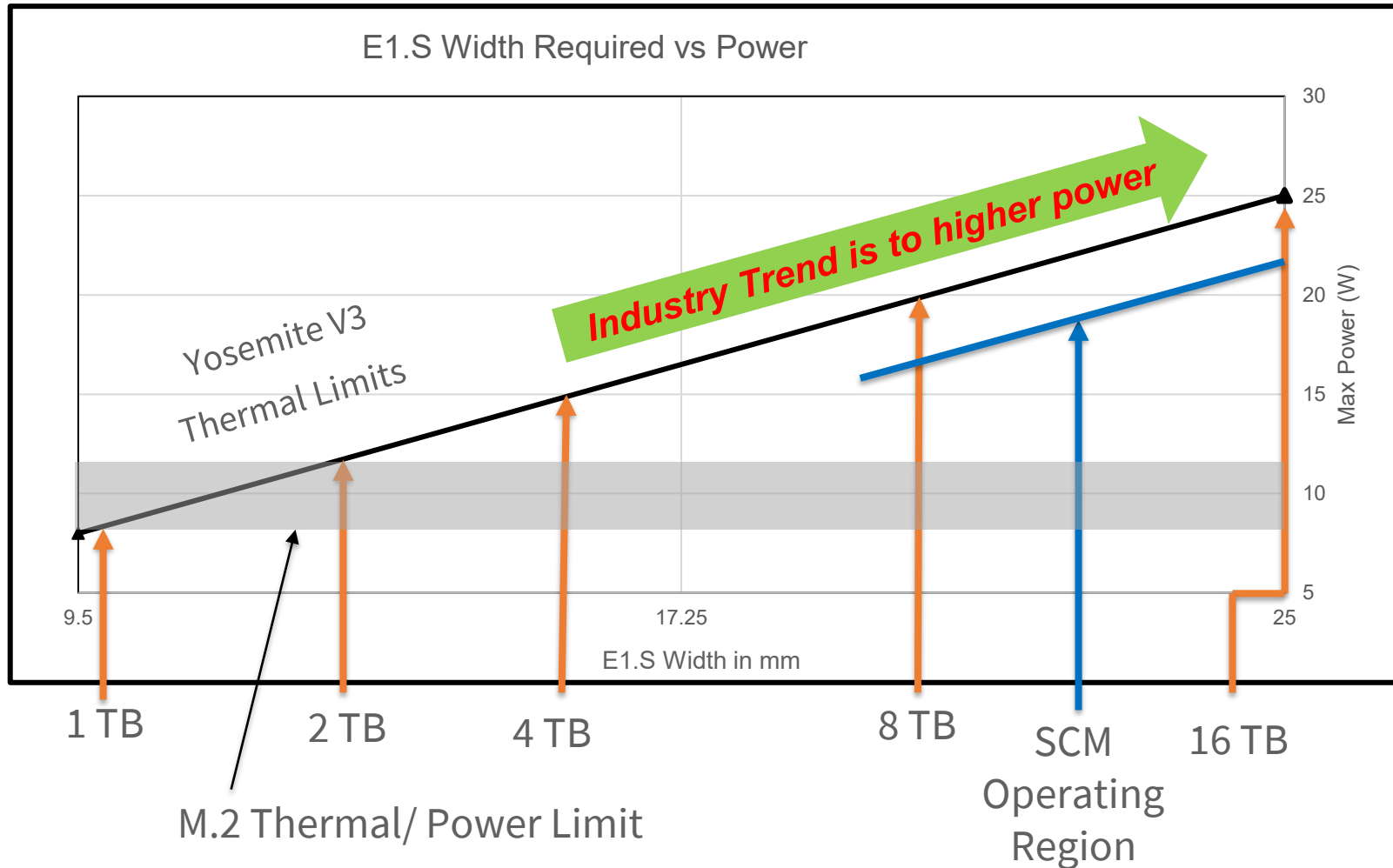
	1 OU Blade Platform	2 OU Blade Platform
Chassis Height	40U	
SSD	25mm E1.S @25W	
Number of SSDs per Blade	4	6
Number of SSDs per Chassis	48	36
Max Blade TLC Capacity	64 TB	96 TB
Max Chassis TLC Capacity	768 TB	576 TB
SCM support	Y	
Use Case	High Performance/ Direct Attached Storage	High Capacity/ Disaggregated Storage

The Future Is Now

Chassis with 2 OU Blades



Yosemite V3 E1.S Thermal Performance



Key

- TLC SSD
- SCM SSD

Conclusion:

- ❖ 25mm E1.S will scale to meet:
 - Power/Performance Requirements
 - Low Airflow Thermal Requirements
- ❖ M.2 will not scale at 4 TB or larger:
 - Performance throttled due to lack of power
 - Can not meet IOPS/ TB performance requirements
- ❖ Less than 25mm E1.S will not scale at low airflow:
 - Can not meet IOP/TB due to power/thermal limitations

Thermal Limit based on Facebook Data Center Requirement of 0.145 CFM/W

E1.S: Form Factor For The Future

Power/ Thermal/ Performance Scales

- Able to activate all dies for fully unthrottled performance

Standard thermal solution

- Scales for both front and back of the box solutions

Serviceability

- Support hot plug
- Case for EMI/ ESD

Resource

- Optimized for mainstream

Dense

- Fits in dense 1U storage and server applications

25mm E1.S:

Scaling For The Next Generation Of Storage



Thank
You



EDSFF E1.S: Cloud, Enterprise, and Edge



Jonathan Hinkle

Executive Director and Distinguished
Researcher - System Architecture
Lenovo

Jonathan Hinkle is the Distinguished Researcher - System Architecture in Lenovo Research, where he leads the investigation of new data center systems architectures and technologies.

Jonathan is an industry leading technical expert in memory, storage devices, and data center systems architecture with over 20 years of experience. In the JEDEC standards organization, Jonathan serves on the Board of Directors, is Vice-Chair of Marketing and Chairs the Hybrid DIMM Task Group standardizing NVDIMMs. He holds BS and MS degrees in Computer Engineering from North Carolina State University.

E1.S: SFF-TA-1006 Industry Standard Form Factor

Vision:

Create a smaller, high density solid state drive standard that is optimized for the data center

E1.S (EDSFF 1U Short):

- Mainstream NVMe™ drive
- Compact, modular form factor
 - Vertical system fit in 1U rack height (44.45mm)
 - Fits in depth of 2.5" drive
- Supports hot-plug and enterprise feature set
- +12V main power for reduced system cost
- LEDs on-drive for lower cost and easier integration
- High Performance and Capacity (PCIe x4 saturation, 12TB+)



Optimized for the NVMe drive design and use across **all data center and edge systems** to scale as mainstream storage.

