OPEN

Compute Project

OCP Mezzanine card 2.0 Design Specification

~~Draft~~ **Version 0.~~33~~40**

Candidate for Release

~~Subject to Change~~

Author: Jia Ning, Engineer, Facebook

1 Contents

2 Overview

2.1 License

2.2 Background

The original OCP Mezzanine Card for Intel v2.0 Motherboard specification[1] have been developed mainly to serve the use case of Single and Dual port 10G Ethernet card. Adoption of this specification has been seen in OCP community on different server and storage platforms. Over the recent two years, demand of supporting new use cases were raised and the original Mezzanine card specification cannot support those new use cases without modification in order to support different I/O types, increase bandwidth of data and management, and support higher power controller IC.

Mezzanine card 2.0 specification is developed based on original OCP Mezzanine card. It extends the card mechanical and electrical interface to enable new uses cases for Facebook and other users in OCP community. The extension takes backward compatibility to existing OCP platforms into consideration, and some tradeoffs are made between backward compatibility and new requirements.

2.3 New Use Cases

These new major use cases are taken into consideration in this specification.

- Single and Dual QSFP port 40G Ethernet NIC
- Quad SFP+ Port 10GE NIC
- Quad port 10GBase-T NIC
- x16 PCIe lane to baseboard

[1] http://www.opencompute.org/assets/download/Intel-Mezzanine-Card-Design-Specification-v0.5.pdf

- 16x KR to baseboard
- NIC controller with high TDP that needs more heatsink volume
- Management sideband to support use case such as remote System Firmware update
- Baseboard and Mezzanine card identification

The Mezzanine card 2.0 specification makes change on as needed base to maximize backward compatibility to existing OCP platforms. Some modification impacts backward compatibility to existing OCP platforms and compatibility check need to be done.

2.4 Major Changes to Form Factor

To accommodate the new uses cases above, major changes to form factor are listed as below. More detailed description can be found in Chapter 3.

- Extend PCB area to support 1x extra connector to baseboard
- Extend PCB area to support I/O interface
- Add option to have I/O on Secondary side to support I/O interface
- Add 12mm stacking option to support higher volume heatsink

2.5 Major Changes to Electrical Interface

To accommodate the new uses cases above, major changes to electrical interface are listed as below. More detailed description can be found in chapter 4 ~~chapter 4.~~.

- Modify original 120pin connector to have NC-SI signals; this is the original OCP Mezzanine card connector; it is referred to Connector A in this specification.
- Add 80 pin connector on Mezzanine card in order to expend PCIe lane width from x8 to x16; it is referred to Connector B in this specification.
- Add card ID mechanism for baseboard to identify different types of Mezzanine cards
- Add definition of thermal reporting interface to support temperature based system fan speed control
- Add NC-SI interface to support higher bandwidth management traffic

3 Mezzanine Card Form Factor

Mezzanine card form factor is described in this chapter. Vendor should refer to 2D DXF and 3D files for dimension, tolerance, and height restriction details.

3.1 Primary and Secondary Side Definition

Primary side and secondary side are used to refer to the two sides of mezzanine card in this document. Primary side is the side with Mezzanine board to board connector. Example of primary side and secondary side is shown in Figure 1 ~~Figure 1~~.

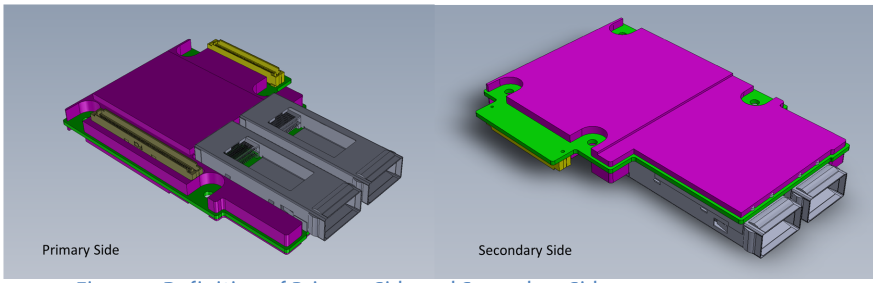


Figure 1: Definition of Primary Side and Secondary Side

3.2 Mezzanine Card Connectors

For connector A, FCI/61083-124402LF or equivalent connector is mounted on Mezzanine card. FCI/61082-121402LF (or equivalent) for 8mm stack height, or FCI/61082-122402LF (or equivalent) for 12mm stack height is mounted on baseboard side. PCI-E x8 Gen3, I2C and NC-SI sideband signals, and power are assigned in connector A. Connector A can also be used for up to 8x KR.

For Connector B, FCI/61083-084402LF or equivalent connector is mounted on Mezzanine card. FCI/61082-081402LF (or equivalent) for 8mm stack height, or FCI/61082_082402LF (or equivalent) for 12mm stack height is mounted on baseboard side. PCI-E x8 Gen3, which can be combined to x16 with Connector A are provisioned in Connector B. Connector B can also be used for up to 8x KR.

	Mezzanine card	Baseboard (8mm stack)	Baseboard (12mm stack)
Connector A	FCI/61083-124402LF	FCI/61082-121402LF	FCI/61082-122402LF
Connector B	FCI/61083-084402LF	FCI/61082-081402LF	FCI/61082-082402LF

Connector A and Connector B is shown in Figure 2~~Figure 2~~, viewing from secondary side.

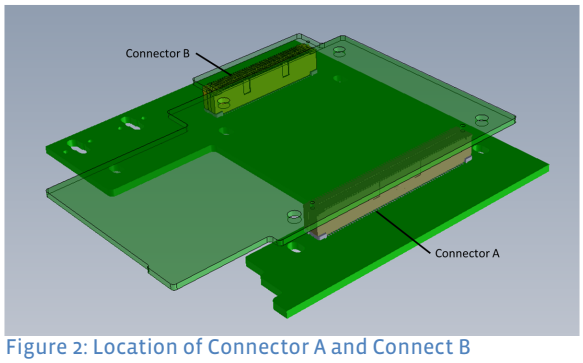


Figure 2: Location of Connector A and Connect B

3.3 Form Factor Definition in Horizontal Plane

This section defines the Mezzanine card form factor in horizontal direction, i.e. from top or bottom view.

There are two optional PCB areas, and the usage depends on the connection needed for host interface side, and I/O. Figure 3~~Figure 3~~ illustrates Mezzanine PCB from primary side with two optional PCB area shown.
By default vendor should implement the Mezzanine card in form factor in Figure 4~~Figure 4~~. Vendor should extend PCB to the optional areas if and only if the extension is necessary to achieve the purposes mentioned below in this section. It is to maximize mechanical compatibility to existing platforms.

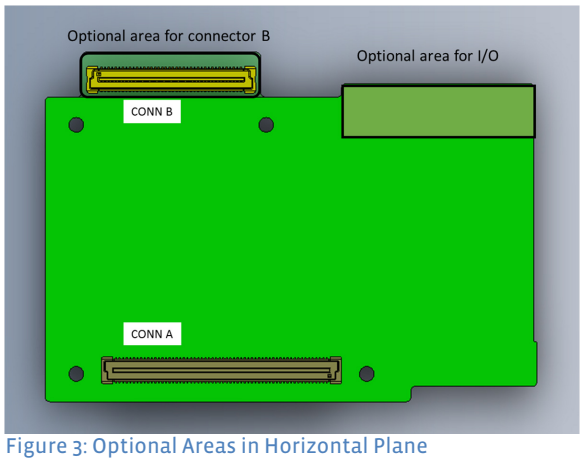


Figure 3: Optional Areas in Horizontal Plane

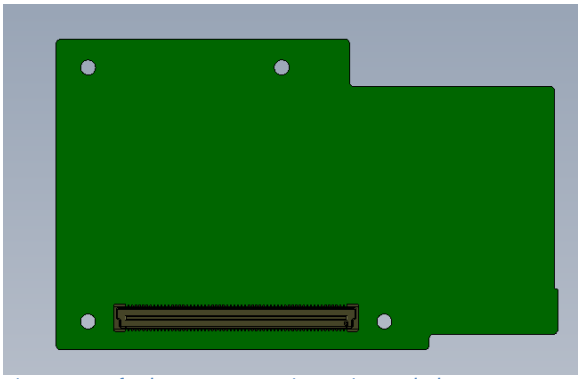


Figure 4: Default Form Factor in Horizontal Plane

3.3.1 Optional Area for Connector B

This area is extended to increase channels or lanes to baseboard.

