Open. Together.
Learning of Designing Project
Olympus JBOF with EDSFF SSD

Mark A Shaw
Principal Hardware Engineer Manager, Microsoft

Antson Chang
Assistant Technical Director, Wiwynn Corporation
Agenda

- Project Olympus FX-16 JBOF
- Design tips
- Power consumption and thermal efficiency
- Conclusion
### Summary of FX-16

#### Device
- 1U JBOF
- Total capacity is 256TB (or higher)

#### I/O
- 2 PCIe x16 uplinks in mini-SAS HD form factor running at PCIe gen3
- 1 Ethernet port for management

#### SSD slot
- 16 slots and per slot supports up to x8
- EDSFF 1U Long (E1.L), 18mm width

#### SSD
- Hot-pluggable and front loaded
- Multi-sourced

#### Power consumption
- <400 watts
### Storage Trends

**Performance**

<table>
<thead>
<tr>
<th>Type</th>
<th>Interface</th>
<th>Read (IOPs)</th>
<th>Write (IOPs)</th>
<th>Read (GB/s)</th>
<th>Write (GB/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDD</td>
<td>SAS/SATA</td>
<td>200</td>
<td>200</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>SSD</td>
<td>SATA</td>
<td>95K</td>
<td>25K</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>SAS</td>
<td>420K</td>
<td>50K</td>
<td>2.2</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>PCIe</td>
<td>750K</td>
<td>180K</td>
<td>3.5</td>
<td>3.0</td>
</tr>
</tbody>
</table>

**Density**

<table>
<thead>
<tr>
<th>Year</th>
<th>Density per 1U (Tera-Byte)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HDD</td>
</tr>
<tr>
<td>2018</td>
<td>210-350</td>
</tr>
<tr>
<td>2019</td>
<td>245-395</td>
</tr>
<tr>
<td>2020</td>
<td>280-450</td>
</tr>
</tbody>
</table>

**Cost**

<table>
<thead>
<tr>
<th>Media Price ($/GB)</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business critical HDD</td>
<td>$0.026</td>
<td>$0.023</td>
<td>$0.019</td>
<td>$0.017</td>
</tr>
<tr>
<td>Mission critical HDD</td>
<td>$0.117</td>
<td>$0.105</td>
<td>$0.093</td>
<td>$0.087</td>
</tr>
<tr>
<td>Data center SSD</td>
<td>$0.45</td>
<td>$0.26</td>
<td>$0.17</td>
<td>$0.13</td>
</tr>
</tbody>
</table>

Source: IBM Mural Lyer's presentation slide: NVMe Overview
Trade-off for what Data Center really wants!

- System and SSD features we want......
  - Hot-pluggable
  - Balanced airflow impedance and thermal efficiency
  - Total capacity
  - Flexibility for high performance/density SSDs
  ......etc.

- Put several M.2 together as SSD group?
- U.2 => M.2 carrier => EDSFF SSD
New SSD form factor: EDSFF

- *In the end of 2016*, Wiwynn worked with Microsoft on *M.2 Carrier*. In the same time, Samsung (M.3/NF1/NGSFF) and Intel (Ruler SSD) proposed similar form factors.
- *In the first half of 2017*, Intel invites Microsoft, Lenovo, Facebook, Dell/EMC, Samsung, HPE …etc. to join *EDSFF* (Enterprise & Datacenter SSD Form Factor)
- *In 2018 Q1*, Olympus FX-16 first proto-type supporting 16 E.1L 18mm SSDs was demonstrated at OCP Summit.
- *In 2019 Q2*, FX-16 is *ready for mass production*. 
### Dimension defined in EDSFF

<table>
<thead>
<tr>
<th>SSD Form factor</th>
<th>Dimension (mm)</th>
<th>Maximum Power (watt) (Recommended)</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Height</td>
<td>Width</td>
<td>Depth</td>
</tr>
<tr>
<td>EDSFF, 1U (x4/x8)</td>
<td>Short</td>
<td>31.5</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td>Short</td>
<td>31.5</td>
<td>8.01</td>
</tr>
<tr>
<td></td>
<td>Long</td>
<td>38.4</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td>Long</td>
<td>38.4</td>
<td>18</td>
</tr>
<tr>
<td>EDSFF, 3 inch (x4/x8/x16)</td>
<td>Short</td>
<td>76</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>Short</td>
<td>76</td>
<td>16.8</td>
</tr>
<tr>
<td></td>
<td>Long</td>
<td>76</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>Long</td>
<td>76</td>
<td>16.8</td>
</tr>
<tr>
<td>Other existing SSD form factor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.2 (2.5”)</td>
<td>70</td>
<td>7/15</td>
<td>100</td>
</tr>
<tr>
<td>M.2</td>
<td>22</td>
<td>4.3</td>
<td>110</td>
</tr>
<tr>
<td>NGSFF (M.3)</td>
<td>30.5</td>
<td>4.3/4.8</td>
<td>110</td>
</tr>
</tbody>
</table>

**EDSFF form factor**

**U.2 form factor**

**M.2 or NGSFF form factor**
Agenda

• Project Olympus JBOF FX-16

• Design tips
  • Power consumption and thermal efficiency
  • Conclusion
## System Implementation with EDSFF

<table>
<thead>
<tr>
<th>Backplane</th>
<th>Paddle Card</th>
<th>Orthogonal conn. (Press fit)</th>
<th>Orthogonal conn. (SMT with assembling)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Backplane" /></td>
<td><img src="image" alt="Paddle Card" /></td>
<td><img src="image" alt="Orthogonal conn. (Press fit)" /></td>
<td><img src="image" alt="Orthogonal conn. (SMT with assembling)" /></td>
</tr>
</tbody>
</table>

**Pros.**
- Support up to PCIe x16 in 2U chassis.
- Clear airway enhances thermal efficiency.
- SMT connectors. Easier placement and SI routing
- PCIe Gen4/5 support
- Connector availability & sourcing
- Clear airway enhances thermal efficiency.
- Less complicated mechanical design
- Easier assembly process
- Clear airway enhances thermal efficiency.
- SMT connectors. Easier placement and SI routing
- PCIe Gen4/5 support
- Screw fixing makes system mechanical design easier.

