Storage

Device.next

Microsoft, Dell EMC, HPE

December 12, 2019

Today's Storage Device Landscape
One Drive Form Factor does not meet the needs of all use cases

- **Best $/GB** – 3.5" HDD
  - 3.5" Remains best $/GB for foreseeable future
    - Will be a required Server and Storage FF as long as these remain best $/GB
  - Moderate performance – Archive & Cold data
  - Large form factor mandates increased chassis depth

- **Performance** – 2.5" HDD & SSDs
  - Primary Performant Storage
  - Multiple interfaces - NVMe for best performance
  - Moderate Capacity
  - Reduced size originally driven to increase performance for rotating media
    - SSD was an evolution into the legacy 2.5" form factor.
  - Modular systems require the short depth of a 2.5" device

- **Space Constrained Solutions** – M.2
  - Specialty implementations where larger storage form factors do not fit.
  - Boot media, Logging, Primary storage in specialty 1U and Edge servers.
  - Low Power with good enough performance and capacity.
  - Low cost entry point

Industry Architectural Requirements

Define a family of devices with a 10+ year lifetime

- Common connector definition (based on SFF-TA-1002)
  - Support for multiple protocols (PCIe, NVMe, 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 wide device to work in two narrow slots
  - Allow smaller devices to work in larger cases (Russian Dolls)
  - Requires wide devices to be asymmetric (which also allows for dual PCB implementations)

Cohesive Family of Devices

Storage Devices
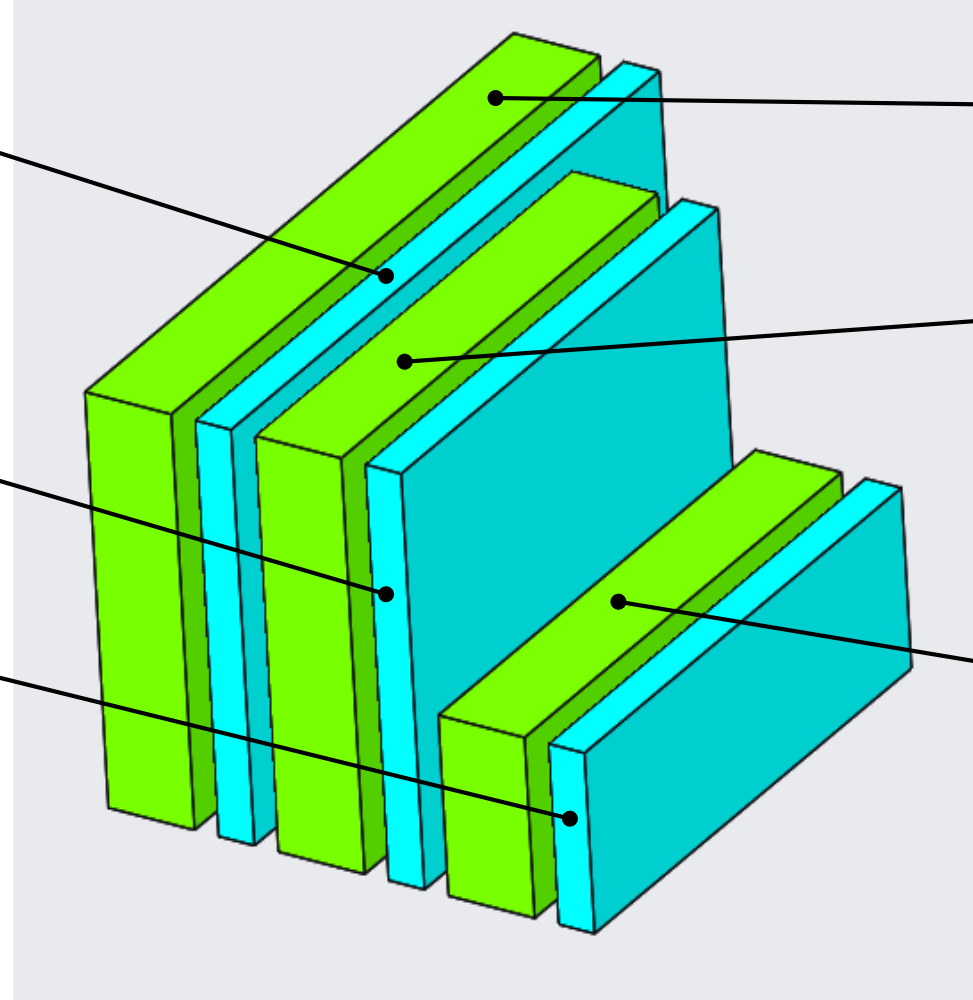
2U, Long, Narrow (40W)
- High capacity NAND storage
- Dense and storage centric chassis

2U, Short, Narrow (25W)
- 1U & 2U Performant primary NAND storage
- Modular and short chassis

1U, Short, Narrow (16W)
- Boot device
- Edge and small form factor chassis
- 1U "Good Enough" NAND storage

High Power Devices

2U, Long, Wide (70W)
- Higher end SCM implementation
- Higher end accelerator
- Computational storage
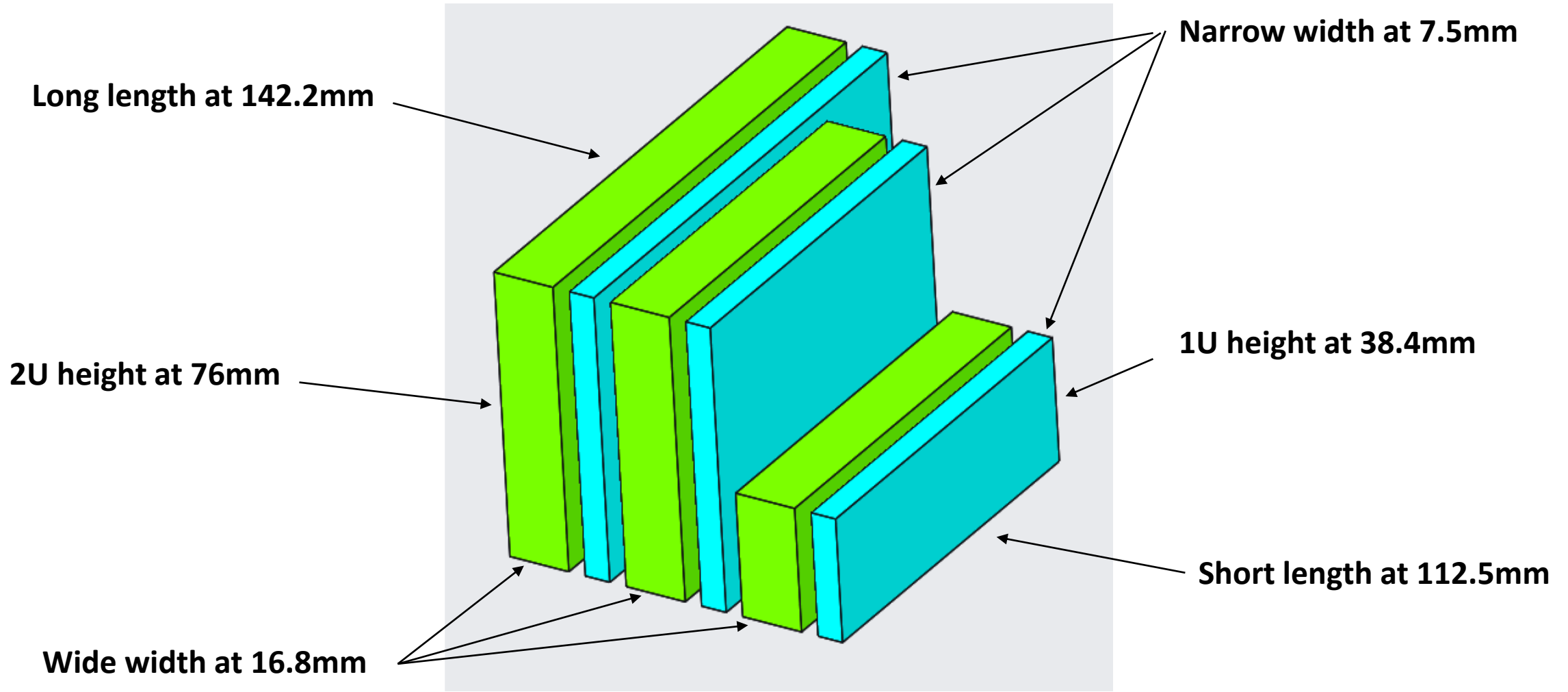
2U, Short, Wide (40W)
- Mid-range SCM implementation
- Mid-range accelerator
- Computational storage