Connector B is an 80 pin connector (FCI/61083-084402) and provides extra x8 PCIe lanes (or 8x KR for ~~PHY~~KR Mezzanine card), PERST# signals and clocks. Definition of the Connector B is in Chapter 04.

Mezzanine card that only uses signals in Connector A should not extend PCB to this area.

Mezzanine card that uses more than x8 PCIe lane should extend PCB to this area.

3.3.2 Optional Area for I/O

This area is extended to accommodate more I/O types.

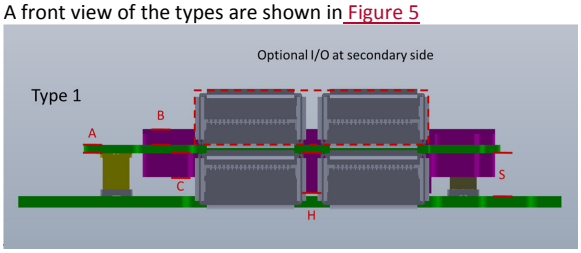
Single/Dual port 10G SFP+ or 40G QSFP Ethernet card should not extend PCB to this area.

Mezzanine card that uses 4x SFP+ ports or 4x 10G Base-T ports is allowed to extend PCB to this area to accommodate I/O connector placement. By doing so, it may break mechanical compatibility of existing platforms.

3.4 Form Factor Vertical Stack Definition

There are 3 options to implement mezzanine card with different placement height restriction, I/O connectors' location, and mezzanine connector stacking height to baseboard.

A front view of the types are shown in Figure 5



~~Figure 5~~, Figure 6~~Figure 6~~, and Figure 7~~Figure 7~~.

For Figure 5~~Figure 5~~ and Figure 6~~Figure 6~~, it shows a possible placement of 4xQSFP connectors. This it is not a current use case; the placement has manufacture concern due to using belle to belle placement of QSFP cage on 1.57mm PCB. It may need a customized QSFP connector for the use case of 4xQSFP.

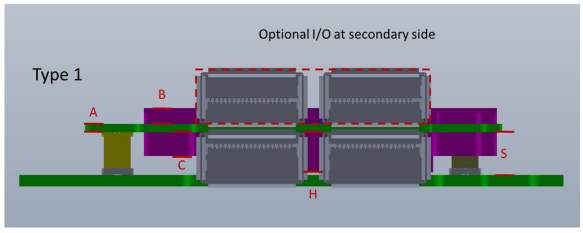


Figure 5: Type 1 Vertical Stack Front View

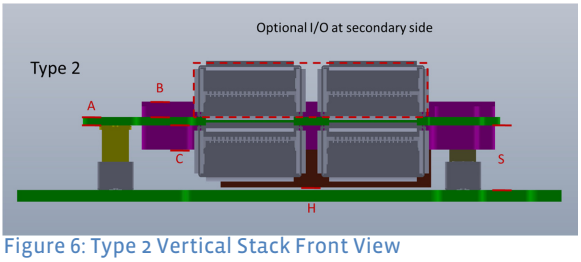


Figure 6: Type 2 Vertical Stack Front View

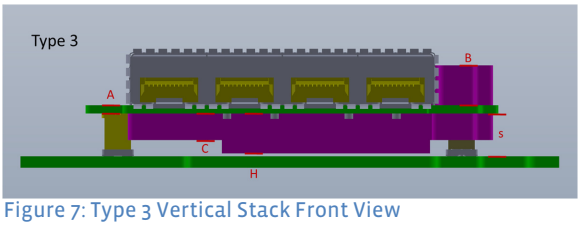


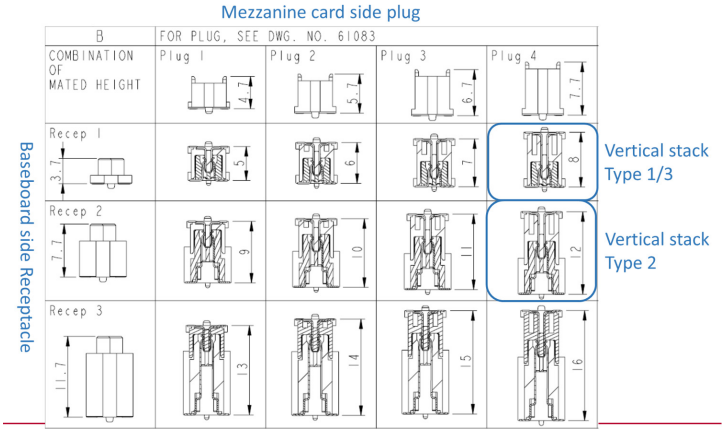
Figure 7: Type 3 Vertical Stack Front View

A summary of major dimension and height restriction across 3 types are shown in Table 1

~~Table 1~~.

Table 1: Mezzanine Card Vertical Stack Types Dimension Comparison

TYPE	A(typ)	B(max)	C(max)	H(max)	S(typ)	I/O	Controller IC
TYPE 1	1.57mm	2.9mm /2.0mm	4.5mm	7.5mm	8mm	Primary side Secondary side optional	Primary side
TYPE 2	1.57mm	2.9mm /2.0mm	4.5mm	11.5mm	12mm	Primary side Secondary side optional	Primary side
TYPE 3	1.57mm	7.5mm	4.5mm	7.5mm	8mm	Secondary side	Primary side/ secondary side



~~Figure 8~~ Figure 8 describes the connector selection for baseboard and mezzanine card to achieve different vertical stack types.[2]

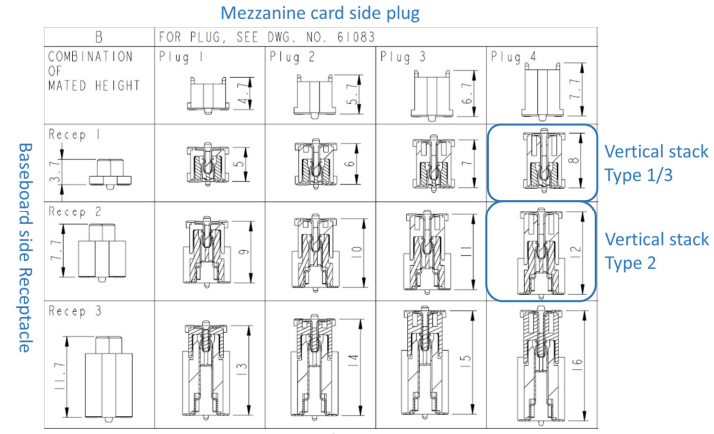


Figure 8: Mezzanine Connector Selection Matrix

3.5 Implementation Examples

This section gives examples of Mezzanine 2.0 implementation. The implementation is not limited to the examples given here as long as it follows the specification.

The 3D screen shots shown below is to illustrate the design. Vendor should follow 3D drawings for detail height restrictions.