E1.S Options: Dimensions and Power

Choose your standard enclosure :
(or your own custom mech)



Density in 1U,
edge servers

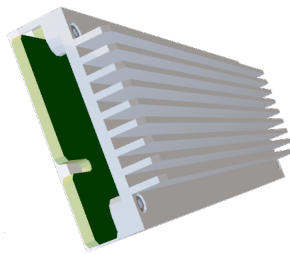
Fits 24 across 1U,
similar to u.2 in 2U

Plenty of thermal
headroom, power.

Standard NVMe
drive card



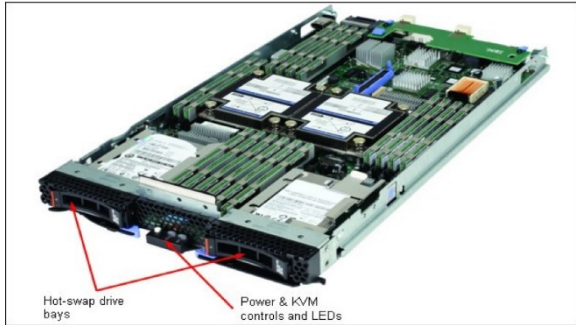
Not widely used



Enclosure Parameter	5.9mm Device	Device with Heat Spreader (8.01mm)	Device with Symmetric Enclosure (9.5mm)	Device with Asymmetric Enclosure (15mm)	Device with Asymmetric Enclosure (25mm)
Recommended sustained power (W)	12	16	20	20+	25+
Enclosure Max Inlet air temperature, 950 m to 3050 m (° C)	35 - (1° C for 175 m of elevation gain)	35 - (1° C for 175 m of elevation gain)	35 - (1° C for 175 m of elevation gain)	35 - (1° C for 175 m of elevation gain)	35 - (1° C for 175 m of elevation gain)
Add in card to add in card pitch (mm)	9	11	13	17	26
Recommended Fan Pressure loss across device (Pascal)	83	52	64	40	21
Airflow, average min per device (CFM). 1 CFM = 1.7 m3/h)	1.41 – (0.01 CFM for every 1° C below 35° C inlet temp)	1.71 – (0.06 CFM for every 1° C below 35° C inlet temp)	2.02 - (0.02 CFM for every 1° C below 35° C inlet temp)	1.5 - (0.02 CFM for every 1° C below 35° C inlet temp)	4.10 - (0.04 CFM for every 1° C below 35° C inlet

E1.S – use cases

Common x4 slots for NVMe drives are ideal:
no stranded lanes
no customer limitations and ultimate scalability
power per device (in front of CPU) at reasonable level



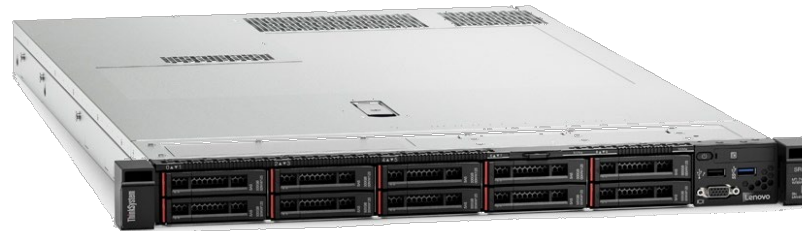
Smaller footprint systems:

Blade servers
Edge compute and storage
Dense, scaled servers

NEED: performance, excellent thermal

USE: **u.2** and **m.2** replacement

Typical: 6-12 x E1.S



1U systems:

General purpose enterprise compute
Rack configurations with 32-40 servers
Hyperscale environments

NEED: scaled performance and capacity, excellent thermal

USE: mainstream **u.2** replacement

Typical: 16-32 x E1.S



2U and larger systems:

Storage appliance
IO and storage-rich server / database
Performance-oriented storage systems

NEED: scalability of performance with capacity, lower cost/performance

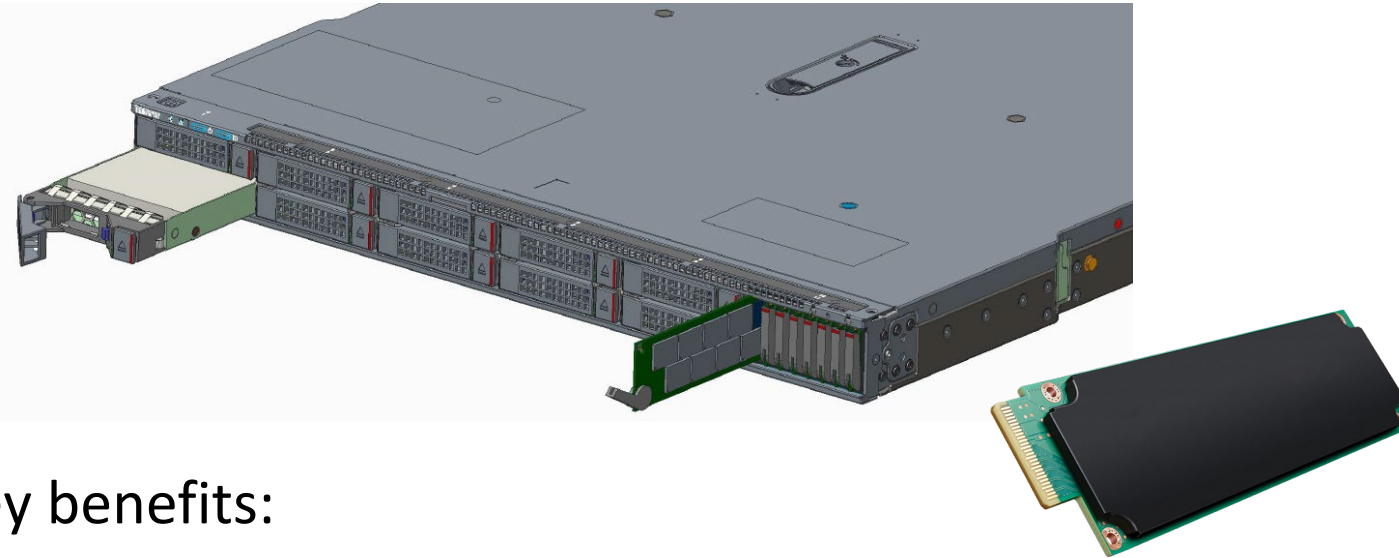
USE: mainstream **u.2** replacement

Typical: 48-64 x E1.S

Alt: Thick E3.S for highest capacity

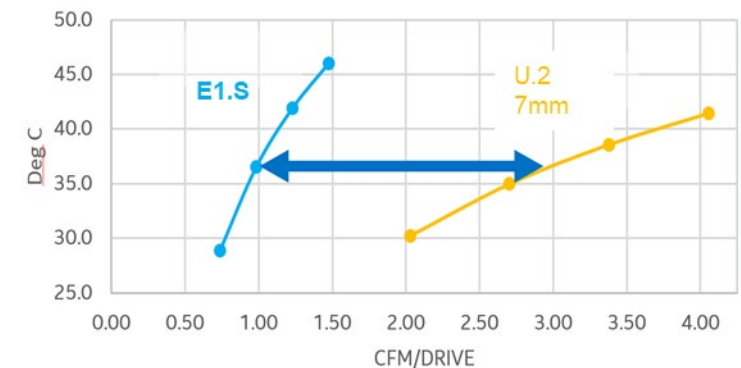
EDSFF 1U Short (E1.S) - Key System Benefits