1U, Short, Wide (25W)
- High performance, low density 1U NAND storage
- Minimum viable SCM implementations
- Low end 1U accelerator

Interoperable Device Sizes

Smaller devices fit within larger envelopes

Narrow width at 7.5mm

Long length at 142.2mm

1U height at 38.4mm

2U height at 76mm

Short length at 112.5mm

Wide width at 16.8mm

Device pitch of 9.3mm allows for a 1.8mm air gap

Common Connector Placement

Common connector registration allows for smaller devices to be used in larger device bays

- 1U x4 device fits in a 2U x4, x8, or x16 slot
- 2U x4 or x8 device fits in a 2U x8 or x16 slot
- Higher lane count devices work in lower lane count slots
- A wide device will work in a pair of narrow slots
- Requires wide devices be asymmetric
- Asymmetric design allows for dual PCB implementations in wide devices

Rear View

PCIe x16   PCIe x8   PCIe x4
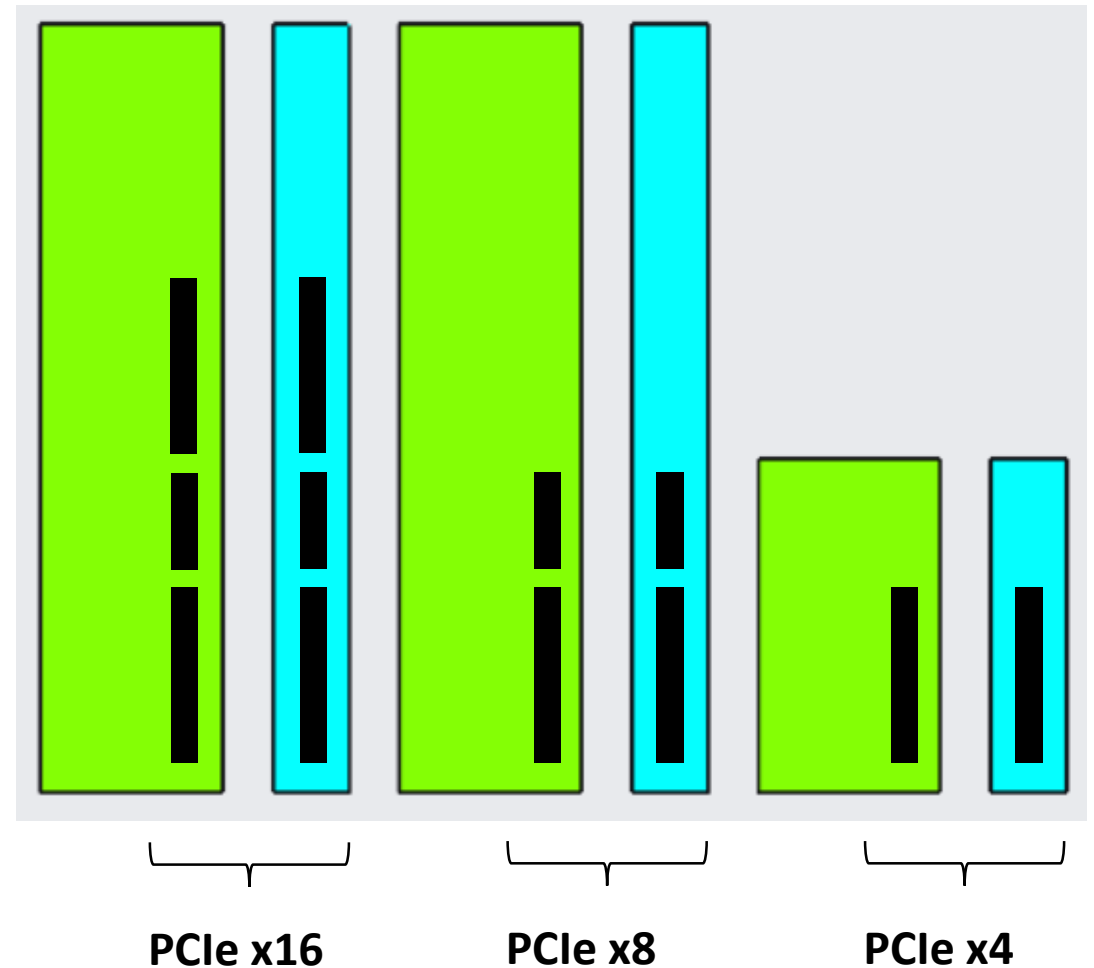
Common LED placement for SSD

Common LED placements allow for smaller devices to be used in larger device bays while maintaining common ID look and feel

- 1U short/narrow fits in a 2U short/narrow slot
- 1U short/wide fits in a 2U short/wide slot
- 1U short/wide fits in two 1U short/narrow slots
- Short devices fit in long slots (with light-pipes)

Front View

Green LED

Amber LED
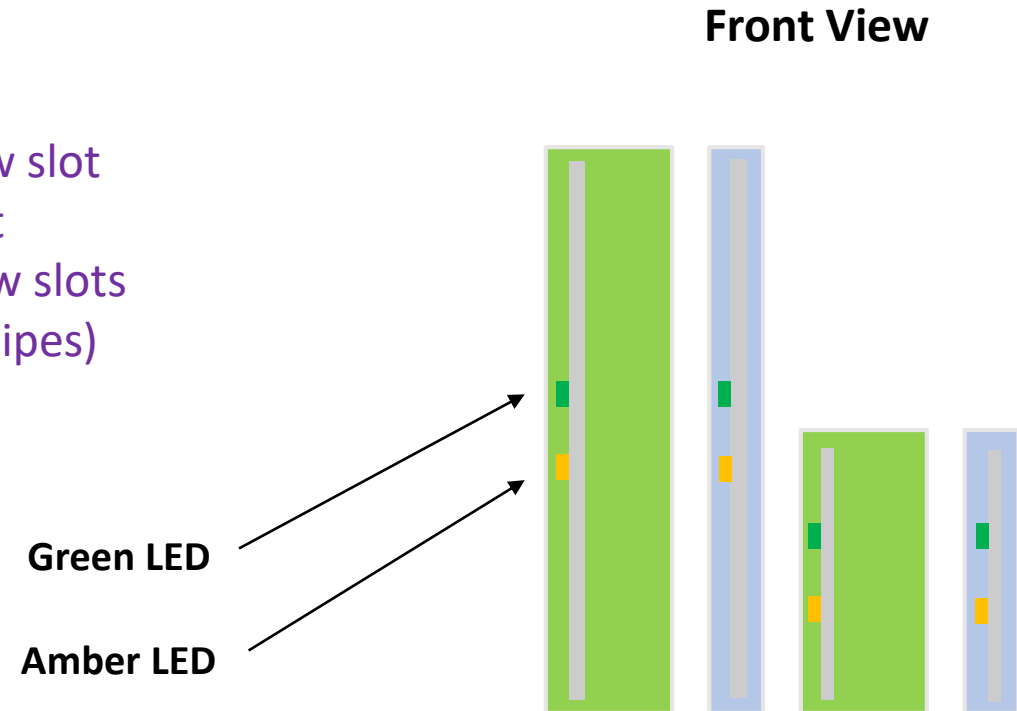
Potential Implementations (with cases)

