OCP NIC 3.0 Collaboration

- An Open Hardware development Story

Joshua Held / Mechanical Engineer
Yueming Li / Thermal Engineer
Jia Ning / Hardware Engineering Manager
Facebook, Inc
Agenda

• Overview of project in the past one year
• OCP NIC 3.0 Mechanicals
• OCP NIC 3.0 Thermal
Background

OCP Mezzanine Cards

<table>
<thead>
<tr>
<th>Specification</th>
<th>Version</th>
<th>Submit Date</th>
<th>Contributor</th>
<th>License</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCP Mezzanine Card v2.0</td>
<td>V2.0-1.0</td>
<td>Dec 15, 2015</td>
<td>Facebook</td>
<td>OWFa 1.0</td>
<td>Added support for x16 (quad x4), NCISI, dual QSFP+, &amp; Quad SFP+</td>
</tr>
<tr>
<td>OCP_Mezz_2.0_rev1.00_20151215b_pub_release.pdf (2.2MB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accepted by OCP IC 2/24/2016</td>
</tr>
<tr>
<td>OCP_Mezz_2.0_rev1.00_20151215b_pub_release_3D_package.zip (88MB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mechanical 20151023_P1-P9_K1-K5.zip (57MB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCP Mezzanine Card v0.5, original standard</td>
<td>V0.5</td>
<td>Oct 8, 2012</td>
<td>Facebook</td>
<td>OWFa 1.0</td>
<td>Defacto standard for the original network mezzanine with a x8 PCIe Gen3 interface</td>
</tr>
<tr>
<td>Mezzanine Card (rev 0.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OCP Mezz v0.5
- Defined ~2012
- 10G Ethernet
- 2x SFP
- X8 PCIe Gen3
- I2C sideband

OCP Mezz v2.0
- Defined ~2015
- 10/25/40/50/100G Ethernet
- Up to 4x SFP28, 2x QSFP28, 4x RJ45
- X16 PCIe Gen3
- NCISI sideband
OCP NIC 2.0 Limitation

- Gates emerging use cases & blocks broader adoption
  - Board space
  - Mechanical and thermal profile
  - Connector placement
  - Specification quality
OCP NIC 3.0 Milestones

2017

Preparation

1

Setup Subgroup
Mailing list
Conference calls
Wiki

Define problem statement
And design boundary

March Summit Workshop

2

Solution exploration
and convergence

Dallas Workshop

3

Specification drafting,
review, and approval

4

Implementation and
Refinement

March Summit Workshop
**Define Problem Statement**

March to Mid-April ‘17

---

<table>
<thead>
<tr>
<th>OCP Partner</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td># I/O Ports</td>
<td>2, (2xSFP/SFP28, 2xQSFP/QSFP28)</td>
<td>2, (2xSFP/SFP28, 2xQSFP/QSFP28)</td>
<td>1-4 ports SFP+ or Base-T, or 1-2 QSFP</td>
<td>1-4 ports SFP+ or Base-T, or 1-2 QSFP</td>
</tr>
<tr>
<td># Major IC (ASIC, FPGA, or other)</td>
<td>1 or 2 (ASIC, ASIC+PHY)</td>
<td>2 or 3 (ASIC + FPGA/SoC + PHY)</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td># DRAM</td>
<td>6 x16</td>
<td>12 x16</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Power envelope of IC#1 (Max power at Tj max)</td>
<td>15</td>
<td>20</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Max T_case of IC#1</td>
<td>95</td>
<td>95</td>
<td>105</td>
<td>110</td>
</tr>
<tr>
<td>IC#1 mechanical dimension (WxLxH)</td>
<td>33x33x2.5</td>
<td>45x45x3.7</td>
<td>25x25x 3.7</td>
<td>45x45x3.7</td>
</tr>
<tr>
<td>Power envelope of IC#2 (Max power at Tj max)</td>
<td>N/A</td>
<td>3W per port</td>
<td>3W per port, total &lt;50W</td>
<td>N/A</td>
</tr>
<tr>
<td>Max T_case of IC#2</td>
<td>N/A</td>
<td>105</td>
<td>110</td>
<td>110</td>
</tr>
<tr>
<td>IC#2 mechanical dimension (WxLxH)</td>
<td>N/A</td>
<td>19x19x~2.5 mm</td>
<td>19x19~2.5 mm</td>
<td>25 x 25</td>
</tr>
<tr>
<td>Power envelope of IC#2 (Max power at Tj max)</td>
<td>N/A</td>
<td>TBD, total &lt; 50W</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Max T_case of IC#2</td>
<td>N/A</td>
<td>105</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>IC#2 mechanical dimension (WxLxH)</td>
<td>N/A</td>
<td>N/A</td>
<td>25x25x3.7</td>
<td>N/A</td>
</tr>
<tr>
<td>DRAM Power (each component at Tj max)</td>
<td>0.33</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Max T_case of DRAM</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td>Optical Module power (each)</td>
<td>0</td>
<td>1.5</td>
<td>1.5 watts (SFP+), 3.5 watts (QSFP)</td>
<td>1.5 watts (SFP+), 3.5 watts (QSFP)</td>
</tr>
<tr>
<td>Optical Module Max Tcase</td>
<td>0</td>
<td>70</td>
<td>70 deg C</td>
<td>85 deg C</td>
</tr>
</tbody>
</table>

