Datacenter Designs Using the EDSFF Form Factors

Mark A. Shaw

Microsoft Azure
EDSFF – Decoupling for workloads

**SFF-TA-1006 (E1.S)**
Performance Density Optimized Services
High Power SCM (i.e. PCM/ReRAM/FRAM)
Cache Coherent Devices (i.e. Gen-Z)
Targets Compute nodes

**SFF-TA-1008 (E3)**
Ultra High-Performance Applications
FPGA or Computational Accelerations
AIC / PCIe HHHL replacement

**SFF-TA-1007 (E1.L)**
Capacity Density Optimized Services
Cheapest Consumer memories (i.e. QLC)
Denali-based pools of flash
Targets Storage
Azure Design Considerations

**Compute**
- Increasing # of VMs per node
- Challenges keeping up with IOP Density
- Varying IOP and Density requirements
- Power, Space and Thermal constrained

**Performance Storage**
- High Density
- Low Latency
- Low $/GB
- On-Line Service

**Cold Storage**
- High Density
- Lowest $/GB
- Low $/Watt
- On-Line Service

**GPU**
- Ultra High-Performance Applications
- FPGA or Computational Accelerations
- PCIe HHHL replacement
Performance Storage – FX16

- Lowest $/TB
- High Density (256 TB or higher)
- Low latency
- Moderate BW/SSD
- Expandable capacity
- On-line Service
- Modular to support all head nodes
- Leverage existing components and architectures
M.2 Carrier in FX16

Advantages
- TTM with existing tech
- Meets current performance requirements

Disadvantages
- Mechanical complexity
- Thermal performance
- 3.3V power scaling
- Supply Chain complexity
- Service complexity
- Increased “spares” support
- Lack of QLC support ($$)
Advantages
- Meets performance requirements
- Reduced mechanical complexity
- Thermal Performance
- Scaleability
- Reduce Supply Chain complexity
- Reduced service complexity
- QLC support $$

Disadvantages
- TTM with new technology
- Risk with new technology
Cloud SSD E1.L

MULTI-SOURCED!!

- E1.L Form Factor
- Configurable platform for multiple work loads
- Single or Dual controller configurable
- Supports LL-NAND, eTLC, QLC
- Supports LBA and Denali Media SSD
## Going Forward

<table>
<thead>
<tr>
<th>M.2</th>
<th>PCIe Gen3</th>
<th>PCIe Gen4</th>
<th>PCIe Gen5</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
</tbody>
</table>

**PCIE Gen3**
- Degraded connector performance
- Carrier topology risks
- Bent but not broken

**PCIE Gen4**
- Degraded connector performance
- Reduced trace lengths
- Riser topology risks
- Increased board material cost
- Bent but not broken

**PCIE Gen5**
- Connector rated for Gen5
- Demonstrated 2 connector topology
- High risk connector performance
- Higher risk carrier topologies
- Likely broken

- Degraded connector performance
- Reduced trace lengths
- Riser topology risks
- Increased board material cost
- Likely broken

- Connector rated for Gen5
- Demonstrated 2 connector topology
Thank You