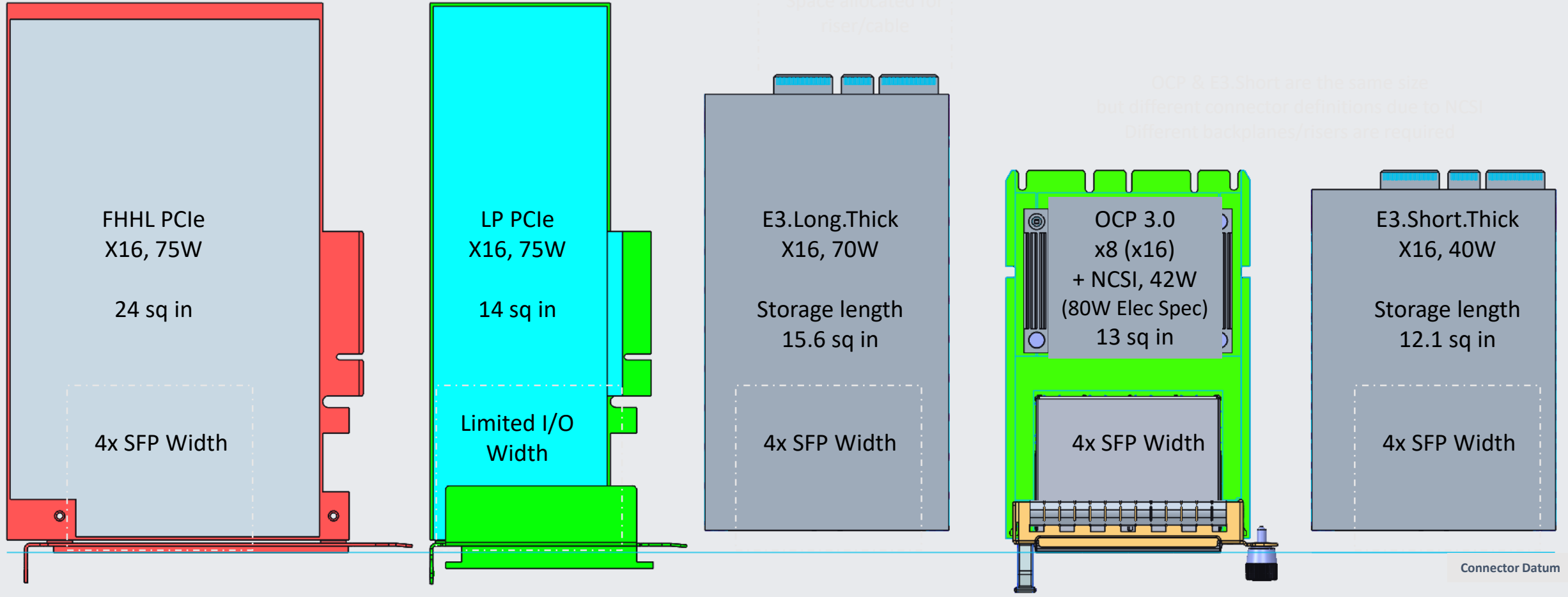|	1U-SN	1U-SW	2U-SN	2U-SW	2U-LN	2U-LW
Height (mm)	38.4	38.4	76	76	76	76
Length (mm)	112.5	112.5	112.5	112.5	142.2	142.2
Thickness (mm)	7.5	16.8	7.5	16.8	7.5	16.8
Power (W)	16	25	25	40	40	70
Interface Width	4/8	4/8	4/8/16	4/8/16	4/8/16	4/8/16
Target NAND placements	8 to 12	n/a	16 to 24	n/a	32 to 40	n/a

1U-SN: 1U short/narrow
1U-SW: 1U short/wide
2U-SN: 2U short/narrow
2U-SW: 2U short/wide
2U-LN: 2U long/narrow
2U-LW: 2U long/wide

Primary use case for narrow devices is NAND based SSDs (with the exception of 1U-SW)

Topics under discussion

- 1U short/wide width
  - Discussion around changing from 16.8mm to 15.0mm

- Mounting holes
  - Goal is to have 1U devices fit into 2U slots
  - Need to come up with a mounting scheme that allows a 1U device to be mounted into a 2U carrier

- LED placement
  - Goal is to have a common ID look & feel across all devices

PCIe Device Form Factors

FHHL PCIe
X16, 75W

24 sq in

4x SFP Width

LP PCIe
X16, 75W

14 sq in

Limited I/O
Width

E3.Long.Thick
X16, 70W

Storage length
15.6 sq in

4x SFP Width

OCP 3.0
x8 (x16)
+ NCSI, 42W
(80W Elec Spec)
13 sq in

4x SFP Width

E3.Short.Thick
X16, 40W

Storage length
12.1 sq in

4x SFP Width

Space allocated for
riser/cable

OCP & E3.Short are the same size
but different connector definitions due to NCSI
Different backplanes/risers are required
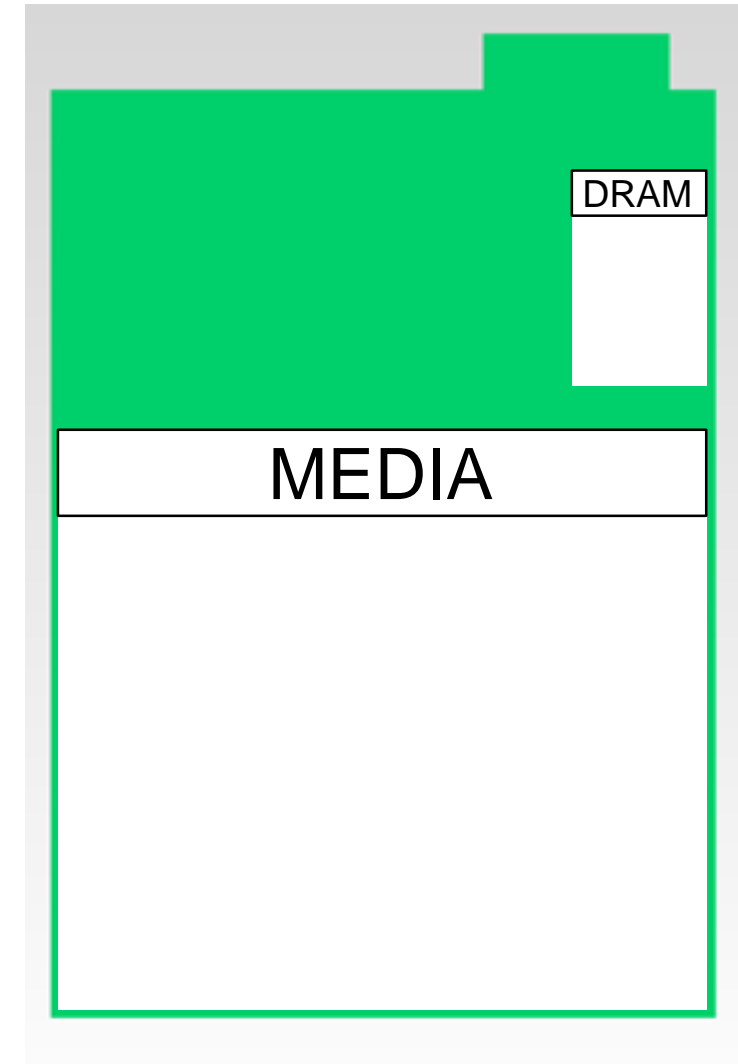
Connector Datum