[2] Refer to complete drawing for more detail:
http://portal.fciconnect.com/Comergent//fci/drawing/61082.pdf (Receptacle-Baseboard side)
http://portal.fciconnect.com/Comergent//fci/drawing/61083.pdf (Plug-Mezzanine card)

Implementation may results in mechanical conflict with existing OCP platforms. It may trigger modification of mechanical design, or limitation on configuration. A mechanical check in system should be done when planning to use a Mezzanine 2.0 with OCP platforms, or enabling a new Mezzanine 2.0 card.

3.5.1 Single/Dual Port SFP+ 10G Ethernet Mezzanine Card

This is the original OCP 10G Mezzanine card. Specification can be found at the link below:
http://www.opencompute.org/assets/download/Intel-Mezzanine-Card-Design-Specification-v0.5.pdf

3.5.2 Dual QSFP Port 40G Mezzanine Card (Type 1)

This is a single/dual port QSFP, 40G Ethernet Mezzanine card. Depopulate 2nd port makes it a single port 40G Ethernet Mezzanine card. For single port card which need extra space for component placement, placement of component in the volume of 2nd QSFP port is allowed.

On the primary side, there is a component height restriction of 4.5mm and 7.5mm. The 7.5mm height restriction area is intended for heatsink of controller IC; placement of components other than controller IC and heatsink is allowed in this area. 7.5mm max height makes this card fit Type 1 vertical stack, with 8mm stack height.

On the secondary side, there is a component height restriction of 2.0mm and 2.9mm in different areas. As the electrical interface to baseboard, connector A and connector B provide up to x16 PCIe connection. Connector A is mandatory for this SKU. Connector B provides extra x8 PCIe lanes. Connector B is optional for this SKU.

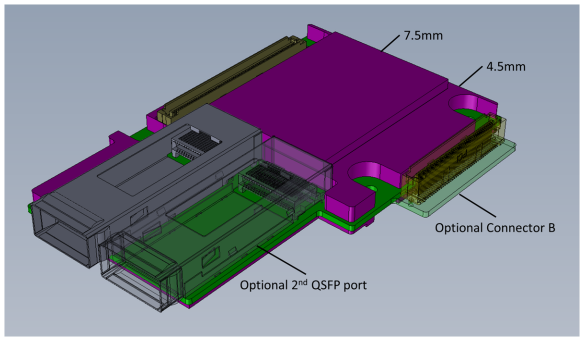


Figure 9: Primary side veiw of dual port QSFP Mezzanine card

3.5.3 Port and LED Location

The port and LED location is shown in Figure 10~~Figure 10~~. Similar to the 10G PCIe, 40G Mezzanine card can depopulated Port 1 to become a single port card.

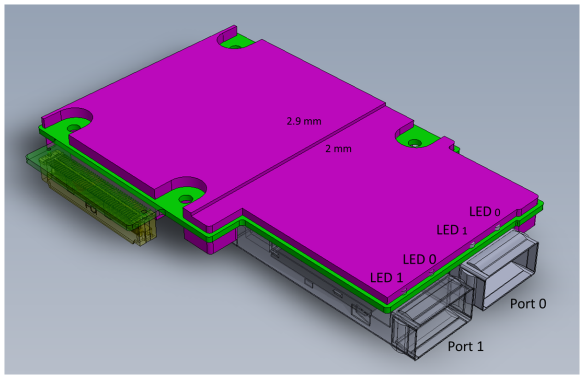


Figure 10: Dual QSFP port Mezzanine card port and LED location

P0-LED0: Port 0, Physical link speed (40G Green/Yellow for other link status)
P0-LED1: Port 0, Logic Link/Activity (Green)
P1-LED0: Port 1, Physical link speed (40G Green/Yellow for other link status)
P1-LED1: Port 1, Logic Link/Activity (Green)

3.5.4 Dual QSFP Port 40G Mezzanine card (Type 2)

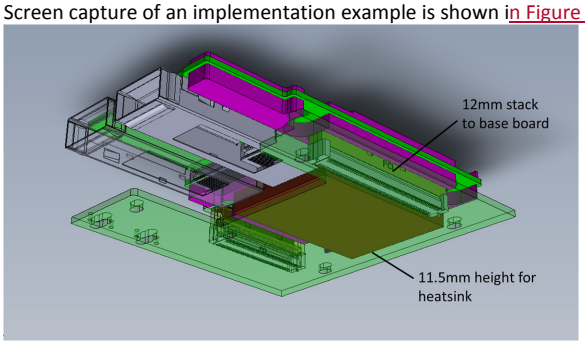
Due to the 8mm stack limitation in Type 1 vertical stack, heat sink height is limited to 7.5mm max and may not be able to provide sufficient cooling to some controller IC. Type 2 vertical stack allows 11.5mm max for heatsink and provide more freedom to thermal design.

This implementation has limitation in system mechanical compatibility due to taking extra volume. Vendor may need to modify mechanical design in order to support it.

Vendor should plan the components in the 11.5mm heatsink area accordingly, if there is a plan to use BOM option to make Type 2 vertical stack fit into Type 1 vertical stack.

The Mezzanine connector is the same for Type 1 and Type 2 on the mezzanine card side. Baseboard side need to use different connectors to support different stacking height for Type 1 and Type 2 as described in section 3.4.

Screen capture of an implementation example is shown in Figure 11.~~n~~



~~Figure 11.~~

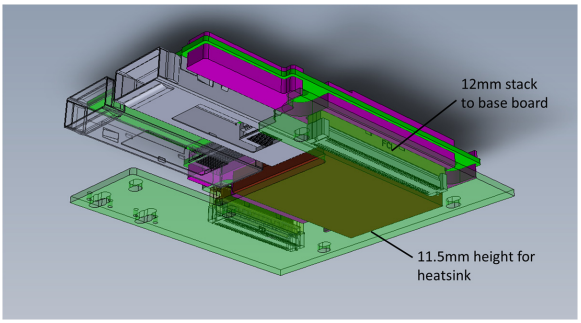


Figure 11: Dual Port QSFP Mezzanine card mounted on a baseboard with 12mm stacking height

3.5.5 Quad SFP+ port 10G Mezzanine card (Type 3)

Quad SFP+ port 10G Mezzanine card can be implement in Type 3 vertical stack up as shown in Figure 12~~Figure 12~~ and Figure 13~~Figure 13~~.

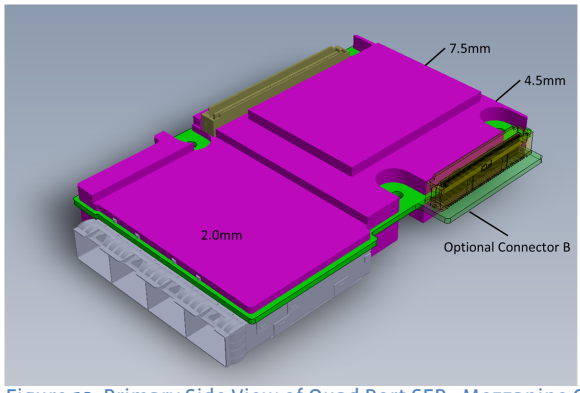


Figure 12: Primary Side View of Quad Port SFP+ Mezzanine Card

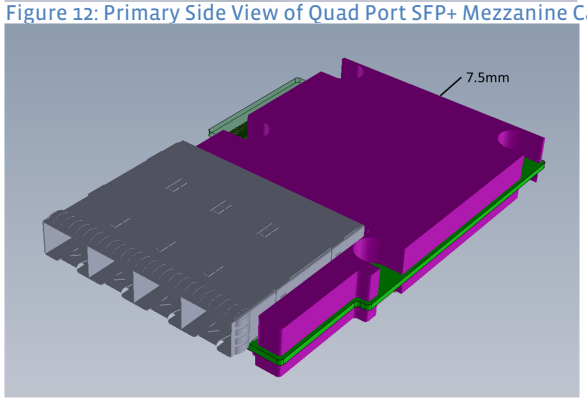
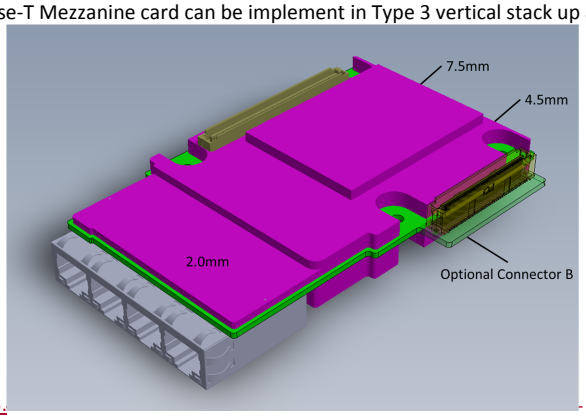