**Cons.**
- More restricted airflow in 1U.
- To connect with MB, several high-speed connectors or cables are needed.
- More complicated mechanical design
- More complicated assembly process
- New (2018 Q4).
- Hard to place components next to connector.
- Back-drill is needed to support up to PCIe Gen4/5.
- Connector Sourcing
- New (2019 Q2).
Mechanical Design with EDSFF

- E1.L supports customized extension kit to aligned with existing panel looking.
- For blind mating
  - Guiding rails
  - Mechanical stopper
  - Latch
- Shelter blocks airflow while SSD is not populated.
Mechanical Design with EDSFF

- E1.L supports customized extension kit to aligned with existing panel looking.

- For blind mating
  - Guiding rails
  - Mechanical stopper
  - Latch

- Shelter blocks airflow while SSD is not populated.
Mechanical Design with EDSFF

• E1.L supports customized extension kit to aligned with existing panel looking.

• For blind mating
  - Guiding rails
  - Mechanical stopper
  - Latch

• Shelter blocks airflow while SSD is not populated.
Mechanical Design with EDSFF

- E1.L supports customized extension kit to aligned with existing panel looking.
- For blind mating
  - Guiding rails
  - Mechanical stopper
  - Latch
- Shelter blocks airflow while SSD is non-populated.
# Electrical Design for hot-plugging

<table>
<thead>
<tr>
<th>Function</th>
<th>Circuitry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isolation between slots</td>
<td>Suggest 100% power and signal isolation between slots to ensure that there is no power or signal interaction between slots. Insertion or removal of one slot cannot affect the functionality of other slots.</td>
</tr>
<tr>
<td>Power enabling</td>
<td>MOS + soft start (or hot-swap controller)</td>
</tr>
<tr>
<td>Inrush current protection</td>
<td>Extra logics for power enabling based on presence pin Note: There is no pre-charge pin in EDSFF pin definition.</td>
</tr>
<tr>
<td>Over Current Protection</td>
<td>Polyfuse (or hot-swap controller)</td>
</tr>
</tbody>
</table>
How much time is needed for SMbus access after plugging?

- ~10msec.
- ~1000msec.
- >100msec.
- Many seconds
Dual port implementation

- **Clock and PERST# input** for Port B are needed.
- Pull-down DualPortEN# (B9 pin) by host **before enabling SSD 12V power.**
- Make sure **which lanes** connect to Port B.
Agenda

- Olympus JBOF FX-16
- Design tips
- **Power consumption and thermal efficiency**
- Conclusion
## Power and Thermal

<table>
<thead>
<tr>
<th>Stress tool (Diskspd)</th>
<th>Block size (Byte)</th>
<th>Ambient 25°C</th>
<th>Ambient 35°C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>System fans (6pcs)</td>
<td>System fans (6pcs)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PWM / Air flow / Power</td>
<td>PWM / Air flow / Power</td>
</tr>
<tr>
<td></td>
<td></td>
<td>System power (watts)</td>
<td>System power (watts)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CFM * kW</td>
<td>Fan power / System Power</td>
</tr>
<tr>
<td>100% Sequential Write</td>
<td>512K</td>
<td>36% / 23.4 CFM / 6.50 W</td>
<td>341</td>
</tr>
<tr>
<td></td>
<td>4K</td>
<td>26% / 16.3 CFM / 4.00 W</td>
<td>249</td>
</tr>
<tr>
<td>50% Sequential Write/Read</td>
<td>512K</td>
<td>34% / 22.0 CFM / 5.75 W</td>
<td>315</td>
</tr>
<tr>
<td></td>
<td>4K</td>
<td>24% / 14.9 CFM / 3.75 W</td>
<td>212</td>
</tr>
<tr>
<td>100% Sequential Read</td>
<td>512K</td>
<td>22% / 13.6 CFM / 3.50 W</td>
<td>193</td>
</tr>
<tr>
<td></td>
<td>4K</td>
<td>22% / 13.6 CFM / 3.50 W</td>
<td>162</td>
</tr>
</tbody>
</table>

*: Microsoft specification is less than 158 CFM/kW

*=Microsoft specification is less than 158 CFM/kW
Agenda

• Olympus JBOF FX-16
• Design tips
• Power consumption and thermal efficiency

• Conclusion
Conclusion

• EDSFF provides:
  - front access and hot pluggable capability,
  - high density system design
  - efficient thermal design
  - Scalable options for x4, x8 or x16
  - Single or dual port feature
  - Same connector definition supporting PCIe Gen3/4/5

Join us to use EDSFF for the next decade for SSDs!
Agenda

• Project Olympus FX-16 JBOF
  - Summary of FX-16,
  - Storage Trends,
  - Trade-off for what Data Center really wants!
  - New SSD form factor: EDSFF, and it’s dimension definition

• Design tips
• Power consumption and thermal efficiency
• Conclusion
Agenda

- Project Olympus JBOF FX-16

- Design tips
  - System implementation with EDSFF,
  - Mechanical and Electrical design,
  - SMbus access,
  - Dual-port implementation

- Power consumption and thermal efficiency

- Conclusion