Dell/EMC Thermal Study

EDSFF Short Component Zones

- Component layout zones used in modeling

- 3" short: 25 W total power
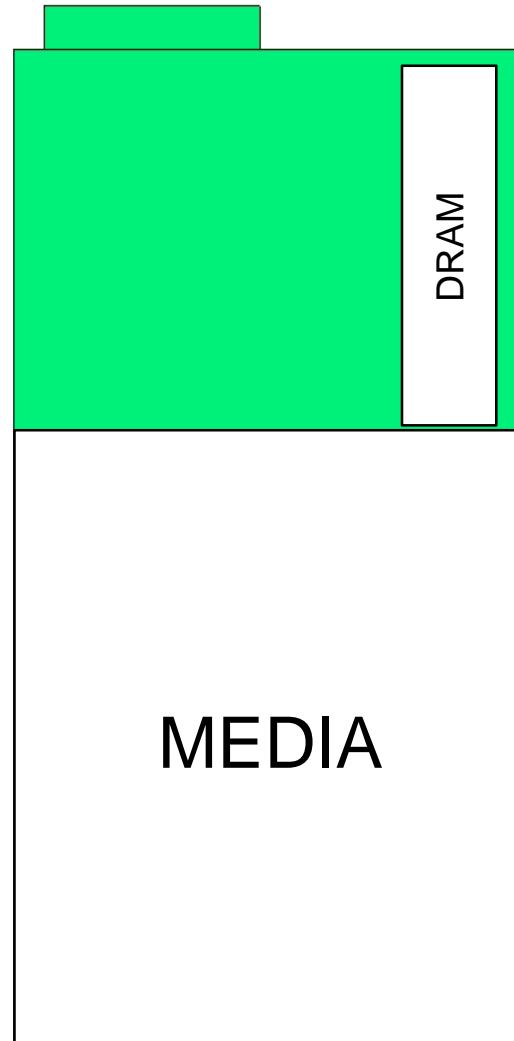
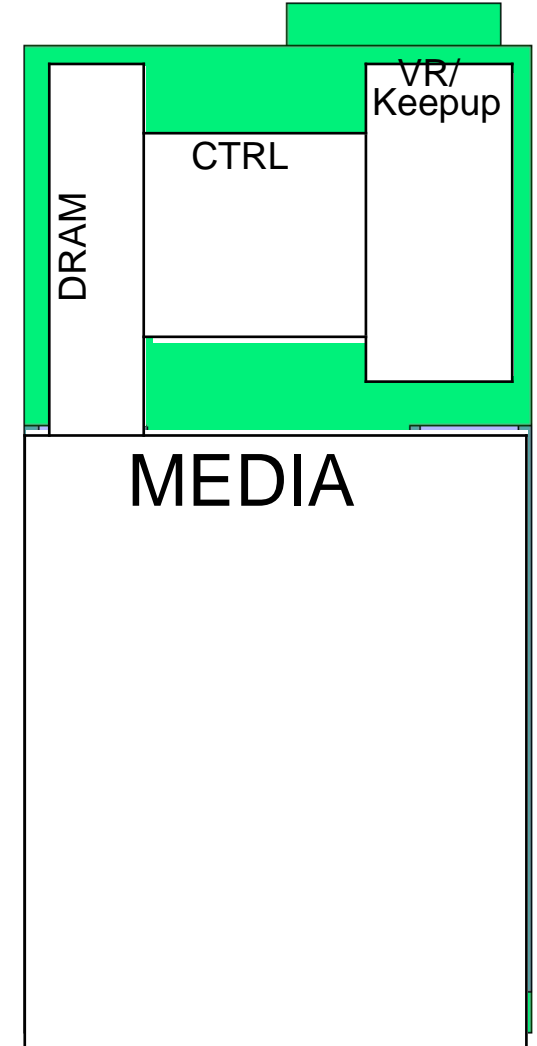- x4 Connector Modeled
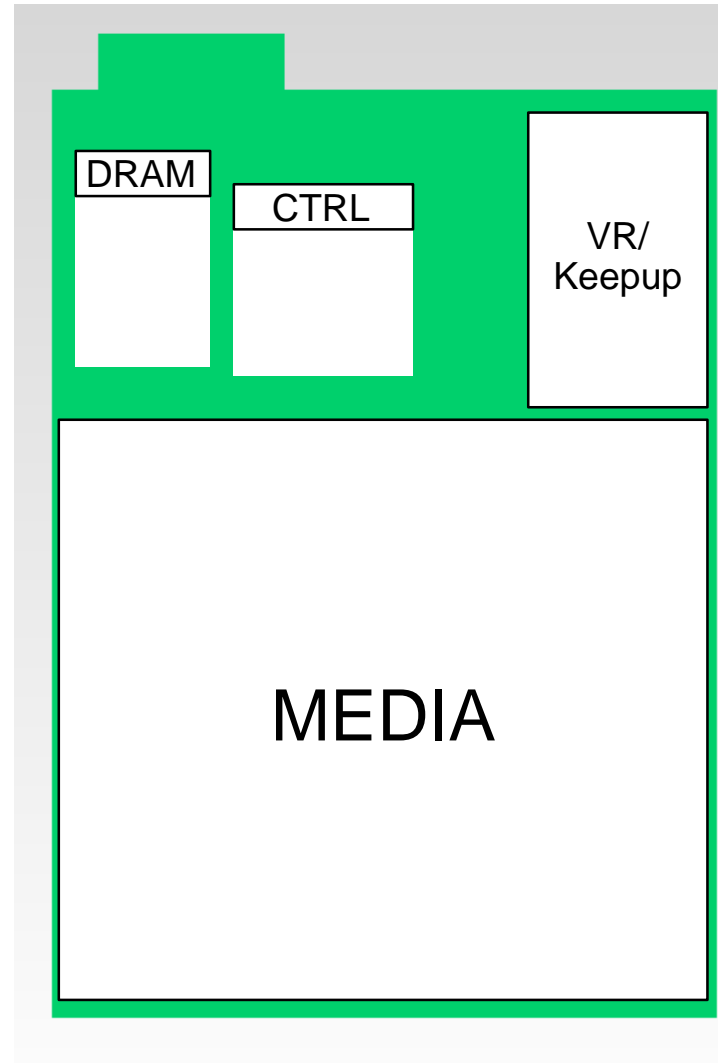
Top View

Bottom View

EDSFF Long Component Zones

- Component layout zones used in modeling

- 3" long: 35 W total power

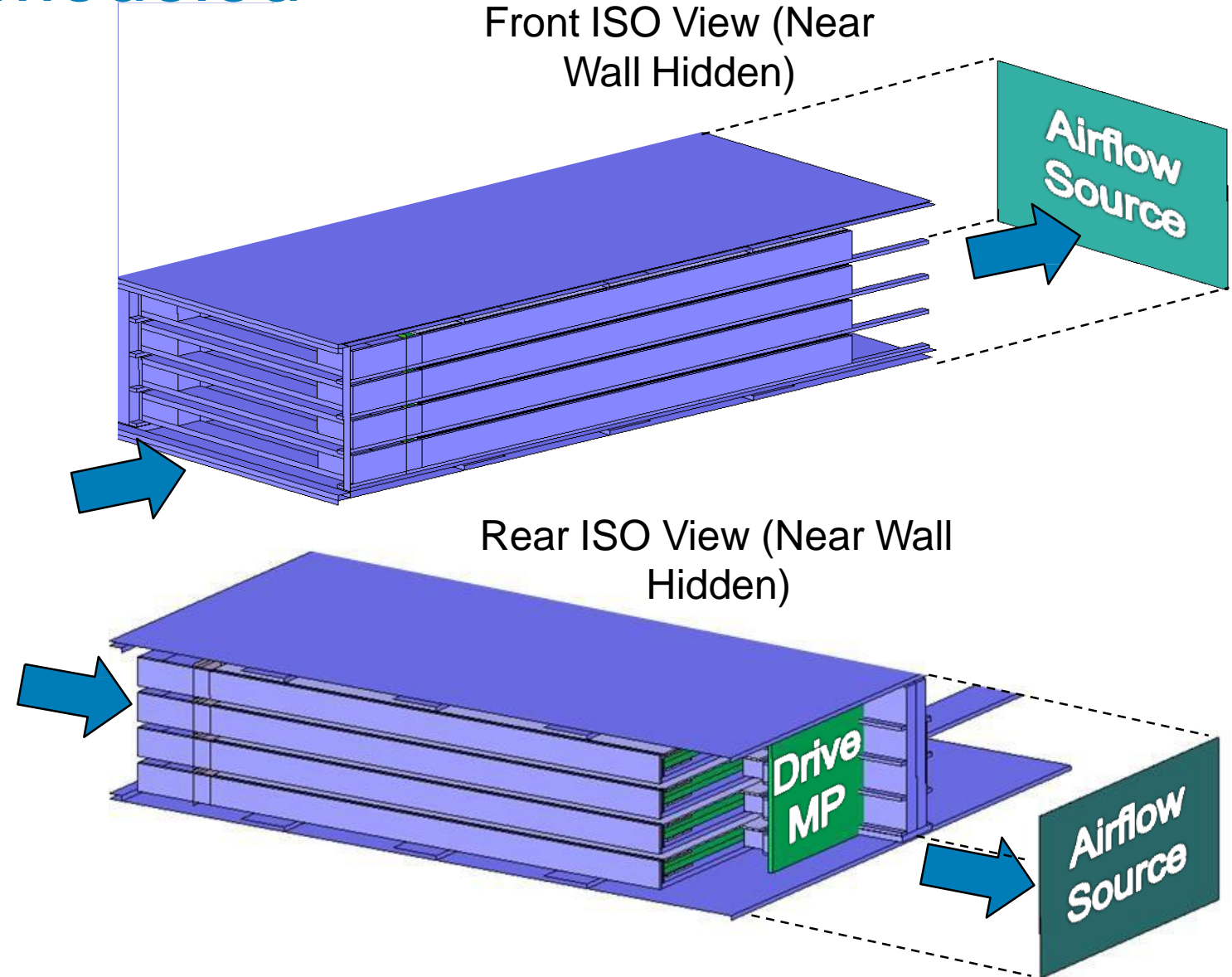- x8 Connector Modeled

Top View

Bottom View

Solution Domain Modeled

- Domain modeled was a 4 module stack in a 1U chassis height

- Represents 1/5 width of a chassis (5 wide x 4 module tall config)