Industry Standard datacenter-optimized NVMe™ drive that provides significant new system benefits



- Key benefits:

- Much **smaller** enabling high **density** storage
- Significantly improved **system airflow** and thermal solution
- Most **efficient** modular **scaling** of NVMe capacity and performance
- Enhanced feature set in **space-constrained** servers
- **Lower** base system infrastructure and drive **costs** (high volume, common building block)



E1.S Future Use Cases

- E1.S is ideal with its versatility as a common, high volume building block FF across systems and use cases.
- In addition to scaling resources in larger datacenter system use, edge and IoT use cases are also ideal to leverage its small, modular form factor
- Optimal future use focuses on system scaling of devices with low to moderate power (6-25W): providing balanced power and better cooling to enable higher power CPU and GPU in servers
- Upcoming E1.S uses beyond NVMe storage include DRAM memory, Persistent Memory, computational storage, PCIe accelerators and IO fabric connections

Enterprise and Data Center SSD Form Factor – E3

Paul is the Storage Architect for the Future Server Architecture Team at HPE and is responsible for researching future technologies and defining the server-storage strategy for ProLiant servers.



Paul Kaler
Storage Architect, HPE

Bill is a platform Architect with 30 years of experience in the definition, design, marketing, and sales of storage and server system platforms. Responsible for storage subsystem architecture for Dell PowerEdge servers.

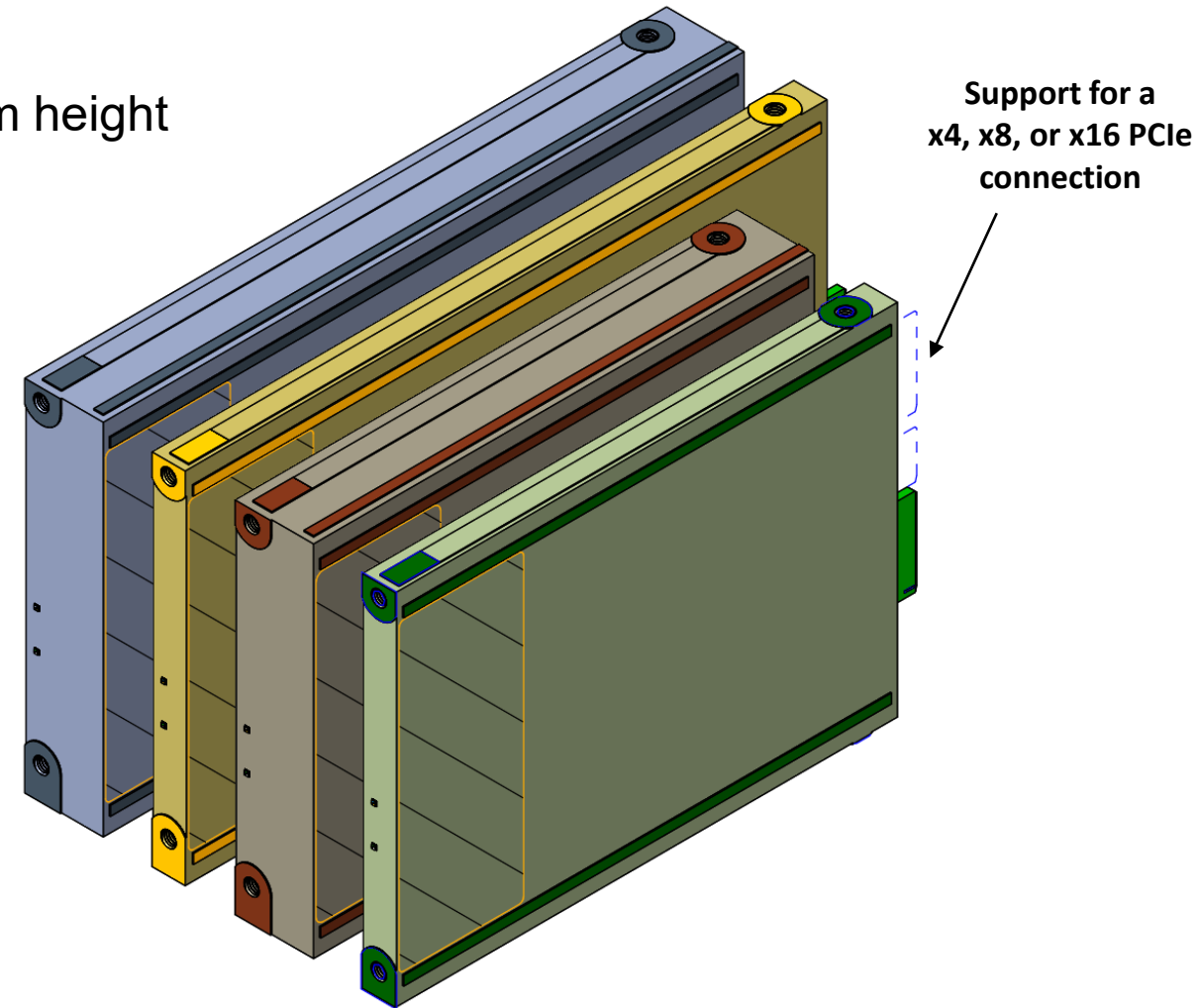


Bill Lynn
Server Advanced Engineering
Architect, Dell

EDSFF E3 for Dummies

- E3 is a family of four form factors with a common 76mm height

- E3.S
 - 76mm x 112.75mm x 7.5mm
 - Target to support from 20W to 25W
 - Optimized for primary NAND storage in Servers
- E3.S, 2x
 - 76mm x 112.75mm x 16.8mm
 - Target to support from 35W to 40W
 - Support for higher power devices like CXL based SCM
- E3.L
 - 76mm x 142.2mm x 7.5mm
 - Target to support up to 40W
 - Support for higher capacity NAND storage
- E3.L, 2x
 - 76mm x 142.2mm x 16.8mm
 - Target to support up to 70W
 - Support for higher power devices like FPGAs and accelerators