Figure 13: Secondary Side View of Quad Port SFP+ Mezzanine Card

3.5.6 Quad port 10G Base-T Mezzanine Card (Type 3)

Quad 10G Base-T Mezzanine card can be implement in Type 3 vertical stack up as shown in Figure 14 and Figure 15.



~~Figure 14 and 15.~~

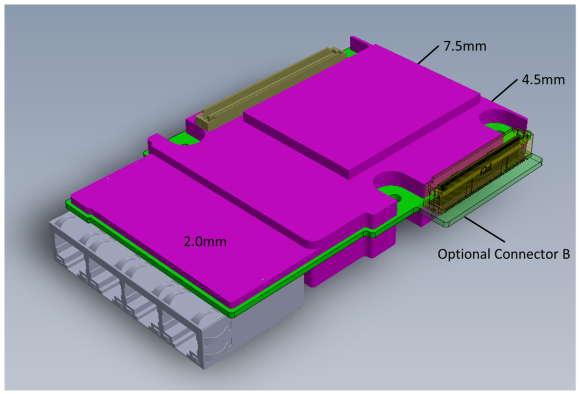


Figure 14: Primary Side View of Quad port 10G Base-T Mezzanine card

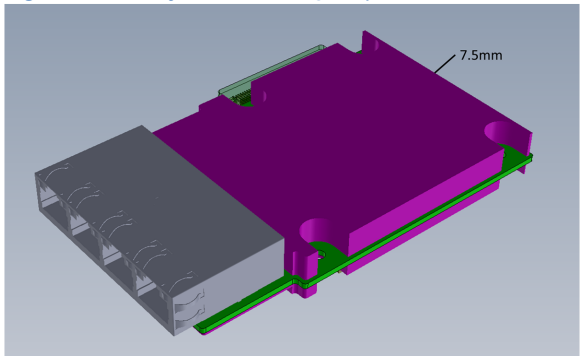


Figure 15: Secondary Side View of Quad port 10G Base-T Mezzanine card

3.6 MAC Address label requirement

MAC address label(s) must be scannable when 10GbE Mezzanine card is installed in server, rack, etc. by system vendor, rack integrator, and DC user without interrupt of normal operation.

For 2x MAC addresses, 2x 2D bar codes and 2x human readable texts for MAC address need to be placed within 10mm from Mezzanine card PCB edge as shown in Figure 16.

For 3x MAC addresses, 3x 2D bar codes and 3x human readable texts for MAC address need to be placed within 18mm from Mezzanine card PCB edge.

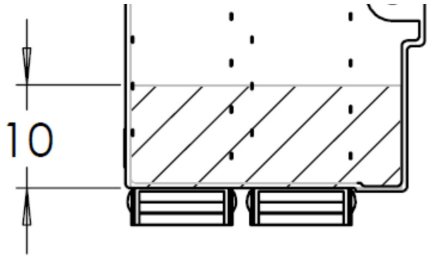


Figure 16: MAC address label placement

The scanned bar code should not include ".". Example: "AA.BB.CC.00.11.20" should scan as "AABBCC001120". Implementation example is shown in Figure 17.

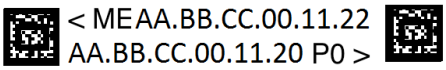


Figure 17: MAC address label implementation example

4 Mezzanine Card to Baseboard Electrical Interface

4.1 Power Capability and Status

Baseboard supplies power to power pins on Mezzanine card connector. There are four power rails available. The current capability and power status is as the table below. Normal power is available at on state S0 only. Auxiliary power is available at all power states including hibernate state S4 or off state S5.

Power Rail	Voltage Tolerance	# of pins	Current Capability	Status
P12V_AUX	±8%(max)	3@A only 5@A+B	2.4A 4.0A	Auxiliary Power
P5V_AUX	±9%(max)	3	2.4A	Auxiliary power
P3V3_AUX	±5%(max)	2	1.6A	Auxiliary power

P3V3	±5%(max)	8	6.4A	Normal power

4.2 Pin Definition of Mezzanine Connector

4.2.1 x16 PCIe Mezzanine Card

Pin definition of a mezzanine card with up to x16 PCIe lanes is in Table 2~~Table 2~~. The direction of the signals are from the perspective of the baseboard.

For mezzanine card with x8 or less PCIe lanes, only Connector A is required. Connector B and its optional PCB area should not be implemented as mentioned in 3.3.1.