- Uniform airflow source

  - downstream of modules

- Inlet conditions:
  35°C, 950m altitude

- 3" Short uses x4 connector/drive MP

- 3" Long uses x8 connector/drive MP

Front ISO View (Near Wall Hidden)

Airflow Source

Rear ISO View (Near Wall Hidden)

Drive MP

Airflow Source

Touch Temperature

- UL62368
  - New touch temp spec in place now or coming into place soon 2019 EU
  - Drives while hot should <u>not</u> be accessible by a user in a consumer application (TS1), which would be **70℃**
    Drive transient simulations suggest hotspot cool-down of ~20C in 1 min
  - Drive should be considered a TS2 device, a device that would be touched while hot by someone who is "instructed"→ **80℃** maximum touch temperature for metal surfaces

Component Limits

- NAND:
  - 85C Junction

- Controller:
  - 105C Junction

- Enclosure Temperature:
  - 80 ℃ for metallic surface

EDSFF 3" Short Temperatures

- The plot to the right shows key component temperatures for the EDSFF 3" Short proposal

- The 80C enclosure temperature required the highest airflow to achieve

Volume Flows [CFM]		
To 105C TJ, CONT	To 85C TJ, NAND	To 80C Max. Enclosure
1.35	1.75	2.40

- At 2.4 CFM / module and 25W each, temperature rise would be 21C at 950m altitude
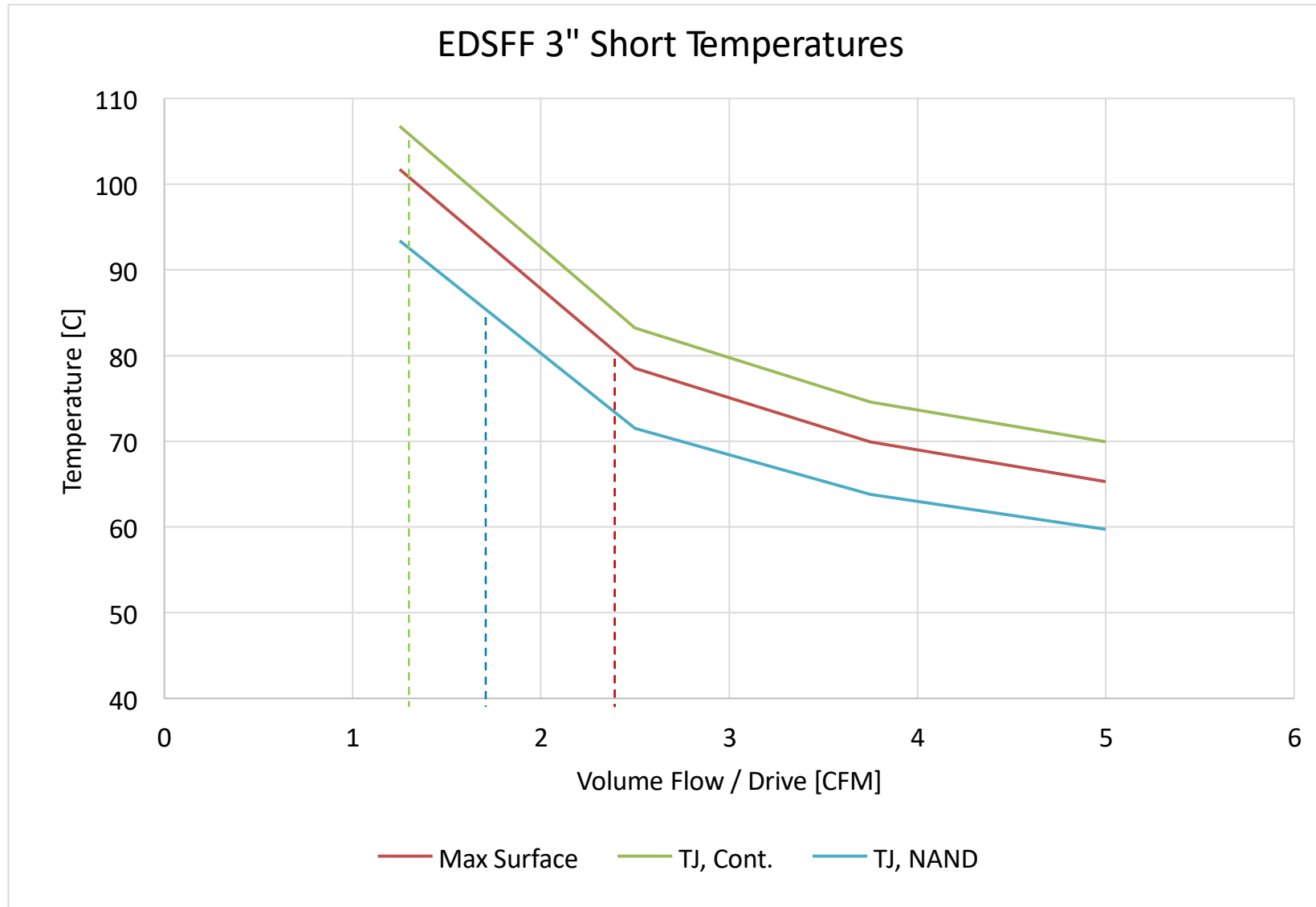

EDSFF 3" Short Temperatures

EDSFF 3" Long Temperatures

- The plot to the right shows key component temperatures for the EDSFF 3" Long proposal

- The 80C enclosure temperature required the highest airflow to achieve

Volume Flows [CFM]		
To 105C TJ, CONT	To 85C TJ, NAND	To 80C Max. Enclosure
2.05	2.20	3.65

- At 3.65 CFM / module and 35W each, temperature rise would be 20C at 950m altitude