---

*Note: Graph shows trends for 90/10 and 50/50 goals for TCASE-TDP.*
Solution Exploration and Convergence

Mid-April to end-of-September’17

14x solution options proposed and evaluated
Dallas Workshop

Sep 25th, 2017

• First F2F workshop
  • Met each other in person
  • Had an open discussion
  • Accelerated the collaboration
Solution Overview

- 2x Form factors (SFF and LFF)
- SFF-TA-1002 connector
- 32 lanes of PCIe Gen4
  - 4x of OCP NIC 2.0
- EMI containment
- Front service
- 80W/150W power delivery
- Larger thermal potential in similar profile
- NIC management features

Draft – Review - Approval

Oct’17 to Jan’18

15x General specification working sessions
11x Mechanical specific working sessions
6x Thermal specific working sessions
41x doc revisions

1 Specification
Mechanical 3D models
Mechanical 2D models
Thermal simulation models
Thermal test fixture model

Community partners
Amphenol Corporation
Broadcom Limited
Dell, Inc.
Facebook, Inc.
Hewlett Packard Enterprise Company
Intel Corporation
Lenovo Group Ltd
Mellanox Technologies, Ltd
Netronome Systems, Inc.
Quanta Computer Inc.
TE Connectivity Corporation
And many more!
Implementation and Refinement

0v70 – Initial release
Jan 25th, 2018

0v80 – Hot fixes
End of March’18

0v90 – Add Signal Integrity Guideline and Conformance
TBD

1v00 – Add Implementation Learning
Q4’18

Subgroup Wiki with latest specification: http://www.opencompute.org/wiki/Server/Mezz
Mailing list: http://lists.opencompute.org/mailman/listinfo/opencompute-mezz-card
OCP NIC 3.0 Mechanicals
### OCP Mezz 2.0 vs OCP NIC 3.0

<table>
<thead>
<tr>
<th></th>
<th>Mezz 2.0</th>
<th>NIC 3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Size</td>
<td>Non-Rectangle</td>
<td>76x115</td>
</tr>
<tr>
<td>Small Area</td>
<td>8000 mm²</td>
<td>8740 mm²</td>
</tr>
<tr>
<td>Large Size</td>
<td>NA</td>
<td>139x115</td>
</tr>
<tr>
<td>Large Area</td>
<td>NA</td>
<td>15985 mm²</td>
</tr>
<tr>
<td>Expansion Direction</td>
<td>NA</td>
<td>Side</td>
</tr>
<tr>
<td>Connector style</td>
<td>Mezz</td>
<td>Edge (.6mm pitch)</td>
</tr>
<tr>
<td>PCB Orientations</td>
<td>Parallel</td>
<td>Parallel</td>
</tr>
<tr>
<td>Installation</td>
<td>In Chassis</td>
<td>Front/Rear Panel</td>
</tr>
<tr>
<td>Installation Action</td>
<td>Parallel to Front/Rear Panel</td>
<td>Perpendicular to Front/Rear Panel</td>
</tr>
<tr>
<td>Hot Swap</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>EMI Containment for Serviceability</td>
<td>High Difficulty</td>
<td>Low Difficulty</td>
</tr>
</tbody>
</table>
NIC 3.0 Configuration
NIC 3.0 Mechanical Goal:

Develop universal form factors which shall include mechanicals and EMI containment.