Table 2: x16 PCIe Mezzanine Card Pin Definition

Connector A				Connector B			
Signal	Pin	Pin	Signal	Signal	Pin	Pin	Signal
P12V_AUX	A61	A1	MEZZ_PRSNTA1_N /BASEBOARD_A_ID	P12V_AUX	B41	B1	MEZZ_PRSNTB1_N /BASEBOARD_B_ID
P12V_AUX	A62	A2	P5V_AUX	P12V_AUX	B42	B2	GND
P12V_AUX	A63	A3	P5V_AUX	RSVD	B43	B3	MEZZ_RX_DP<8>
GND	A64	A4	P5V_AUX	GND	B44	B4	MEZZ_RX_DN<8>
GND	A65	A5	GND	MEZZ_TX_DP<8>	B45	B5	GND
P3V3_AUX	A66	A6	GND	MEZZ_TX_DN<8>	B46	B6	GND
GND	A67	A7	P3V3_AUX	GND	B47	B7	MEZZ_RX_DP<9>
GND	A68	A8	GND	GND	B48	B8	MEZZ_RX_DN<9>
P3V3	A69	A9	GND	MEZZ_TX_DP<9>	B49	B9	GND
P3V3	A70	A10	P3V3	MEZZ_TX_DN<9>	B50	B10	GND
P3V3	A71	A11	P3V3	GND	B51	B11	MEZZ_RX_DP<10>
P3V3	A72	A12	P3V3	GND	B52	B12	MEZZ_RX_DN<10>
GND	A73	A13	P3V3	MEZZ_TX_DP<10>	B53	B13	GND
LAN_3V3STB_ALERT_N	A74	A14	NCSI_RCSDV	MEZZ_TX_DN<10>	B54	B14	GND
SMB_LAN_3V3STB_CLK	A75	A15	NCSI_RCLK	GND	B55	B15	MEZZ_RX_DP<11>
SMB_LAN_3V3STB_DAT	A76	A16	NCSI_TXEN	GND	B56	B16	MEZZ_RX_DN<11>
PCIE_WAKE_N	A77	A17	PERST_N0	MEZZ_TX_DP<11>	B57	B17	GND
NCSI_RXER	A78	A18	MEZZ_SMCLK	MEZZ_TX_DN<11>	B58	B18	GND
GND	A79	A19	MEZZ_SMDATA	GND	B59	B19	MEZZ_RX_DP<12>
NCSI_TXD0	A80	A20	GND	GND	B60	B20	MEZZ_RX_DN<12>
NCSI_TXD1	A81	A21	GND	MEZZ_TX_DP<12>	B61	B21	GND
GND	A82	A22	NCSI_RXD0	MEZZ_TX_DN<12>	B62	B22	GND
GND	A83	A23	NCSI_RXD1	GND	B63	B23	MEZZ_RX_DP<13>
CLK_100M_MEZZ1_DP	A84	A24	GND	GND	B64	B24	MEZZ_RX_DN<13>
CLK_100M_MEZZ1_DN	A85	A25	GND	MEZZ_TX_DP<13>	B65	B25	GND
GND	A86	A26	CLK_100M_MEZZ2_DP	MEZZ_TX_DN<13>	B66	B26	GND
GND	A87	A27	CLK_100M_MEZZ2_DN	GND	B67	B27	MEZZ_RX_DP<14>
MEZZ_TX_DP_C<0>	A88	A28	GND	GND	B68	B28	MEZZ_RX_DN<14>
MEZZ_TX_DN_C<0>	A89	A29	GND	MEZZ_TX_DP<14>	B69	B29	GND
GND	A90	A30	MEZZ_RX_DP<0>	MEZZ_TX_DN<14>	B70	B30	GND
GND	A91	A31	MEZZ_RX_DN<0>	GND	B71	B31	MEZZ_RX_DP<15>
MEZZ_TX_DP_C<1>	A92	A32	GND	GND	B72	B32	MEZZ_RX_DN<15>
MEZZ_TX_DN_C<1>	A93	A33	GND	MEZZ_TX_DP<15>	B73	B33	GND
GND	A94	A34	MEZZ_RX_DP<1>	MEZZ_TX_DN<15>	B74	B34	GND
GND	A95	A35	MEZZ_RX_DN<1>	GND	B75	B35	CLK_100M_MEZZ3_DP
MEZZ_TX_DP_C<2>	A96	A36	GND	GND	B76	B36	CLK_100M_MEZZ3_DN
MEZZ_TX_DN_C<2>	A97	A37	GND	CLK_100M_MEZZ4_DP	B77	B37	GND
GND	A98	A38	MEZZ_RX_DP<2>	CLK_100M_MEZZ4_DN	B78	B38	PERST_N1
GND	A99	A39	MEZZ_RX_DN<2>	GND	B79	B39	PERST_N2

MEZZ_TX_DP_C<3>	A100	A40	GND	MEZZ_PRSNTB2_N	B80	B40	PERST_N3
MEZZ_TX_DN_C<3>	A101	A41	GND				
GND	A102	A42	MEZZ_RX_DP<3>				
GND	A103	A43	MEZZ_RX_DN<3>				
MEZZ_TX_DP_C<4>	A104	A44	GND				
MEZZ_TX_DN_C<4>	A105	A45	GND				
GND	A106	A46	MEZZ_RX_DP<4>				
GND	A107	A47	MEZZ_RX_DN<4>				
MEZZ_TX_DP_C<5>	A108	A48	GND				
MEZZ_TX_DN_C<5>	A109	A49	GND				
GND	A110	A50	MEZZ_RX_DP<5>				
GND	A111	A51	MEZZ_RX_DN<5>				
MEZZ_TX_DP_C<6>	A112	A52	GND				
MEZZ_TX_DN_C<6>	A113	A53	GND				
GND	A114	A54	MEZZ_RX_DP<6>				
GND	A115	A55	MEZZ_RX_DN<6>				
MEZZ_TX_DP_C<7>	A116	A56	GND				
MEZZ_TX_DN_C<7>	A117	A57	GND				
GND	A118	A58	MEZZ_RX_DP<7>				
GND	A119	A59	MEZZ_RX_DN<7>				
MEZZ_PRSNTA2_N	A120	A60	GND				

For x16 PCIe, lane 0~7 is mapped to connector A and lane 8~15 is mapped to connector B.

For the case of multiple root ports are connected to Mezzanine interface on base board, or the case of multiple end points are connected to Mezzanine interface on Mezzanine card, follow bifurcation rule as showing in Table 3.

Table 3: Bifuracation rule of PCIe in connector A and Connect B

Bifurcation		Lane numbering															
# of ports	lane/port	*Pin 1 Connector A								*Pin 1 Connector B							
1	x16	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	x8	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7
4	x4	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
8	x2	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
16	x1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

4.2.2 16x KR ~~Phy~~ Mezzanine card

Pin definition of a Mezzanine card with up to 16 KR lanes from a baseboard. There are PHY or retimer ~~PHY~~ on this Mezzanine card for connecting to rack level network.

Pin definition of 16x KR ~~Phy~~ Mezzanine card is shown in Table 4~~Table 4~~. The direction of the signals are from the perspective of the baseboard.

For KR~~Phy~~ Mezz that uses 8x or less KR pairs, only Connector A is required. Connector B, and its optional PCB area, should not be implemented as mentioned in 3.3.1.

~~Repeaters can also be used on Phy Mezz, to replace Phy for system or IC with build-in Phy.~~

Table 4: 16x KR ~~Phy~~ Mezzanine Card Pin Definition

Connector A				Connector B			
Signal	**Pin**	**Pin**	**Signal**	**Signal**	**Pin**	**Pin**	**Signal**
P12V_AUX	A61	A1	MEZZ_PRSNTA1_N /BASEBOARD_A_ID	P12V_AUX	B41	B1	MEZZ_PRSNTB1_N /BASEBOARD_B_ID
P12V_AUX	A62	A2	P5V_AUX	P12V_AUX	B42	B2	GND
P12V_AUX	A63	A3	P5V_AUX	RSVD	B43	B3	KR_RX_DP<12>
GND	A64	A4	P5V_AUX	GND	B44	B4	KR_RX_DN<12>
GND	A65	A5	GND	KR_TX_DP<12>	B45	B5	GND
P3V3_AUX	A66	A6	GND	KR_TX_DN<12>	B46	B6	GND
GND	A67	A7	P3V3_AUX	GND	B47	B7	KR_RX_DP<13>
GND	A68	A8	GND	GND	B48	B8	KR_RX_DN<13>
P3V3	A69	A9	GND	KR_TX_DP<13>	B49	B9	GND
P3V3	A70	A10	P3V3	KR_TX_DN<13>	B50	B10	GND
P3V3	A71	A11	P3V3	GND	B51	B11	KR_RX_DP<4>
P3V3	A72	A12	P3V3	GND	B52	B12	KR_RX_DN<4>
GND	A73	A13	P3V3	KR_TX_DP<4>	B53	B13	GND
LAN_3V3STB_ALERT_N	A74	A14	NCSI_RCSDV	KR_TX_DN<4>	B54	B14	GND
SMB_LAN_3V3STB_CLK	A75	A15	NCSI_RCLK	GND	B55	B15	KR_RX_DP<5>
SMB_LAN_3V3STB_DAT	A76	A16	NCSI_TXEN	GND	B56	B16	KR_RX_DN<5>
RSVD	A77	A17	RSVD	KR_TX_DP<5>	B57	B17	GND
NCSI_RXER	A78	A18	MEZZ_SMCLK	KR_TX_DN<5>	B58	B18	GND
GND	A79	A19	MEZZ_SMDATA	GND	B59	B19	KR_RX_DP<14>
NCSI_TXD0	A80	A20	GND	GND	B60	B20	KR_RX_DN<14>
NCSI_TXD1	A81	A21	GND	KR_TX_DP<14>	B61	B21	GND
GND	A82	A22	NCSI_RXD0	KR_TX_DN<14>	B62	B22	GND
GND	A83	A23	NCSI_RXD1	GND	B63	B23	KR_RX_DP<15>
RSVD	A84	A24	GND	GND	B64	B24	KR_RX_DN<15>
RSVD	A85	A25	GND	KR_TX_DP<15>	B65	B25	GND
GND	A86	A26	RSVD	KR_TX_DN<15>	B66	B26	GND
GND	A87	A27	RSVD	GND	B67	B27	KR_RX_DP<6>
KR_TX_DP<8>	A88	A28	GND	GND	B68	B28	KR_RX_DN<6>
KR_TX_DN<8>	A89	A29	GND	KR_TX_DP<6>	B69	B29	GND
GND	A90	A30	KR_RX_DP<8>	KR_TX_DN<6>	B70	B30	GND
GND	A91	A31	KR_RX_DN<8>	GND	B71	B31	KR_RX_DP<7>
KR_TX_DP<9>	A92	A32	GND	GND	B72	B32	KR_RX_DN<7>
KR_TX_DN<9>	A93	A33	GND	KR_TX_DP<7>	B73	B33	GND
GND	A94	A34	KR_RX_DP<9>	KR_TX_DN<7>	B74	B34	GND
GND	A95	A35	KR_RX_DN<9>	GND	B75	B35	RSVD
KR_TX_DP<0>	A96	A36	GND	GND	B76	B36	RSVD
KR_TX_DN<0>	A97	A37	GND	RSVD	B77	B37	GND
GND	A98	A38	KR_RX_DP<0>	RSVD	B78	B38	RSVD
GND	A99	A39	KR_RX_DN<0>	GND	B79	B39	RSVD
KR_TX_DP<1>	A100	A40	GND	MEZZ_PRSNTB2_N	B80	B40	RSVD
KR_TX_DN<1>	A101	A41	GND				
GND	A102	A42	KR_RX_DP<1>				
GND	A103	A43	KR_RX_DN<1>				
KR_TX_DP<10>	A104	A44	GND				
KR_TX_DN<10>	A105	A45	GND				
GND	A106	A46	KR_RX_DP<10>				
GND	A107	A47	KR_RX_DN<10>				
KR_TX_DP<11>	A108	A48	GND				
KR_TX_DN<11>	A109	A49	GND				
GND	A110	A50	KR_RX_DP<11>				
GND	A111	A51	KR_RX_DN<11>				

