What a mess!

We can do better
What do we care about in a form factor:

• Physically large enough to meet required capacity
• Thermally capable to cool the device with desired airflow
• Serviceability – easy to install and replace

Lessons learned:
• m.2 is challenging for thermals, serviceability, and capacity
• Messy carrier cards and thermal interface pads

Looking forward:
• SSD market is fragmenting into multiple formfactors, that have divergent design-in requirements
• E1.S (multiple widths), E1.L (multiple widths), E3 (multiple lengths and widths), U.2
• On the bright side, convergence on connectors!

Can we find commonality and reduce the proliferation in formfactors?
**E1 vs E3**

- **E1 can fit in 1U or 2U platforms**
  - Use orthogonal connector to avoid backplane
  - Optimized for 1U platforms
- **E3 can also fit in 1/2U**
  - Requires a backplane in a 1U
  - Optimized for 2U platforms

---

**E1**  
- 7.5 mm

**E3**  
- 7.5 mm

**E1.S**  
- 9.5 mm

**E3**  
- 7.5 mm

---

**OCP NIC 3.0**

---

**E1**  
- 7.5 mm

**E3**  
- 7.5 mm

---

**OCP NIC 3.0**

---

**E1**  
- 7.5 mm

**E3**  
- 7.5 mm

---

**OCP NIC 3.0**

---

**E1**  
- 1.5 mm

**E1**  
- 1.5 mm

---

**OCP NIC 3.0**

---

**E1**  
- 1.5 mm

**E1**  
- 1.5 mm

---

**OCP NIC 3.0**
Microsoft Azure’s opinion

E1.S for performance and moderate capacity, E1.L for high capacity

E1.L with only 2 variants meets our needs

E1.S on paper looks good, but we have challenges
- 9.5mm isn’t thermally capable enough to meet SSD performance requirements
- 25mm is too large to optimize front of server real estate
- Why not create a mid-size option that works for us, and others

Proposal
- What is the minimum width of E1.S that can handle 20-25W @ 35C ambient conditions?
- Enables highest front panel density
- Define a double width version for higher power needs
- Need to scope SCM
Decoding terminology & Simulation Strategy

What we did
We kept a consistent gap of 2mm
Simulated several widths from 9.5 to 15mm

What is next?
Need to simulate with multiple vendors inputs
Ensure we simulate future SSD needs (8, 16TB+)
Storage class memory

<table>
<thead>
<tr>
<th>Device Width (SSD)</th>
<th>9.5</th>
<th>11.5</th>
<th>12</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gap between SSDs</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Effective pitch</td>
<td>11.5</td>
<td>13.5</td>
<td>14</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>Double width device</td>
<td>21</td>
<td>25</td>
<td>26</td>
<td>30</td>
<td>32</td>
</tr>
</tbody>
</table>
Building on E1

A chassis supporting E1 is simply a PCIe device container
Can we leverage this work for other applications?
Computational Storage?
Accelerators?

For accelerators, I propose we find what works with E1.S widths, and only grow the height....
But let’s focus on the SSDs first, then figure out accelerators
Thermal Analysis Boundaries

Spatial boundary conditions

- Three 25mm SSD’s with 27mm pitch set 83mm simulation domain.
- Mechanical assumption is that thinner SSD’s will be placed in the domain if there’s room between the domain wall and the drive itself.
  - Even if the drive has less of an air gap between the fin/bottom of the drive and the domain walls.

Airflow Constraints

- Inlet airflow is held constant in simulation.
- Airflow velocity approach was to determine least amount of airflow required to hit the NAND composite temperature for the center drives powered to 20W workload target.
  - Next round of simulations will determine maximum airflow required to cool outer SSD’s
Thermal Simulation Results

**Assumptions**

- SSD Power is scaled by percent t-rise for the NAND.
  - Reported SSD power is normalized to the NAND composite’s temperature spec. The reported total device power is corrected as if it were operating at the NAND component’s temperature spec.
    - Example: If result is above temperature spec, power is scaled down. If result is below temperature spec, power is scaled up.
- NAND temperature set’s power limits due to it having the least temperature margin.
- Inlet temperature (35C), inlet airflow velocity, and domain area (83 x 33.8 mm^2) are held constant.

<table>
<thead>
<tr>
<th>Component</th>
<th>Component Temperature limit</th>
<th>SSD Module Power (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SoC</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>NAND</td>
<td>80</td>
<td>~19.88</td>
</tr>
<tr>
<td>DDR</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>
• 15mm SSD’s result in highest total domain power for simulations to date.
  • 15mm also meets 30W+ target.
• 25mm SSD has high power density, but lower overall domain power capacity.
  • 25mm is 12W higher in power density than 15mm, but 28W less in overall total domain SSD power.

Simulation notes: Airflow held constant. Input: SSD width; Output: composite temperatures.
Flat Plate SSD Pitch Study

11.5mm (2mm gap) vs 13.5mm (4mm gap) Pitch Study for Flat Plate SSD's

- 9.5mm SSD width held constant.
  - Varied pitch in analysis.
- Pitch has little cooling improvement due to flat plate design.
• Inlet airflow has been adjusted by 25% to address concerns about optimal system airflow or increased system impedance.
• Chose 15mm based on total domain power and ability to meet 30W thermal design power target.
Overall Thermal Learnings

• From 9.5mm to 25mm there is a point of diminishing returns for overall domain power utilization.

• Required airflow will vary between width’s due to amount of energy needing to be cooled in a constrained domain.

• Discussions around target per drive LFM have been 70-100 LFM. Simulation shows an average of 73 LFM per drive when considering both individual SSD TDP and total domain power within mechanical constraints.