EDSFF 3" Long Temperatures
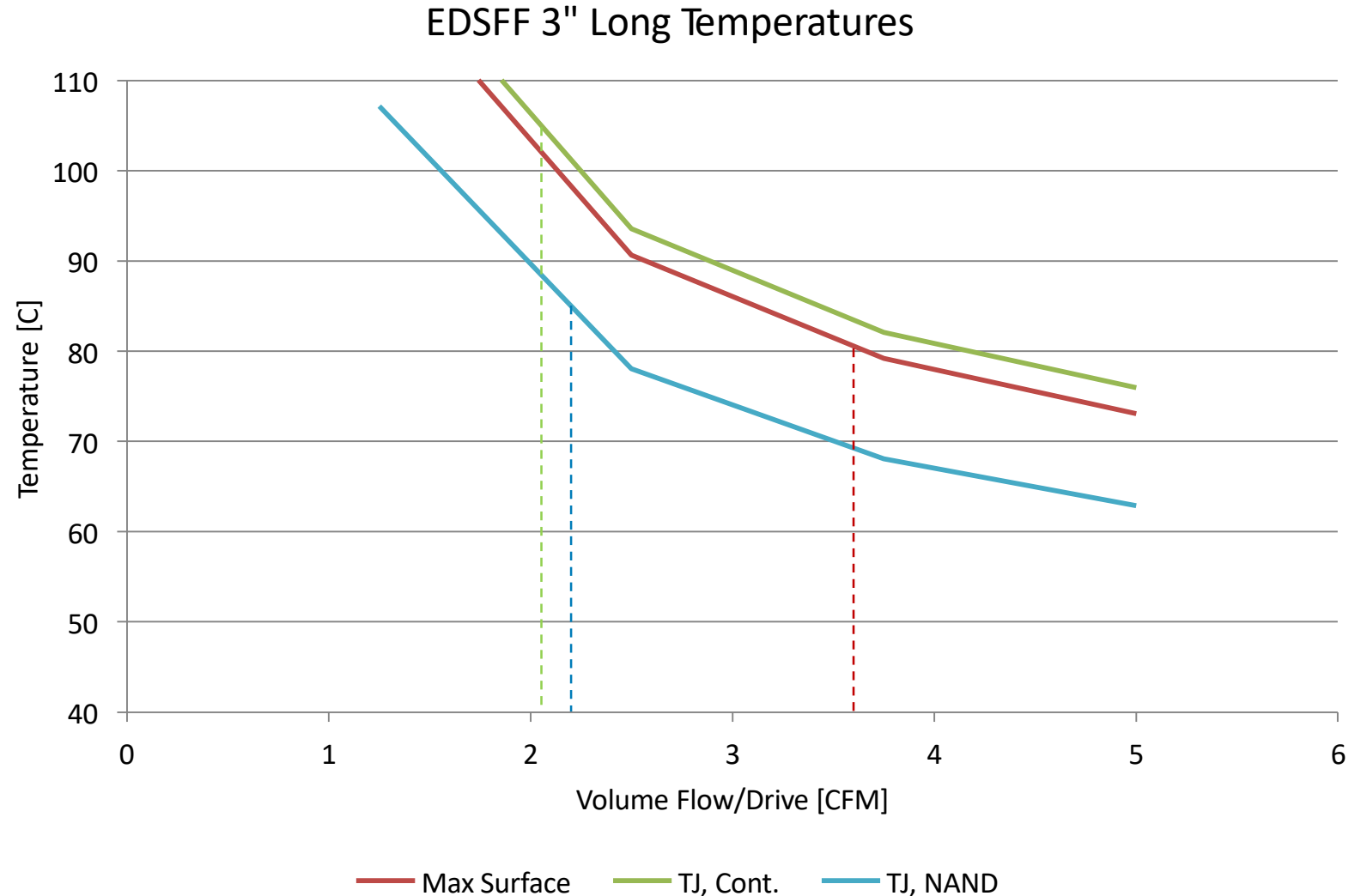
Pressure Drop

- The plot shows the comparison of measured pressure drop of 25 U.2 drives + midplane in a 2U enclosure v. 40 EDSFF 3" short and long modules + drive connector boards simulation data

- Simulated data has been translated to Sea Level Pressure

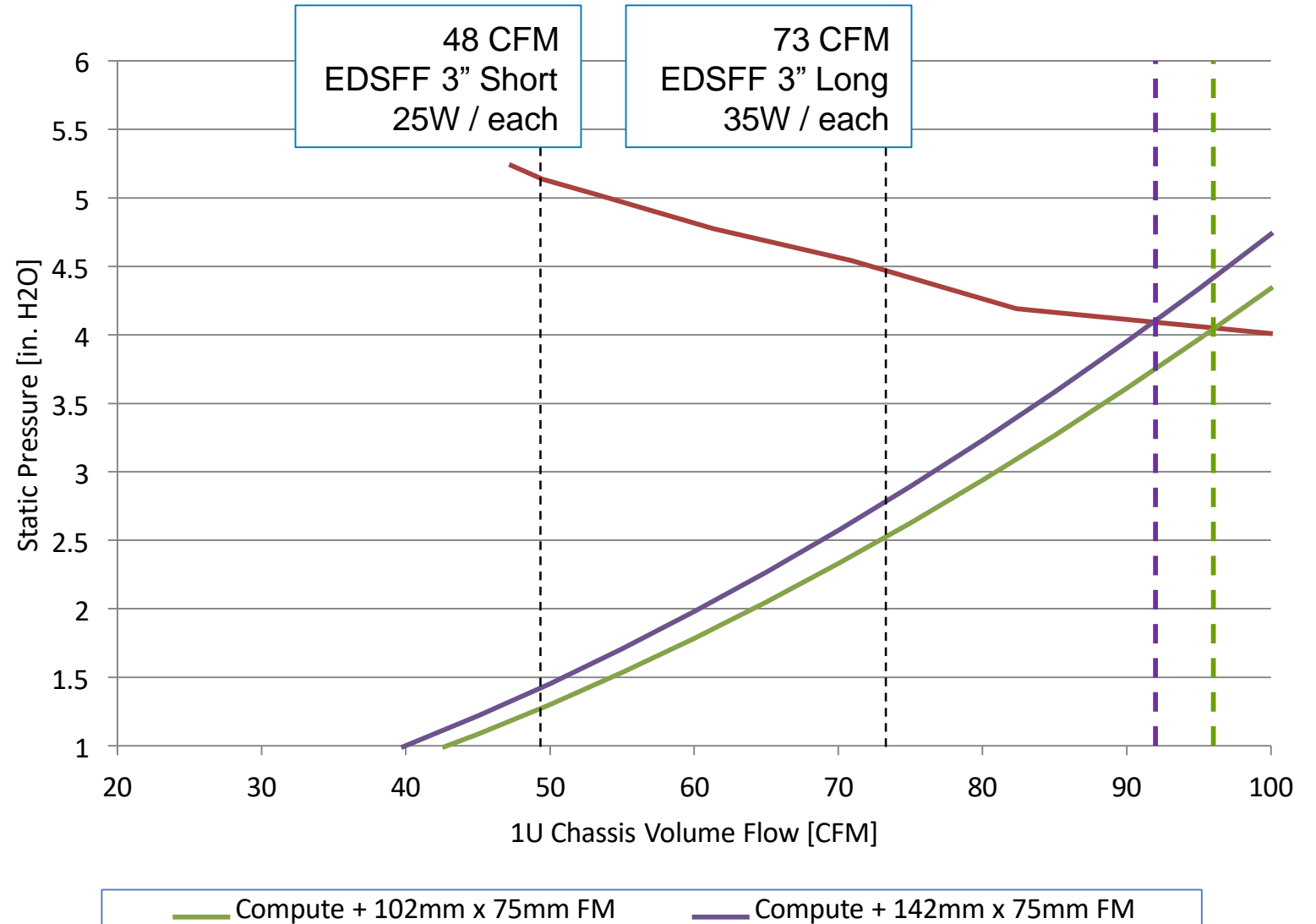- The EDSFF 3" solutions are expected to have lower pressure drop

Pressure Drop Comparison to Standard 2U

Y-axis: Pressure Drop [in H2O]
X-axis: Volume Flow [CFM]