KR_TX_DP<2>	A112	A52	GND
KR_TX_DN<2>	A113	A53	GND
GND	A114	A54	KR_RX_DP<2>
GND	A115	A55	KR_RX_DN<2>
KR_TX_DP<3>	A116	A56	GND
KR_TX_DN<3>	A117	A57	GND
GND	A118	A58	KR_RX_DP<3>
GND	A119	A59	KR_RX_DN<3>
MEZZ_PRSNT2_N	A120	A60	GND

For KR~~Phy~~ Mezz with 4, 8, or 16 KR channels, follow Table 5 to assign the sequence.

Table 5: KR/Repeater numbering sequence

KR/Repeater Numbering Sequence																
# of KR	*Pin 1		Connector A						*Pin 1		Connector B					
4 KR			0	1			2	3								
8 KR			0	1			2	3			4	5			6	7
16 KR	8	9	0	1	10	11	2	3	12	13	4	5	14	15	6	7

4.3 Mezzanine Card Pin Description

Mezzanine card pin description is shown in Table 6~~Table 6~~; input output direction is in the prospective of baseboard.

Table 6: Mezzanine Card Pin Description

Signals on Connector A	Type	Description
GND	Ground	Ground return; total 51 pins on Connector A
P12V_AUX	Power	12V Aux power; total 3 pins on Connector A
P5V_AUX	Power	5V Aux power; total 3 pins on Connector A
P3V3_AUX	Power	P3V3 Aux Power; total 2 pins on Connector A
P3V3	Power	P3V3 power; total 8 pins on Connector A
MEZZ_PRSNTA1_N/BASEBOARD_ID_A	Output	Connector A Present Pin; connect to MEZZ_PRSNTA2_N on Mezz with 0 Ohm; Use as baseboard ID during power up
MEZZ_PRSNTA2_N	Input	Connector A Present Pin; connect to MEZZ_PRSNTA1_N on Mezz with 0 Ohm
LAN_3V3STB_ALERT_N	Input	SMBus Alert for OOB management; 3.3V AUX rail
SMB_LAN_3V3STB_CLK	Output	SMBus Clock for OOB management; 3.3V AUX rail; Share with thermal reporting interface
SMB_LAN_3V3STB_DAT	Bidirectional	SMBus Data for OOB management; 3.3V AUX rail; Share with thermal reporting interface
NCSI_RXER	Input	NC-SI for OOB management
NCSI_RCSDV	Input	NC-SI for OOB management
NCSI_RXD[1..0]	Input	NC-SI for OOB management
NCSI_RCLK	Output	NC-SI for OOB management
NCSI_TXEN	Output	NC-SI for OOB management

NCSI_TXD[1..0]	Output	NC-SI for OOB management
PCIE_WAKE_N	Input	PCIe wake up signal
PERST_N0	Output	PCIe reset signal or Node 1 PCIe reset signal
MEZZ_SMCLK	Output	PCIe SMBus Clock for Mezz slot/EEPROM; 3.3V AUX rail; Share with thermal reporting interface
MEZZ_SMDATA	Bidirectional	PCIe SMBus Data for Mezz slot/EEPROM; 3.3V AUX rail; Share with thermal reporting interface
CLK_100M_MEZZ[2..1]_DP/N	Output	MB clock output for PCIe devices; total 2 pairs on Connector A
MEZZ_TX_DP/N_C<7..0>	Output	PCIe TX; total up to 8 lanes on Connector A; optional with KR signals
MEZZ_RX_DP/N<7..0>	Input	PCIe RX; total up to 8 lanes on Connector A; optional with KR signals
KR_TX_DP/N<15..8>	Output	KR TX; total up to 8 lanes on Connector A; optional with PCIe signals
KR_RX_DP/N<15..8>	Input	KR RX; total up to 8 lanes on Connector A; optional with PCIe signals
RSVD	TBD	Reserved for Future use

Signals on Connector B	**Type**	**Description**
GND	Ground	Ground return; total 36 pins on Connector B
P12V_AUX	Power	12V Aux power; total 2 pins on Connector B
MEZZ_PRSNTB1_N/ BASEBOARD_ID_B	Output	Connector B Present Pin; connect to MEZZ_PRSNTB2_N on Mezz with 0 Ohm Use as baseboard ID during power up
MEZZ_PRSNTB2_N	Input	Connector B Present Pin; connect to MEZZ_PRSNTB1_N on Mezz with 0 Ohm
PERST_N[3..1]	Output	PCIe reset signal or Node[3..1] PCIe reset signal for baseboard with more than 1 nodes
CLK_100M_MEZZ[4..3]_DP/N	Output	MB clock output for PCIe devices; total 2 pairs on Connector B
MEZZ_TX_DP/N_C<15..8>	Output	PCIe TX; total up to 8 lanes on Connector B; optional with KR signals
MEZZ_RX_DP/N<15..8>	Input	PCIe RX; total up to 8 lanes on Connector B; optional with KR signals
KR_TX_DP/N<7..0>	Output	KR TX; total up to 8 lanes on Connector B; optional with PCIe signals

KR_RX_DP/N<7..0>	Input	KR RX; total up to 8 lanes on Connector B; optional with PCIe signals
RSVD	TBD	Reserved for Future use

4.3.1 MEZZ FRU EEPROM

MEZZ FRU EEPROM is for baseboard to identify different types of Mezzanine card. MEZZ FRU EEPROM is connected to MEZZ_SMCLK/MEZZ_SMDATA (pin A18, A19) and address is 0xA2 (8bit format). The size of EEPROM is 1Kbits.