Note* - A thick device will fit into two thin slots
- A short device will fit into a long slot

Industry Architectural Requirements

Define a family of devices with a 10+ year lifetime

- Common connector definition (based on TA-SFF-1002)
 - Support for multiple protocols (PCIe, CXL, or GenZ)
 - Support for three link widths (x4, x8, and x16)
 - Support for multiple PCIe generations (at least through Gen 6)
- Size Devices to enable multiple use cases and chassis constraints
 - Optimized for both 1U and 2U chassis heights
 - Device depth(s) optimized to maximize capacity while also meeting chassis depth constraints
 - Device thicknesses that enable maximizing capacity while also enabling high power devices.
- Common connector registration point for all device variations
 - Allow a short device to work in a long slot
 - Allow a thick device to work in two thin slots
 - Allow smaller devices to work in larger cases (Russian Dolls)

Cohesive Family of Devices

High Power Devices

E3.L, 2x (Target 70W)

- Higher end accelerator
- Computational storage

E3.S, 2x (Target 40W)

- High capacity NAND
- Higher end Load/Store SCM/PM (Memory Device)
- Mid-range accelerator
- Computational storage

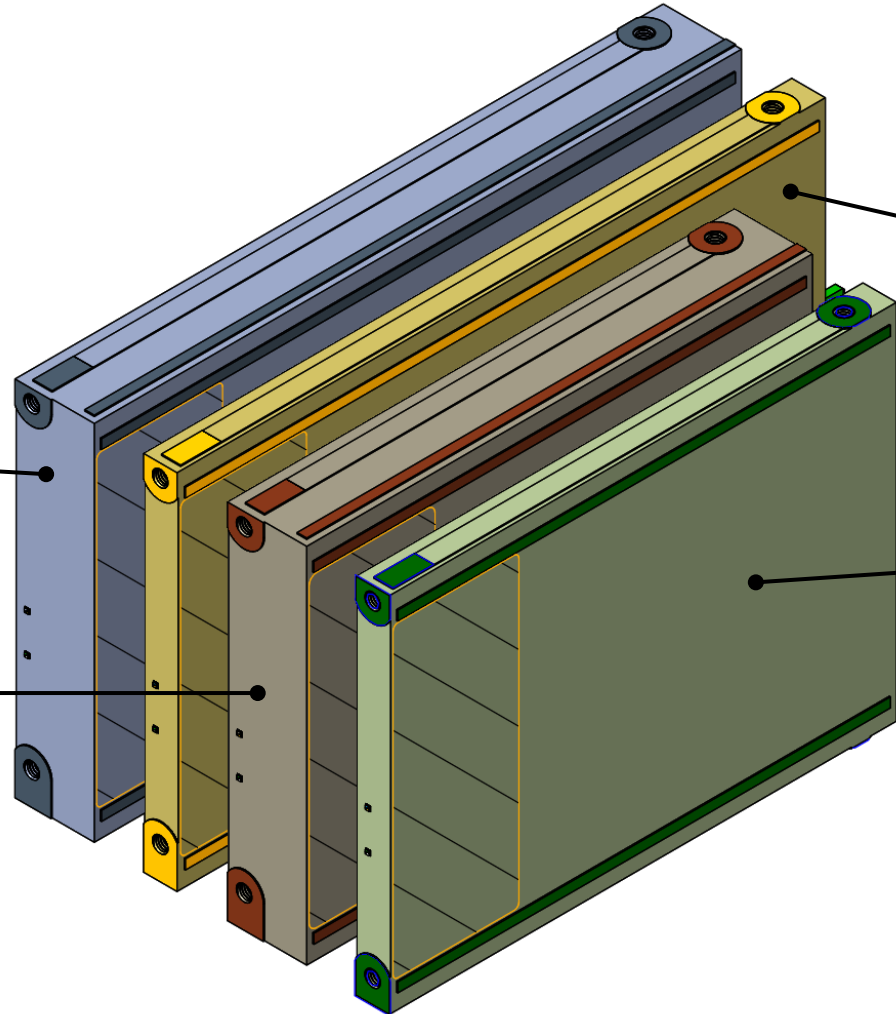
Storage Devices

E3.L (Target 35 to 40W)

- Highest capacity NAND / SCM (Block Storage Device)
- Dense and storage centric chassis

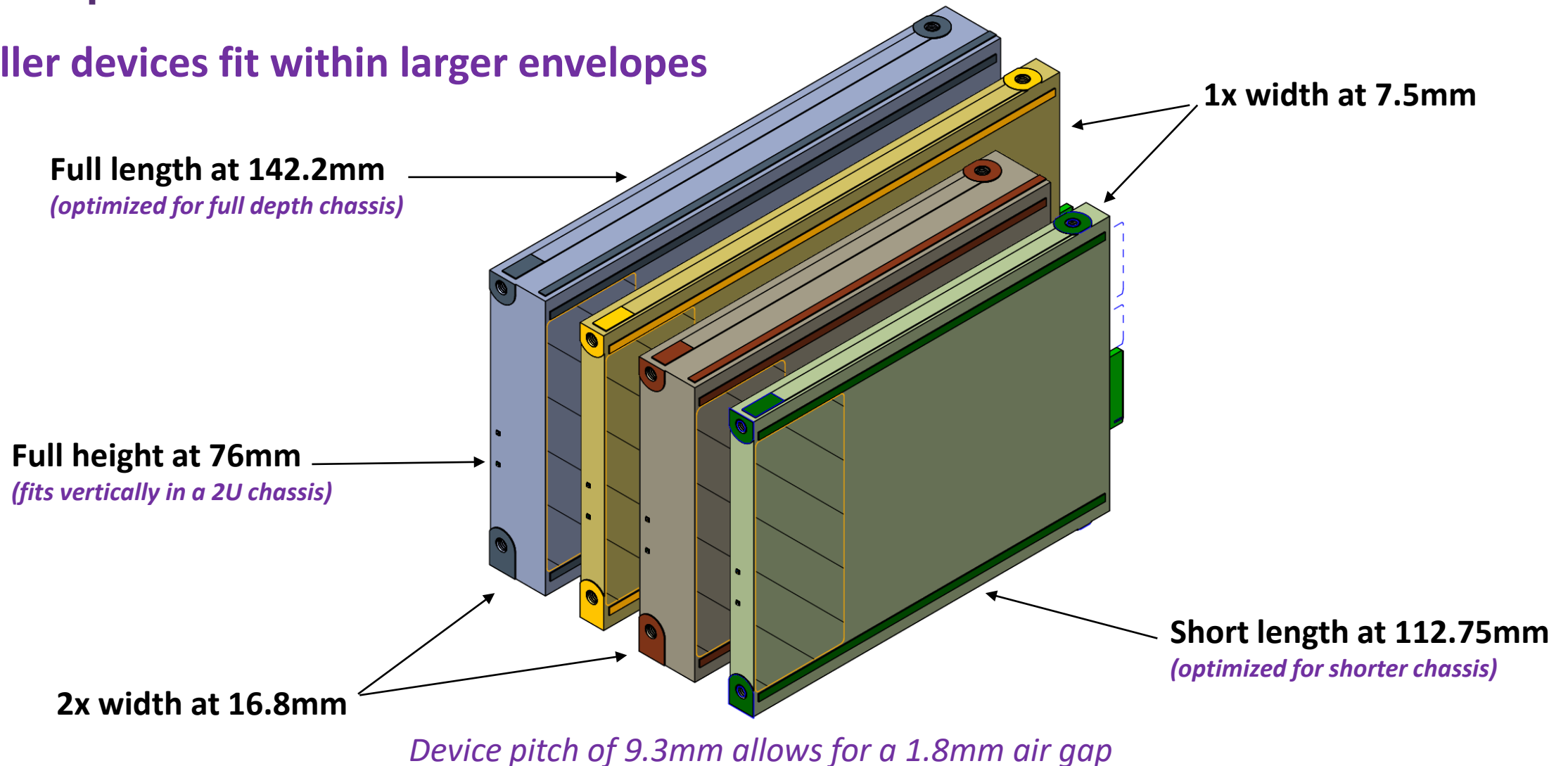
E3.S (Target 20 to 25W)

