WHY M.2 IS UNSUSTAINABLE
(AND WHAT CAN WE DO ABOUT IT)

LEE PREWITT
PRINCIPAL HARDWARE PROGRAM MANAGER
AZURE CSI - MICROSOFT
M.2 IN AZURE

• Each SKU uses multiple m.2s inside the C2010 Open Compute Chassis

• Each VM gets ~90GB of ephemeral storage space

• Each VM can use up to 4,000 IOPs
  • Any mix of 8KiB IOs (from 100% Random Read to 100% Random Write)
M.2 PROS

• Compact formfactor allows for flexible placement in the chassis

• Ubiquitous use across client and cloud allows for a robust supply chain

• ????
M.2 CONS

• Limited number of NAND die placements (4 with use of PLP)
• Limited power envelope
• Limited thermal envelope
• No hot-plug support
• Connector is not robust
• PCB is thin
  • Increases manufacturing cost and increases risk of PCB failures
So is there a better solution?
WHAT ARE THE OPTIONS?

**Adapter Cards**
- Good: High-perf Cache
- Bad: Steals PCIe slots from I/O
- Ugly: $$$, Might fit only two?

**Consumer SSD**
- Good: Small and Modular
- Bad: Low capacity, no hot-plug, connector less reliable
- Ugly: More expensive once adapted to enterprise

**HDD Form Factors**
- Good: Hot-plug, Storage features
- Bad: Designed for HDD, not SSD
- Ugly: Blocks airflow to the hottest components in server
EDSFF: A BRIEF HISTORY

• Enterprise and Datacenter SSD Form Factor

• Formed in 2017 to rethink the SSD from the ground up with a focus on Cloud and Enterprise use cases
OF COURSE THERE CAN NEVER BE ONLY ONE...

**E1.L (SFF-TA-1007)**
- Density Optimized
- 318.75 x 38.4 mm
- Supports > 40W
- Up to 48 Standard NAND sites

**E1.S (SFF-TA-1006)**
- 111.5 x 31.5 mm
- Up to 12 Standard NAND sites
- Supports >12W

**E3 (SFF-TA-1008)**
- Ultra high-performance applications
- (104.9/142.2) x 78mm
- Supports up to 70W
- Up to 48 Standard NAND sites

**EDSFF Advantages**
- **Same** Protocol: NVMe
- **Same** Interface: PCIe
- **Same** Connector: SFF-TA-1002
- **Same** Pinout and Functions

**Different Usages**
**Same** Expectations!
E1.L FOR AZURE STORAGE
### Table 7-1. Thermal guidelines for a 1U short system implementation

<table>
<thead>
<tr>
<th>Enclosure Parameter</th>
<th>5.9mm Device</th>
<th>Device with Heat Spreader (8.01mm)</th>
<th>Device with Symmetric Enclosure (9.5mm)</th>
<th>Device with Asymmetric Enclosure (25mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommended sustained power (W)</td>
<td>12</td>
<td>16</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>Touch point Temperature limit (°C)</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Enclosure Max Inlet air temperature, &lt; 950 m (°C)</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>Enclosure Max Inlet air temperature, 950 m to 3050 m (°C)</td>
<td>35 - (1°C for 175 m of elevation gain)</td>
<td>35 - (1°C for 175 m of elevation gain)</td>
<td>35 - (1°C for 175 m of elevation gain)</td>
<td>35 - (1°C for 175 m of elevation gain)</td>
</tr>
<tr>
<td>Add in card to add in card pitch (mm)</td>
<td>9</td>
<td>11</td>
<td>13</td>
<td>26</td>
</tr>
<tr>
<td>Recommended Fan Pressure loss across device (Pascal)</td>
<td>83</td>
<td>52</td>
<td>64</td>
<td>21</td>
</tr>
<tr>
<td>Airflow, average min per device (CFM), 1 CFM = 1.7 m³/h</td>
<td>1.41 - (0.01 CFM for every 1°C below 35°C inlet temp)</td>
<td>1.71 - (0.06 CFM for every 1°C below 35°C inlet temp)</td>
<td>2.02 - (0.02 CFM for every 1°C below 35°C inlet temp)</td>
<td>4.10 - (0.04 CFM for every 1°C below 35°C inlet temp)</td>
</tr>
</tbody>
</table>