Follow IPMI Platform Management FRU Information Storage Definition v1.0 for data format. Use OEM record 0xC0, offset 1 to store Mezzanine ID definition.

Table 7: Mezzanine ID byte definition

Mezz ID Byte(offset 1)	Usage
0x00	X8 PCIe on Connector A
0x01	X16 PCIe on Connector A and Connector B
0x02	X4 KR with ~~Repeater~~Retimer
0x03	X4 KR with PHY
0x04	Two X4 PCIe devices on Connector A
0x05	One x8 PCIe device on Connector A, one x8 PCIe device on Connector B
All others read back	RFU
No FRU device detected	X8 PCIe on Connector A

If Baseboard cannot find EEPROM on Mezzanine card, Baseboard will assume the Mezzanine cards is original Mezzanine card which has PCIe interface. This provides backward compatibility for Mezzanine card without ID EEPROM.

4.3.2 Baseboard ID

Baseboard with one x8 or one x16 PCIe port does not need to implement Baseboard ID circuit. MEZZ_PRSNTA1_N in connector A and MEZZ_PRSNTB1_N in connector B is connected to ground for this case.

Baseboard ID is an optional feature for special baseboard to identify itself to Mezzanine card. It is only implemented when Mezzanine card needs to have awareness of different baseboard types, and the baseboard Mezzanine card interface is not a single root port PCIe.

The implementation example of Baseboard ID circuit is shown in Figure 18~~Figure 16~~. Connecter A is shown here as an example, and implementation for Connecter B is same.

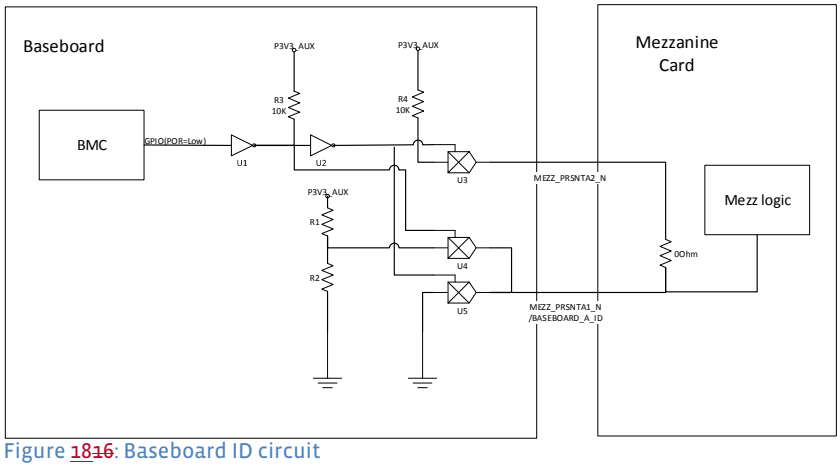


Figure 18~~16~~: Baseboard ID circuit

Mezzanine card identifies different of Baseboard based on the resistor pair R1/R2 shown in Table 8.

Table 8: Baseboard ID definiton

ConnA R1	ConnA R2	ConnB R1	ConnB R2	Baseboard type
NC	0Ohm	Connecter B NC		PCIe x8
NC	0Ohm	NC	0Ohm	PCIe x16
Other				Reserved

Mezzanine card has a 4 seconds window to sample Baseboard ID signal, after both P12V_AUX and P3V3_AUX ready. BMC may disable baseboard ID output by GPIO after 4 seconds and change the multipurpose pin for Mezzanine card present purpose.

5 Management Interface

There are two options of management interfaces on 10GbE or 40GbE Mezzanine card for BMC's out-of-band communication. Both interfaces should be routed from mezzanine card connector to NIC chipset. It is preferred that BMC firmware can choose to hand shake with either interfaces to have out-of-band channel. Only one interface need to be up and running for out-of-band traffic at a given time.

5.1 I2C side band

10Gbe or 40Gbe Mezzanine card implements management interface compatible with Intel's Management Engine (ME) through C600 PCH SMLINK0 port and provides Out of Band (OOB) network access. Vendor should check with Facebook to choose SMBus address for ME OOB access. The hardware and firmware design need to support management capability in both S0 and S5 state.

The same I2C interface should be able to be accessed by BMC (baseboard management controller) on platform that used BMC.

I2C sideband interface should support both IPv4 and IPv6.

5.2 NC-SI side band

An optional NC-SI management interface can be implemented by 10GbE or 40GbE Mezzanine card. It is essential to achieve management feature needs higher bandwidth such as upload and update baseboard firmware and BIOS through OOB. Compare to original OCP Mezzanine card Specification, 8x RSVD pins are redefined to NC-SI interface.

The total length of each NC-SI signal and clock from connector pin to BGA pin on mezzanine card should be greater or equal to 1500mil and less or equal to 3500mil. NC-SI clock and signal should be matched within 1000mil on the mezzanine card.

It is preferred that 10GbE or 40GbE Mezzanine card is able to be connected by different management controllers and interfaces, such as Management engine, BMC I2C, or BMC NC-SI. It is for having backward and forward compatibility with same hardware and firmware.

Management controller should set priority for its capable connection interfaces in a sequence, and scan through the sequence to hand shake with the 1st available management network device. It is to ensure the Mezzanine card without NC-SI side band can still be compatible with baseboard with NC-SI capability. NC-SI sideband interface should support both IPv4 and IPv6.

5.3 MAC address of management interface

MAC address of management network interface should be a positive offset based on the MAC address of data network interface. Different vendor may implement different offset number based on port count, and reserved features. Typical offsets are +1 for single port NIC, +2 for dual port NIC. Some NIC have larger offset due to having more than 2 ports, or having more than one MAC of data or storage on each port.

6 10GbE/40GbE Mezzanine card Data network

6.1 Network booting

Mezzanine NIC card shall support network booting in uEFI system environment. 10GbE Mezzanine card shall support both IPv4 and IPv6 network booting.

7 Thermal Reporting Interface

7.1 Overview of Thermal Reporting Interface

A thermal reporting interface is defined on SMB_LAN_3V3STB_CLK/SMB_LAN_3V3STB_DAT (Connector A, pin A75, A76) or MEZZ_SMCLK/ MEZZ_SMDATA (Connector A, pin A18, A19). The implementation of this requirement will improve the thermal management of system and allow baseboard management device to access key component temperature on Mezzanine card. Baseboard management device needs to scan SMB_LAN_3V3STB_CLK/SMB_LAN_3V3STB_DAT and MEZZ_SMCLK/ MEZZ_SMDATA to determine the location of the thermal reporting interface.

There are two methods to implement thermal reporting described in this section: Emulated method and remote on-die sensing method. Both methods will be treated by baseboard management controller as a TI/TMP421 thermal sensor with slave address 0x3E in 8 bit format.

For Mezzanine card with Thermal Design Power > 5 Watts, this implementation of this interface is required.