- 1U & 2U Performant primary NAND storage
- Volume Load/Store SCM/PM (Memory Device)
- Modular and short chassis



Interoperable Device Sizes

Smaller devices fit within larger envelopes



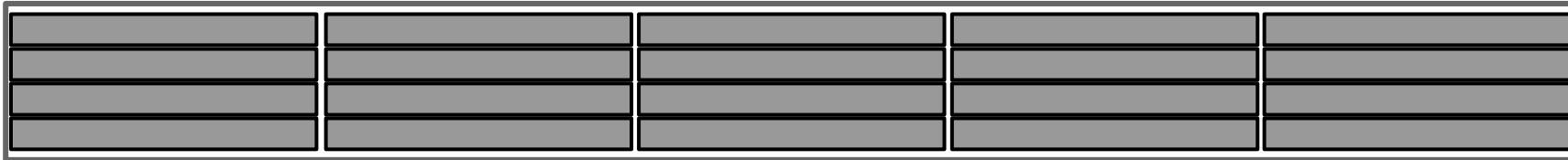
From Factor Design Advantages

E3 family of devices offers many system design advantages

- Support for a wide range of power profiles
 - Current EDSFF pinout for SFF-TA-1002 supports almost 80W maximum power delivery
 - Higher power capabilities allow for more powerful accelerator and other device types
 - E3 will also support higher NVMe power profiles needed for PCIe Gen5 and beyond support
 - Power and thermal limits will be defined by SFF-TA-1023 Thermal Specification for EDSFF Devices
- Support for a wide range of device types
 - Supports host link widths up to x16 PCIe 5 and beyond
 - 2x device thickness allows for the use of standard networking connectors (front facing I/O)
 - Larger PCB surface area allows for larger NAND capacity points and richer device types
- Interoperable device form factors
 - 1x and 2x devices are interchangeable
 - Good balance of device size to infrastructure requirements
 - Right sizing of power profiles to bandwidth capabilities
 - Enable common devices (spares) between 1U and 2U chassis

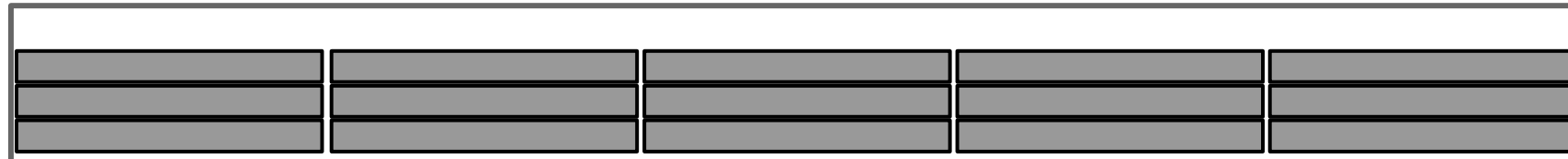
Potential E3 Chassis Configurations (1U)

Storage
Config



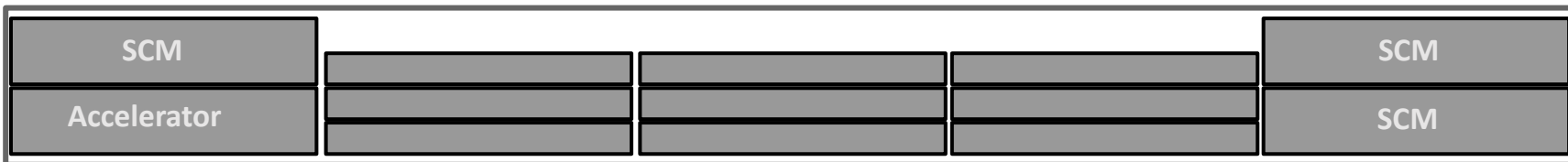
20 E3L 1x Storage Devices

High Power
Server Config



15 E3S 1x Storage Devices with an air channel above

Alternate
Device
Config

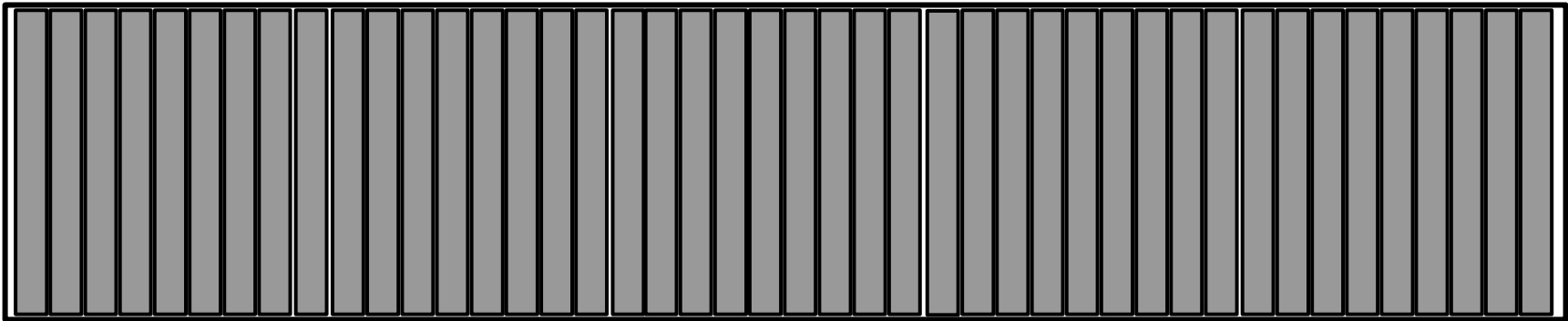


9 E3S 1x Storage Devices and 4 E3S 2x SCMs or accelerators

Potential E3 Chassis Configurations (2U)