Requirements:

• The same NIC design must work in both straddle and right angle configurations
• NICs might be oriented horizontally or vertically
• Retention and guidance must work with specified NIC v3.0 MB thicknesses
• Common mechanical features used across small and large form factors
• Some form of mechanism is required for seating large form factors due to mating forces
• Design of PCB should be flexible enough to support many component and connector configurations without need of mechanical changes
• Recommended mechanical designs will be included in the specification to simplify and reduce barriers to adoption
NIC 3.0 Module Configuration

Features:

• Increased total PCB space
• Simplified component keep-in areas
• Scalable design to support large form factor
• Built in EMI containment
• Available in thumb screw or tool-less configurations with no PCB changes
**NIC 3.0 Module Sizes**

2.75mm card guide keep out

<table>
<thead>
<tr>
<th>PCB Size</th>
<th>Available Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top/Bottom</td>
<td>Routing Inner Layers</td>
</tr>
<tr>
<td>Component Placement</td>
<td></td>
</tr>
<tr>
<td>Placement</td>
<td></td>
</tr>
<tr>
<td>Routing Inner</td>
<td></td>
</tr>
<tr>
<td>Layers</td>
<td></td>
</tr>
<tr>
<td>Small Size (SFF)</td>
<td>76x115</td>
</tr>
<tr>
<td></td>
<td>70.5x109</td>
</tr>
<tr>
<td></td>
<td>76x115</td>
</tr>
<tr>
<td>Large Size (LFF)</td>
<td>139x115</td>
</tr>
<tr>
<td></td>
<td>133.5x109</td>
</tr>
<tr>
<td></td>
<td>139x115</td>
</tr>
</tbody>
</table>

subtract 5.5mm from width (card guide)
subtract 6mm from length (edge connector)
NIC 3.0 SFF Module Versions

2x QSFP

4x SFP

4x RJ45

Single thumb screw version

Tool-less version

Complete 3D CAD available at:
http://www.opencompute.org/wiki/Server/Mezz
NIC 3.0 LFF Module Versions

2x QSFP

4x SFP

4x RJ45

Complete 3D CAD available at: http://www.opencompute.org/wiki/Server/Mezz
NIC 3.0 Chassis Examples

Straddle Mount

Right Angle Mount

Complete 3D CAD available at: http://www.opencompute.org/wiki/Server/Mezz
OCP NIC 3.0 Thermal
Thermal Benefits for NIC 3.0

- More space for heatsink with no bergstak connector on the side
- Up-facing heatsink permits flexibility on heatsink height
- LFF allows thermal potential for high-power ASIC cooling
Thermal Design Guidance – Cold Aisle

- For typical inlet temperature 35°C, SFF:
  - Support 15W ASICs under 100 LFM (Typical)
  - Support 23W ASICs under 200 LFM (High)
• Hot-aisle cooling is more challenging due to higher supply air temperature

• For typical inlet temperature 55°C, SFF:
  - Support 15W ASICs under 200 LFM (Typical)
  - Support 20W ASICs under 300 LFM (High)

Complete CFD model available at: http://www.opencompute.org/wiki/Server/Mezz
Thermal Test Fixture

Purpose:
- Provide standardized test data across different NIC and system vendors

Features:
- Simple and easy adoption by both NIC and system vendors
- Representative thermal data to define cooling tiers across different use cases
- Functional test board for power delivery and reporting interface
Implementation and Refinement

0v70 – Initial release
Jan 25th, 2018

0v80 – Hot fixes (Minor mechanical spec updates)
End of Feb’18

0v90 – Add signal integrity guideline and conformance
TBD

1v00 – Add Implementation learning (planned mechanical expansion)
Q4’18

Subgroup Wiki with latest specification: http://www.opencompute.org/wiki/Server/Mezz
Mailing list: http://lists.opencompute.org/mailman/listinfo/opencompute-mezz-card