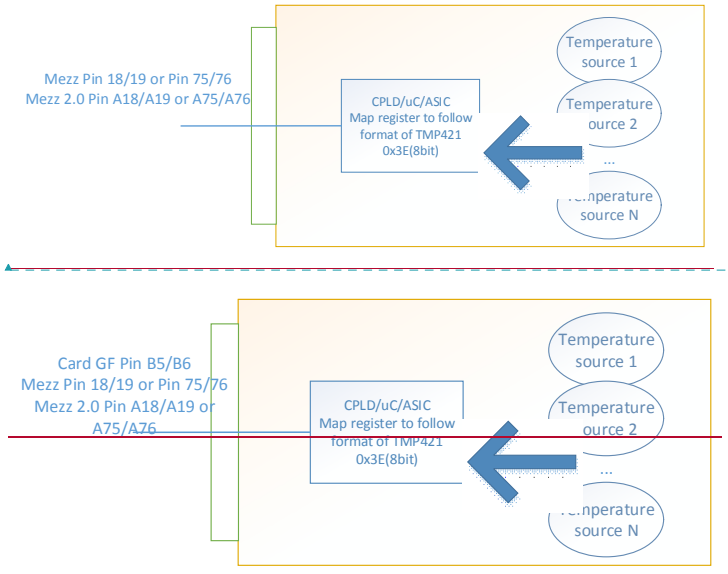
7.1.1 Emulated Thermal Reporting

Mezzanine card should emulate its key temperatures to be accessed from SMBus (Connector A, Pin A18/A19 or pin A75~~,~~/A76; P3V3_STBY rail). The emulation should follow TMP421 register mapping[3]. Baseboard treats the PCIe card thermal sensor as TMP421. Baseboard BMC controller should use 2x separate reads to obtain the MSB and LSB of temperature data. Data obtained is used for system thermal monitoring and fan speed control.

There are two temperatures for TMP421 register mapping, local and remote channel 1. Remote channel 1 is typically used to represent key controller temperature of the card. Local channel is typically used to represent highest of other key components temperature on the card, such as highest temperature of active cable module.

Address of the emulated TMP421 device is fixed at 0x3E in 8bit format.

An implementation block diagram is shown in Figure 19~~Figure 17~~.



[3] TMP421 specification: http://www.ti.com/lit/ds/sbos398c/sbos398c.pdf

Field Code Changed

Figure 1917: Block Diagram for Emulated Thermal Reporting

With firmware change of the controller on baseboard managing thermal data and control, a register mapping of TMP422/TMP423 can be used to support one/two more temperatures without hardware change. The slave address of emulated device is always 0x3E, even it emulates TMP422/TMP423.

Vendor ID and device ID is mapped to offset 0xFE and 0xFF in order for board management controller to detect card types.

Power reporting and power capping is mapped to offset 0xF2 and 0xF3 as an optional feature to achieve device power monitoring and power capping level setting.

Table 9:~~Table 8:~~ describes the register implementation requirement for emulated method.

Table 98: Implementation Requirement for TMP421 Registers

Offset	Description	Original TMP offset	Implementation requirement for emulated method
0x0	Local Temperature (High Byte)	Y	Represents highest temperature of all other key components Required if any of the other key components or modules are critical for thermal design Otherwise it is an optional offset and return 0x00 if not used
0x1	Remote Temperature 1 (High Byte)	Y	Required; represent temperature of main controller
0x2	Remote Temperature 2 (High Byte)	Y	Optional; represent temperature of key component 1; return 0x00 if not used
0x3	Remote Temperature 3 (High Byte)	Y	Optional; represent temperature of key component 2; return 0x00 if not used
0x8	Status Register	Y	Not required
0x9	Configuration Register 1	Y	Not required; Emulated behavior follows SD=0, Temperature Range=0
0x0A	Configuration Register 2	Y	Required; follow TMP423 datasheet to declare the channel supported; RC=1
0x0B	Conversion Rate Register	Y	Not required; Equivalent emulated conversion rate should be >2 sample/s
0x0F	One-Shot Start	Y	Not required
0x10	Local Temperature (Low Byte)	Y	Optional; return 0x00 if not used
0x11	Remote Temperature 1 (Low Byte)	Y	Optional; return 0x00 if not used
0x12	Remote Temperature 2 (Low Byte)	Y	Optional; return 0x00 if not used
0x13	Remote Temperature 3 (Low Byte)	Y	Optional; return 0x00 if not used
0x21	N Correction 1	Y	Not required

0x22	N Correction 2	Y	Not required
0x23	N Correction 3	Y	Not required
0xF0	Manufacturer ID(High Byte)	N	High byte of PCIe vendor ID, if using emulated temperature sensor method
0xF1	Device ID(High Byte)	N	High byte of PCIe device ID, if using emulated temperature sensor method
0xF2	Power reporting	N	Optional; card power reporting; 1LSB=1W; Read only
0xF3	Power capping	N	Optional; card power capping; 1LSB=1W; Read/Write
0xFC	Software Reset	Y	Not required
0xFE	Manufacturer ID	Y(redefined)	Low byte of PCIe vendor ID, if using emulated temperature sensor method
0xFF	Device ID	Y(redefined)	Low byte of PCIe device ID, if using emulated temperature sensor method

7.1.2 Remote on-die sensing

Alternatively, one TMP421 sensor can be used to do on die temperature sensing for IC with thermal diode interface with TMP421 remote sensing channel; Connection diagram is shown in Figure 20~~Figure 18~~.

For NIC needs more than one remote on-die sensing, TMP422/TMP423 can be used and slave address is 0x98(8bit) for this case.

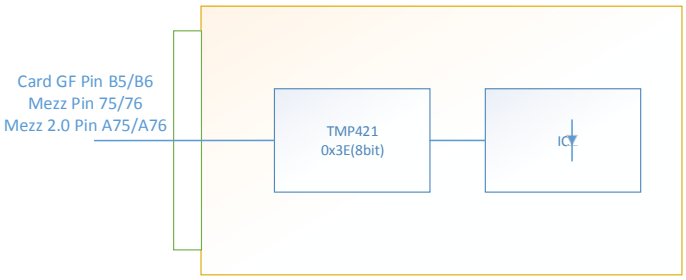


Figure ~~20~~18: Block Diagram for Remote on-die Sensing

8 Environmental

8.1 Environmental Requirements

This Mezzanine card shall meet the same environmental requirements specified in updated Windmill and Leopard Motherboard Design Specification. Minimum airflow is 170 LFM at S0 and 0 LFM at S5. Maximum inlet ambient temperature is 35°C. Connectors are at the inlet of cold air.

8.2 Shock & Vibration

This Mezzanine card shall meet the same shock & vibration requirements specified in updated Windmill and Leopard Motherboard Design Specification.

8.3 Regulation

This Mezzanine card shall meet CE, ~~and~~ CB, FCC Class A, WEEE, ROHS requirements.

9 Revision History

Author	Description	Revision	Date
Jia Ning	-Initial draft for community feedback	0.2	5/18/2014
Jia Ning	-Typical correction and clarification	0.21	5/20/2014
Jia Ning	-Add Mezzanine FRU and Baseboard ID -Remove Mezzanine ID resistor network -Change thermal reporting interface from pin 18, 19 to pin 75, 76 - Update pin define table and description	0.31	6/18/2014
Jia Ning	- Correct Pin number from 18/19 to 75/76 in Figure 18 - Correct description for MEZZ_SMCLK and MEZZ_SMDATA in Table 4	0.32	7/9/2014
Jia Ning	- Add option for TMP422/TMP423 to be used for thermal reporting in Section 7.1.2 - Add clarification of air flow direction and air flow information in section 8.1 - Add clarification of 4x QSFP use case in section 3.4 - Add bifurcation rule in section 4.2.1 - Update Phy Mezz table with new port sequence; add repeater option to Phy Mezz - Update FRU EEPROM format in section 4.3.1	0.33	7/16/14
Jia Ning	- Add new Mezz ID per community feedback	0.34	7/19/2014
Jia Ning	- Add section 3.6 for MAC label requirements - Modify thermal reporting interface to be on A18/A19 or A75/A76. - Add clarification for mechanical compatibility in section 3.2.2 - Add CB, WEEE, ROHS into regulation requirement - Change PHY Mezz to KR Mezz - Format clean up	0.40	8/1/2014