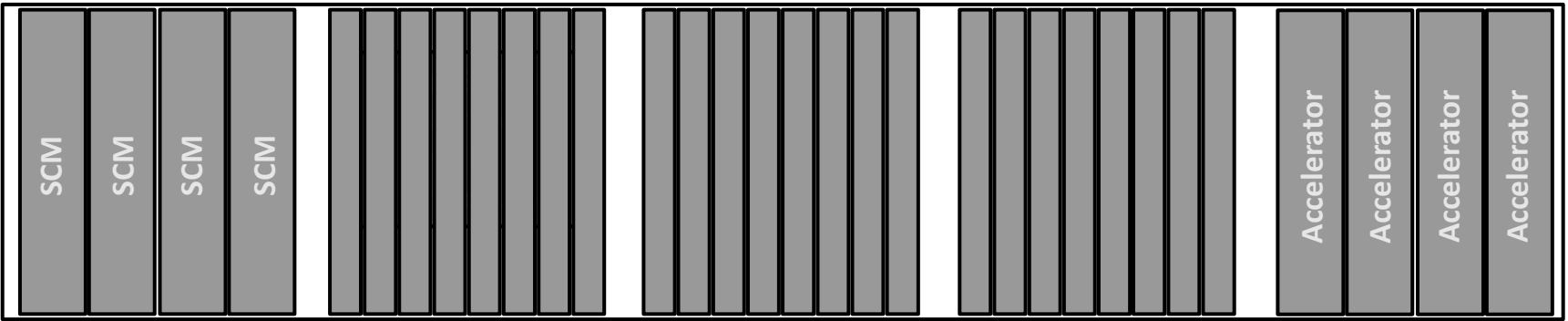
Storage
Config

44x E3L Devices



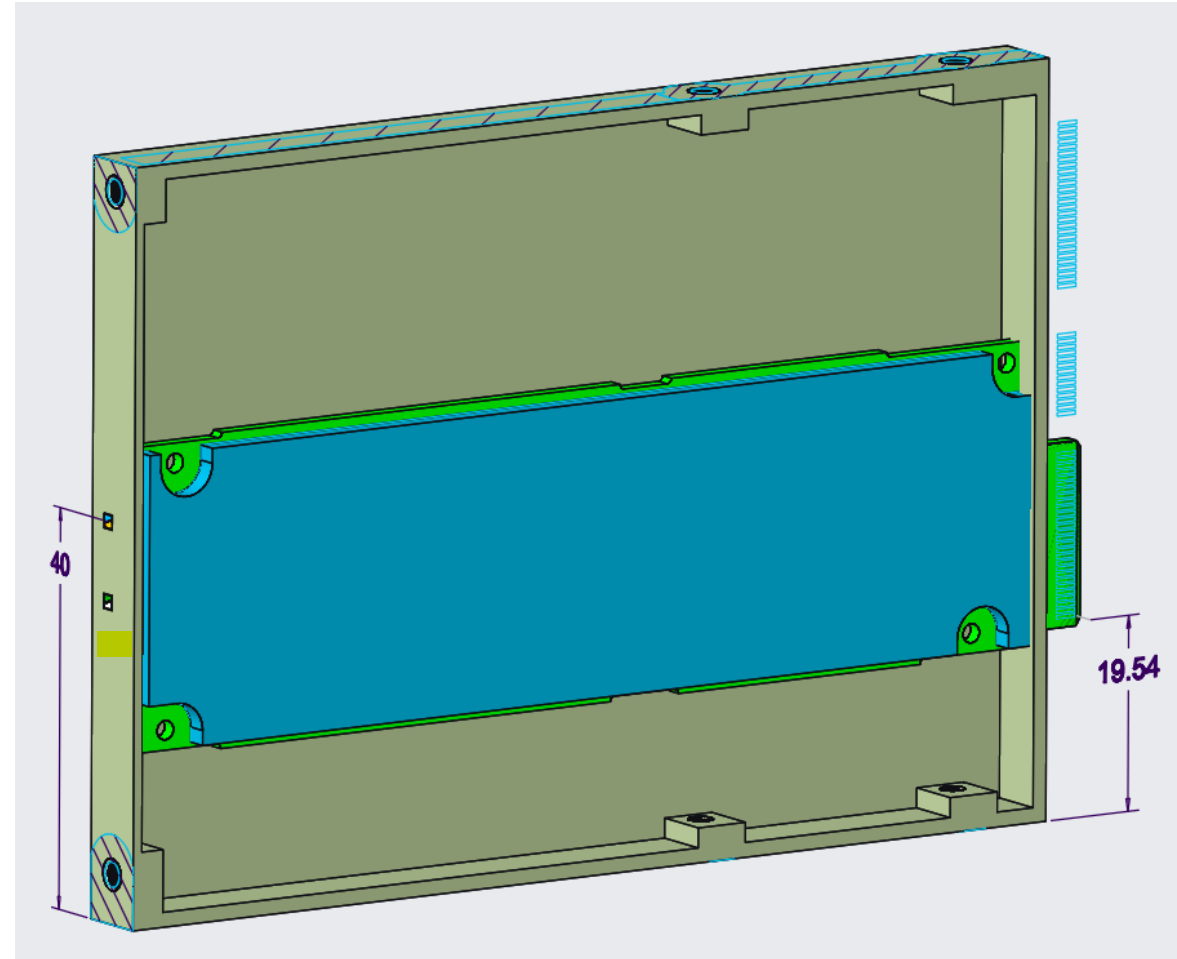
Alternate Device
Config

24x E3S Devices and 8x
SCMs or accelerators



E1 in E3 Case

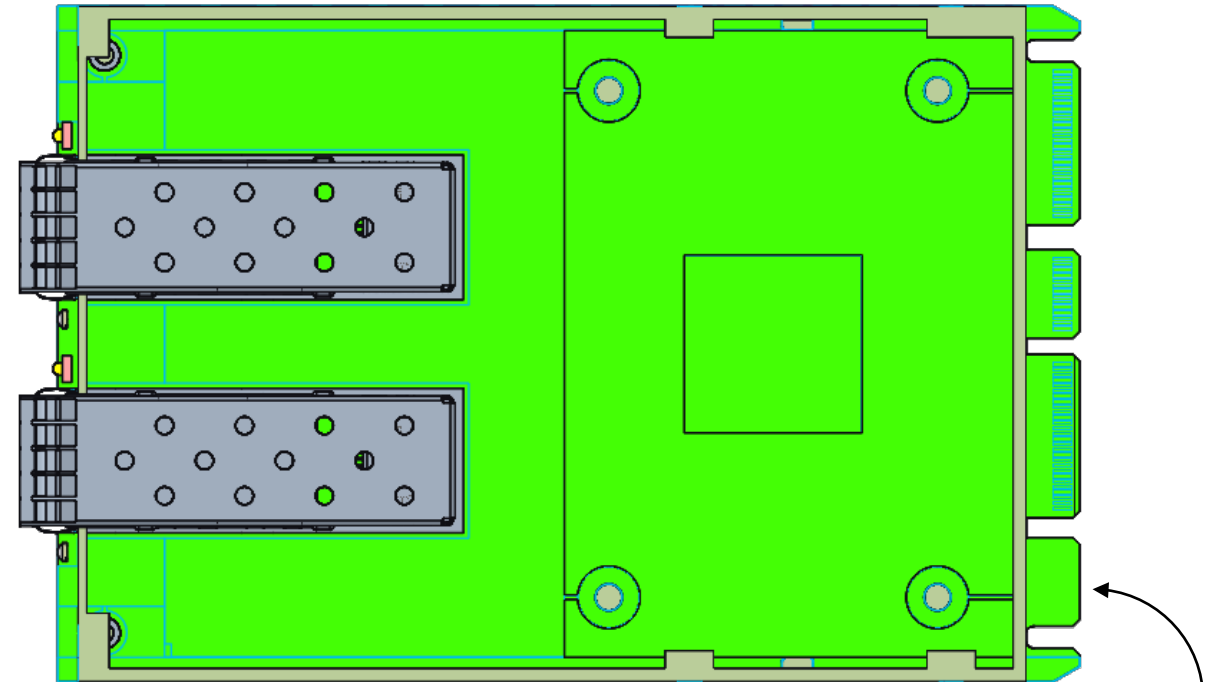
Enables E1.S PCBA designs to be leveraged for lower capacity E3 drives, combining volumes and lowering costs



Raise the E3 connector to 19.54mm so that the E1 PCB clears mechanical interferences and allows centering of the alternate set of LEDs.

Future Device Types

- Moving the E3 connector to 19.54mm allows for the use of a 4C+ connector used by the OCP NIC 3.0
- Should allow leverage of an OCP NIC 3.0 into an E3 case
- Also allows for higher power devices (3x, 4x, etc...) thicknesses



Additional connector space could be used for a 4C+ or a future higher power connector tab