Legend:
- 25 Drives + Midplane
- 40x EDSFF 3" Short 7.5
- 40x EDSFF 3"long7.5
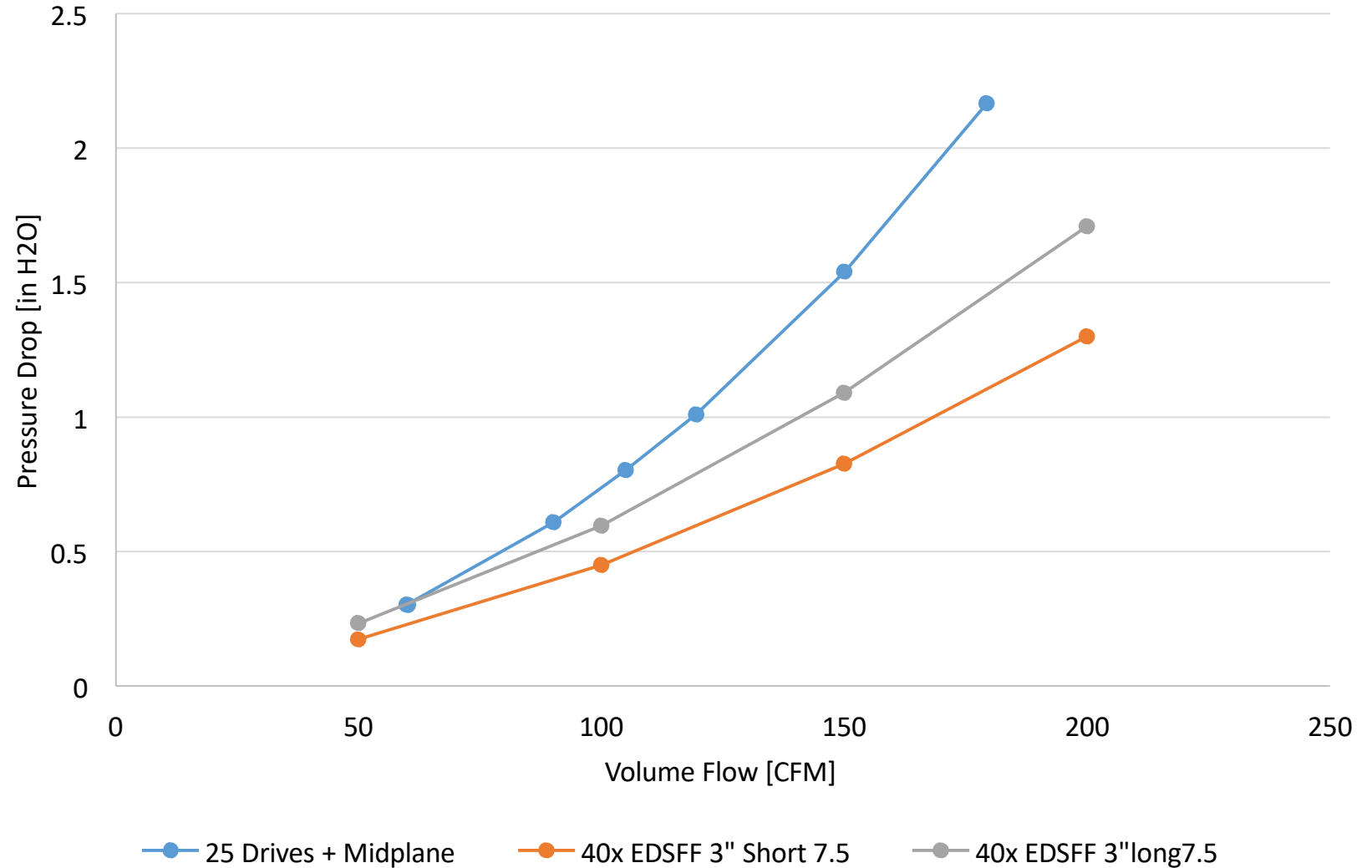
Airflow Feasibility in 1U

- These EDSFF 3" pressure drops are combined with a 1U compute and IO pressure drop with an <u>aggressive</u> fan curve.

- These show that the expected 1U chassis flows are:
  - 92 CFM for EDSFF 3" Long
  - 96 CFM for EDSFF 3" Short

- The ~73 CFM required for the EDSFF 3" Long in 1U should be achievable

1U Server Airflow Resistance v. Source Curve

48 CFM
EDSFF 3" Short
25W / each

73 CFM
EDSFF 3" Long
35W / each

Y-axis: Static Pressure [in. H2O]

X-axis: 1U Chassis Volume Flow [CFM]

— Compute + 102mm x 75mm FM   — Compute + 142mm x 75mm FM
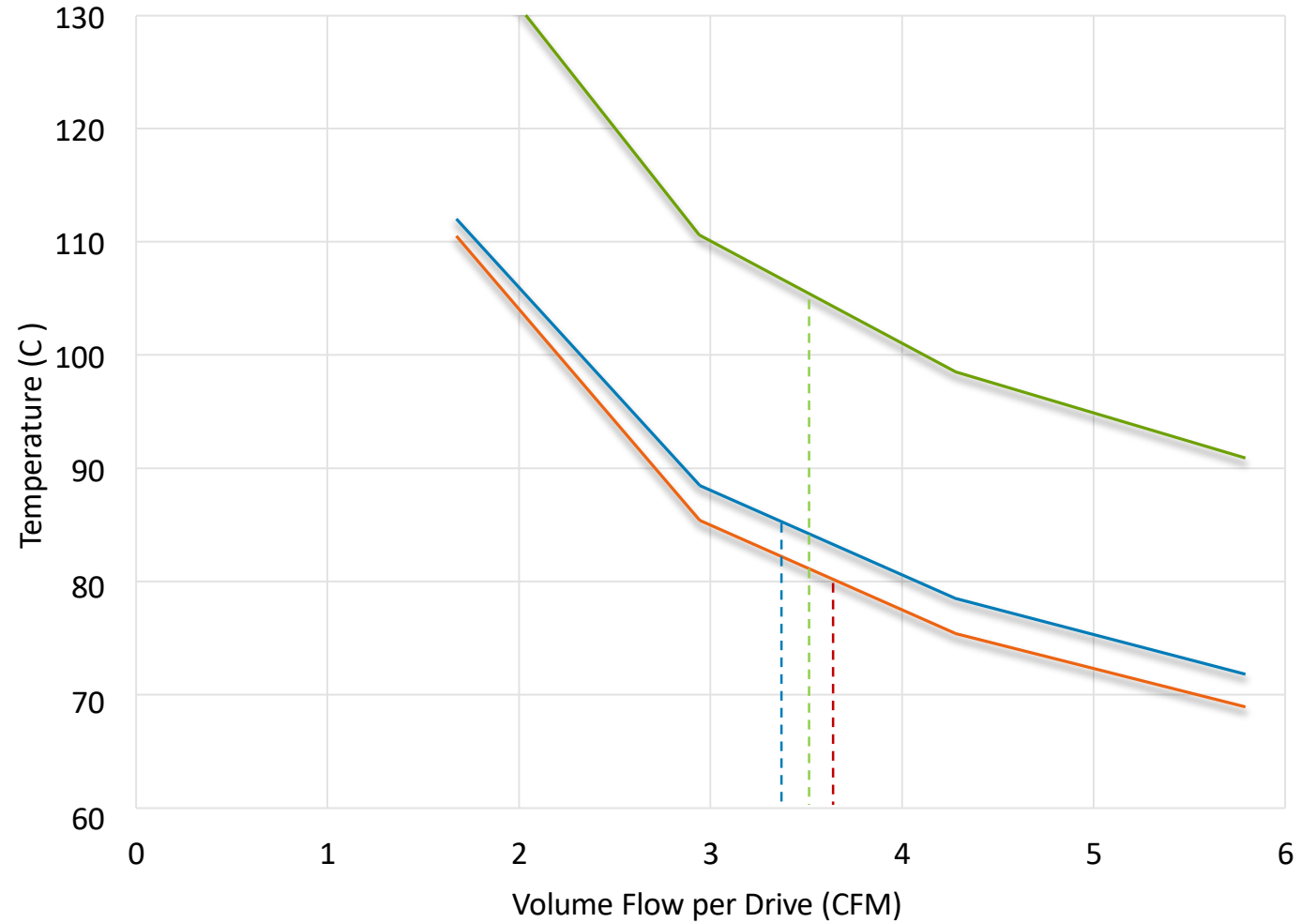
EDSFF 3" 16.8mm
Short Temperatures

- The plot to the right shows key component temperatures for the EDSFF 3" Short 16.8mm proposal

- Below are the volume flows to each component limit

| Volume Flows [CFM]		
To 105C TJ, CONT	To 85C TJ, NAND	To 80C Tc surface
3.4	3.3	3.5

EDSFF 3" 16.8mm Short Temperatures

EDSFF 3" 16.8mm Long Temperatures

- The plot to the right shows key component temperatures for the EDSFF 3" Long 16.8mm proposal
- Below are the volume flows to each component limit