Thank
You



**Hewlett Packard
Enterprise**

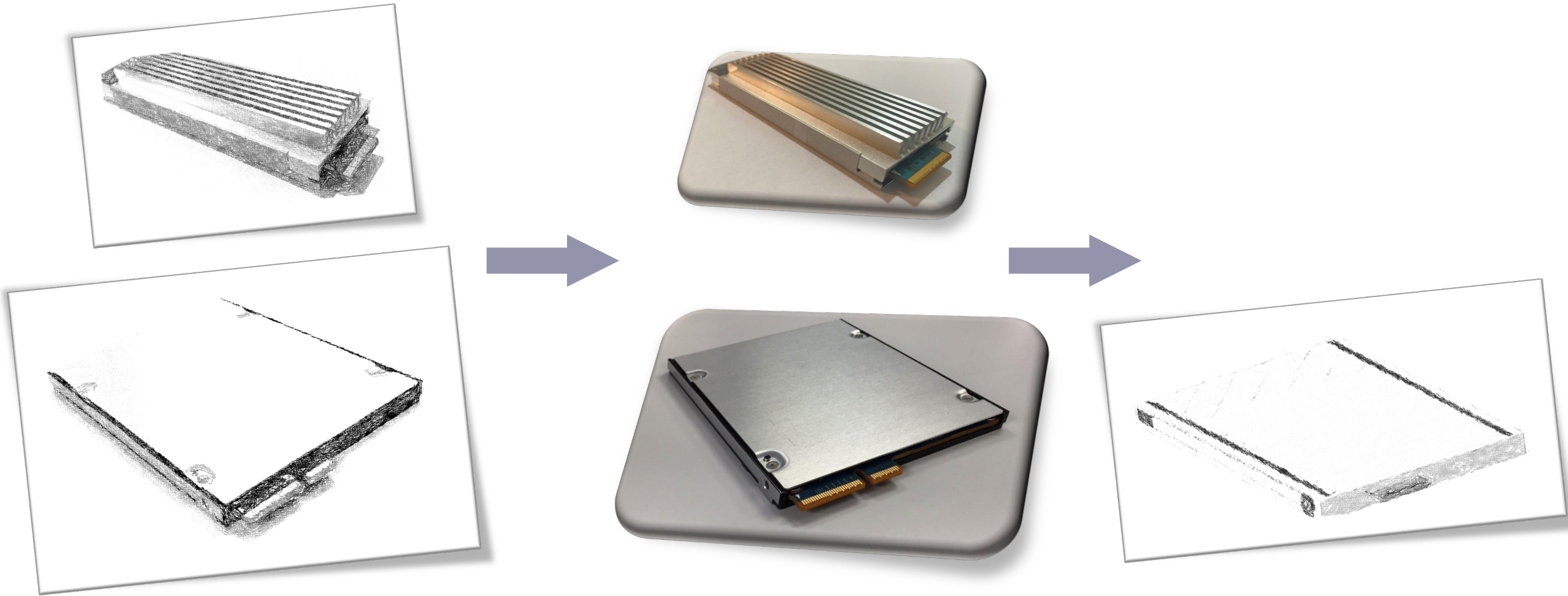


John Geldman

Director, SSD Industry Standards at KIOXIA

- Member Board of Directors, NVM Express
- Currently an active contributor to the following standards organizations:
 - NVM Express, INCITS T10, INCITS T13, JEDEC, OCP, PCI-SIG, SATA IO, SNIA, IEEE SISWG
 - In addition, John's team members are also active in CXL, DMTF, TCG
- Corporate leadership responsibility for standards for multi-billion dollar storage vendors since 2011
- Involved in storage standards since 1992, with an early introduction to standards including the transition from X3T9 to ATA, SCSI, PCMCIA, and CardBus
- An active FMS CAB member for at least 10 years

Standards are doing their job: helping us get from here to there!!!