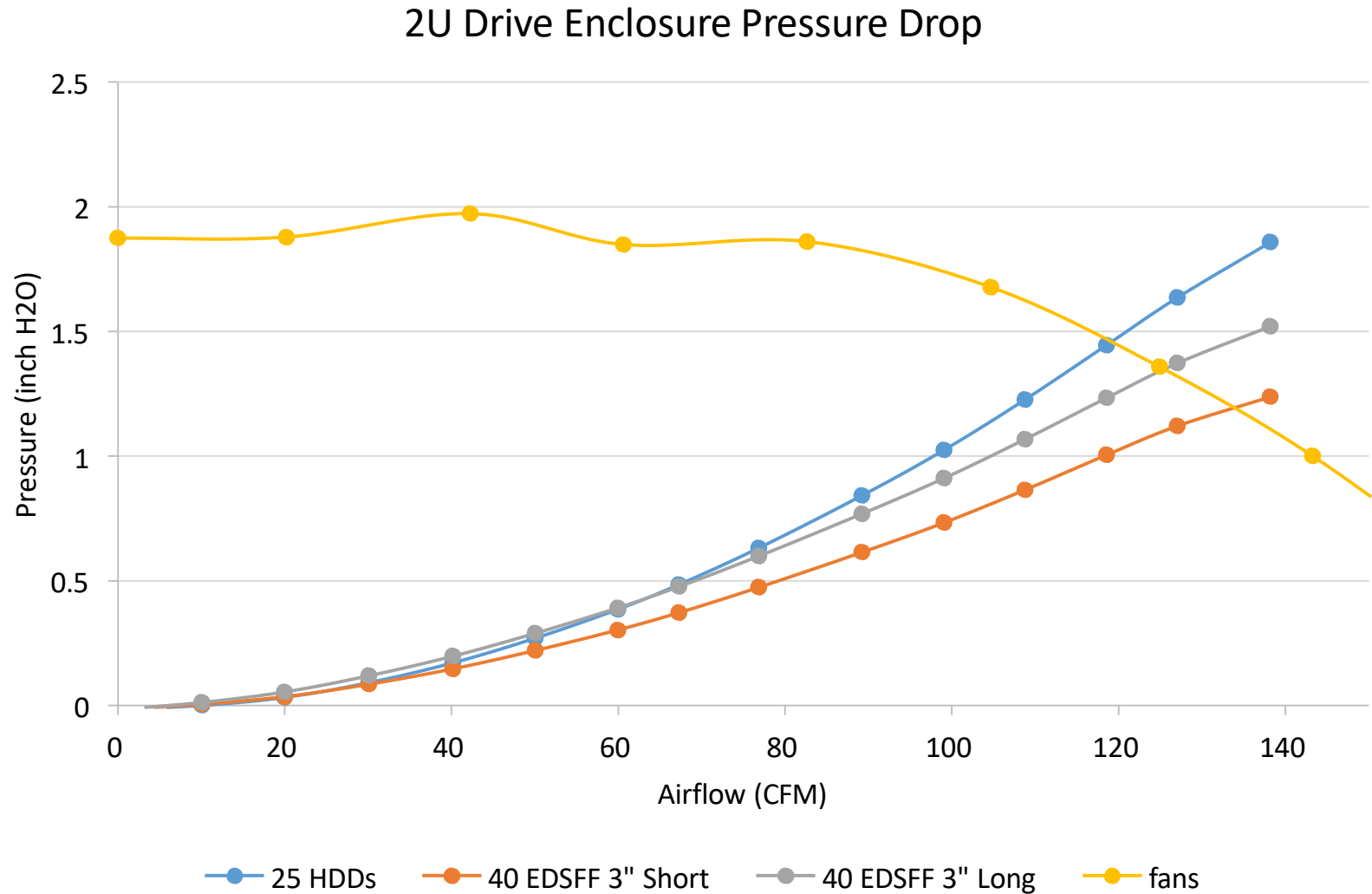| Volume Flows [CFM]		
To 105C TJ, CONT	To 85C TJ, NAND	To 80C Tc surface
4.1	4.7	5

EDSFF 3" 16.8mm Long Temperatures
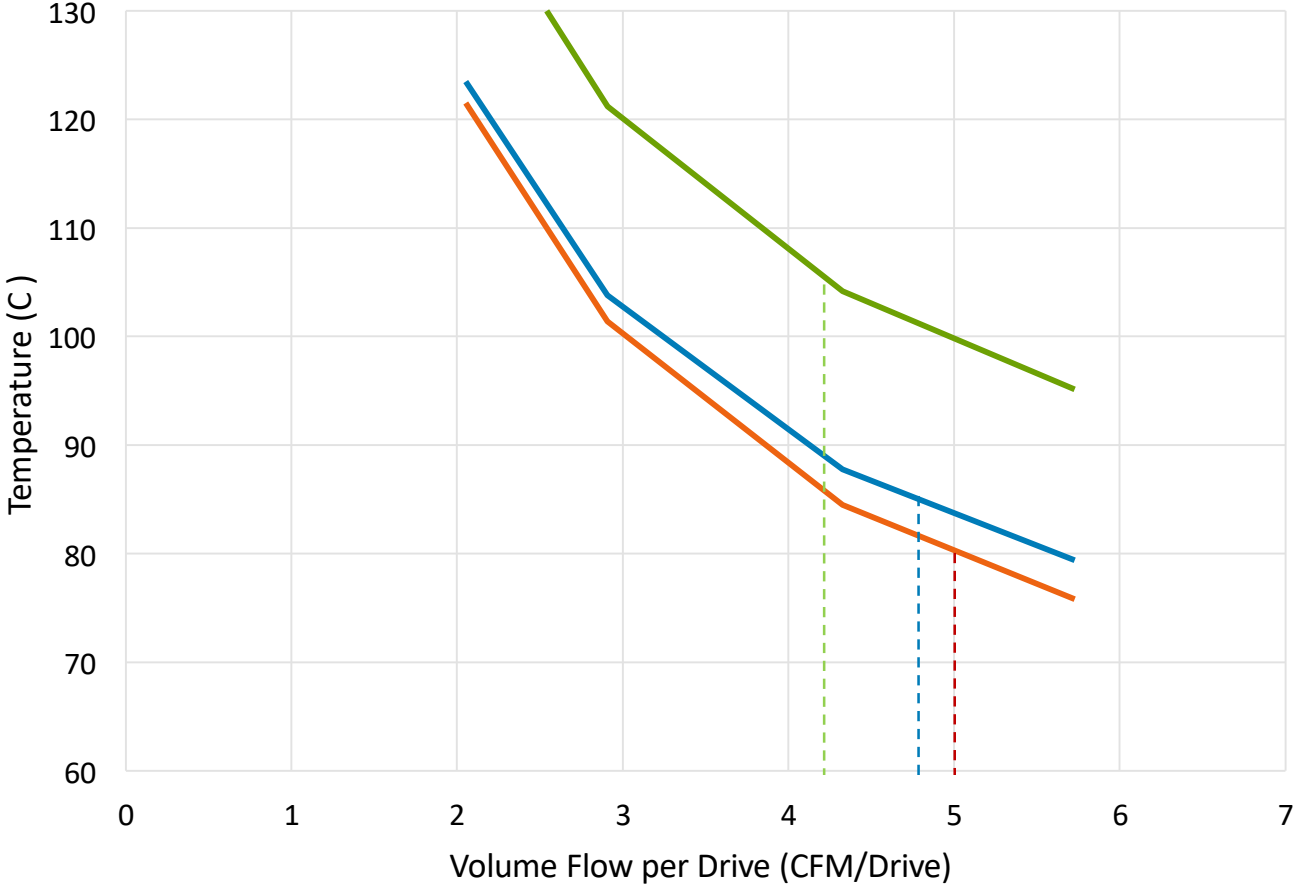
CTRL Tj   NAND Tj

System Level Impacts

- The low pressure drop of the drive module section of the chassis is expected to allow for an increase system airflow

- This additional airflow can enable additional drive power for the same drive exhaust temperature

System airflow	CFM	% Flow Increase
System with 25 HDDs	118	
System with 40 EDSFF 3" Long	125	6%
System with 40 EDSFF 3" Short	135	14%

2U Drive Enclosure Pressure Drop

Pressure (inch H2O) vs Airflow (CFM)

Legend: 25 HDDs — 40 EDSFF 3" Short — 40 EDSFF 3" Long — fans

Summary

- Required airflows for the 7.5"mm modules are updated from previous to accommodate 80C touch temperature.
  - At 25W, 48 CFM/U is required to maintain the EDSFF 3" Short at 35C, 950m
  - At 35W, 73 CFM/U is required to maintain the EDSFF 3" Long at 35C, 950m
  - These airflows should be achievable, especially considering the lower pressure drops

- The reduction in pressure drop observed in simulation should allow for increased heat generation compared to U.2 solutions.