A Standards Developer Point of View on EDSFF for SSDs

- EDSFF is the upcoming form factor family of choice for SSDs
 - The duo of faster PCI Express interfaces and NAND improvements shift the functional requirements
 - The SFF-TA-1002 connector was designed from the ground up for signal integrity challenges of PCI Express® 32 GHz (Gen 5) and PAM4 (Gen 6)
 - The EDSFF form factors were designed from the ground up for higher thermal loads (with defined air flows and methodologies)
- The E3 reboot (the initial release is going away) is a positive change to increase its compatibility and its capabilities
- We are seeing that 2.5" drives may survive for a while in both Server and Storage systems but will run into challenges as:
 - System interface rates migrate to PCIe Gen5 and beyond and SSD power will grow past ~20W
 - PCIe Gen 5 and NAND improvements support higher bandwidth, and higher bandwidth comes at a linear power increase
 - HDD capacities continue to grow for bit cost competition (is a 3.5" HDD focus coming?)

EDSFF: A form factor family affair for SSDs

- News Flash: There continues to be demand for all of the EDSFF form factors for SSDs
 - It is not a case of one winner, all the form factors have significant use cases
 - Other than boot drives, M.2 use cases in servers migrate well to E1.S to realize better performance and higher capacity per drive.
 - 2.5" SSD use cases can migrate to E3.S in servers for better performance
 - E1.L and E3.L are creating new use cases (higher capacities, SCM, higher power accelerators)
 - Signal integrity for the evolving PCI Express interface and thermal capabilities are driving factors
- We (SSD users/makers) are not alone in these form factors
 - EDSFF SSDs will compete for EDSFF server slots for SCM, HDD, Accelerator, and Networking functions
 - The E3 Form Factors (SFF-TA-1008) are being redesigned to be compatible with OCP-NIC form factors

Some standardization work in progress

- The E3 reboot is a non-compatible change to the 2018 SFF-TA-1008
 - The E3 reboot allows slot compatibility with OCP-NIC functions
 - The E3 reboot allows use of E1.S hardware in E3.S slots
 - This may be cost effective for “Low” capacity (e.g., less than 10 TB) needs are well served by existing E1.S SSDs
 - (It’s just fun to talk about “Low” capacity as less than 10 TB...)
- We are seeing standardization efforts to migrate HDDs into systems that support PCIe based EDSFF form factors
 - E3 Form Factors are an obvious option for 2.5” HDDs
 - PCI Express compatible storage transports are moving to support HDDs in multiple standards organizations

An observation of EDSFF SSD Interest

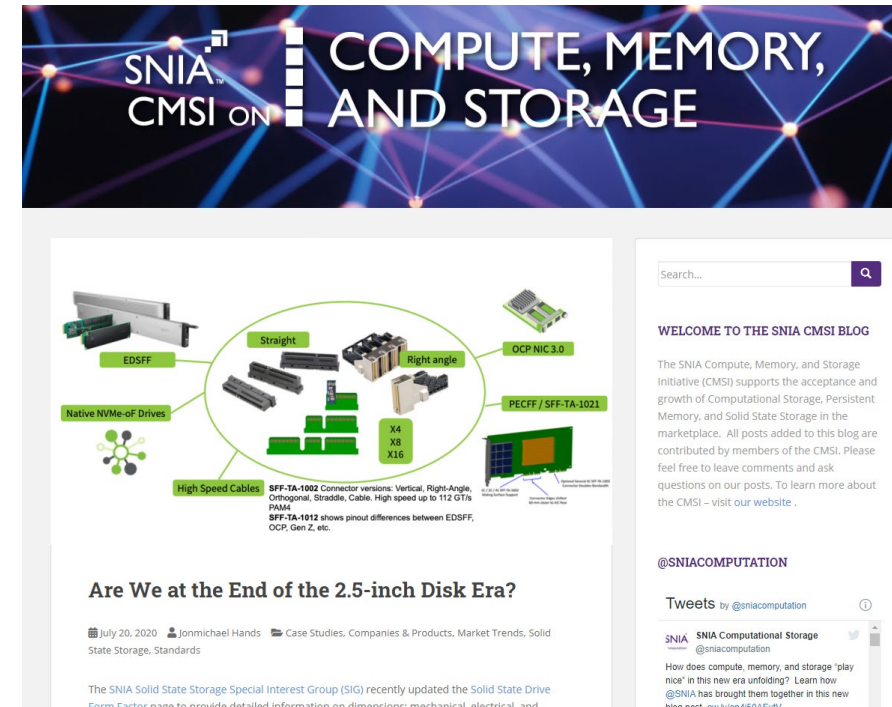
Form Factor	Application area
E1.S (SFF-TA-1006)	Interest in Hyperscaler and Enterprise markets • Form factor for M.2 replacement (for greater capacity and power handling) • Storage and compute applications • Edge servers • Width matters • 9.5 mm width works for many use cases • 15 mm width works well for the added thermal requirements of 4 lane PCIe® Gen 4 (OCP) • 25 mm width may be desired for more than 25 W (e.g., PCIe Gen 5 and accelerator use)
E1.L (SFF-TA-1007)	Interest in Hyperscaler market • Storage applications
E3.S (SFF-TA-1008)	Interest in Enterprise market • Server and storage NAND applications • Server SCM • Accelerator and PCIe Gen 5 power requirements may drive 2x width option
E3.L (SFF-TA-1008)	Interest in Enterprise market • Server and storage NAND applications • Storage SCM power requirements may drive 2X width option

Where To Find Out More About EDSFF

- Website resources
 - www.snia.org/forums/CMSI
- Twitter
 - [@sniasolidstate](https://twitter.com/sniasolidstate)
- Blog
 - [SNIAComputeMemory&Storage](http://www.snia.org/Blog)
- Videos
 - <https://www.youtube.com/user/SNIAVideo/playlists>
- Educational materials
 - <https://www.snia.org/educational-library>
- Joining SNIA and the Compute, Memory, and Storage Initiative
 - https://www.snia.org/member_com/join-SNIA

Finally, Thanks for Watching Our Webcast

- Please rate this [webcast](#) and provide us with feedback
- A link to this webcast and the PDF of the slides are posted to the SNIA Compute Memory and Storage Initiative website at <https://www.snia.org/forums/cmsi/knowledge/articles-presentations>
- You can also find this webcast (<https://www.snia.org/educational-library/enterprise-and-data-center-ssd-form-factor-end-25-inch-disk-era-2020>) and many other videos and presentations on today's topics in the [SNIA Educational Library](#)
- A Q&A from this webcast will be posted to the SNIA CMSI on Compute, Memory, and Storage blog: www.sniasssiblog.org



Questions?

Thank you!