OPEN

Compute Project

25Gb/s Ethernet Mezzanine Card

Mellanox®

TECHNOLOGIES

Connect. Accelerate. Outperform.™

1 Contents

2 Overview

2.1 License

As of April 7, 2011, the following persons or entities have made this Specification available under the Open Web Foundation Final Specification Agreement (OWFa 1.0), which is available at http://www.openwebfoundation.org/legal/the-owf-1-0-agreements/owfa-1-0:

Facebook, Inc.

You can review the signed copies of the Open Web Foundation Agreement Version 1.0 for this Specification at http://opencompute.org/licensing/, which may also include additional parties to those listed above.

Your use of this Specification may be subject to other third party rights. THIS SPECIFICATION IS PROVIDED "AS IS." The contributors expressly disclaim any warranties (express, implied, or otherwise), including implied warranties of merchantability, non-infringement, fitness for a particular purpose, or title, related to the Specification. The entire risk as to implementing or otherwise using the Specification is assumed by the Specification implementer and user. IN NO EVENT WILL ANY PARTY BE LIABLE TO ANY OTHER PARTY FOR LOST PROFITS OR ANY FORM OF INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES OF ANY CHARACTER FROM ANY CAUSES OF ACTION OF ANY KIND WITH RESPECT TO THIS SPECIFICATION OR ITS GOVERNING AGREEMENT, WHETHER BASED ON BREACH OF CONTRACT, TORT (INCLUDING NEGLIGENCE), OR OTHERWISE, AND WHETHER OR NOT THE OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

3 Card Features

- ConnectX®-4 Lx EN based
- Cards following OCP Type1 Stacking Height:
  - Single and Dual Port Card: 25GbE (SFP28)
  - Single and Dual Port Card: 25GbE (SFP28) with Host-Management Support
- Cards following OCP Type4 Stacking Height:
  - Single Port Card: 25GbE (SFP28)
- PCI-e x8 Gen 3.0
- NC-SI (Enabled for cards with Host-Management support)
- Power rail supply of 5VAUX and 3.3VAUX
- 16MB Flash

For more detailed information, please refer to 25GbE User Manual for OCP.

3.1 Form Factor

The following figures illustrate the cards' form factor and dimensions (in mm). Refer to the 2D DXF and 3D files for dimensions, tolerance, and height restrictions.

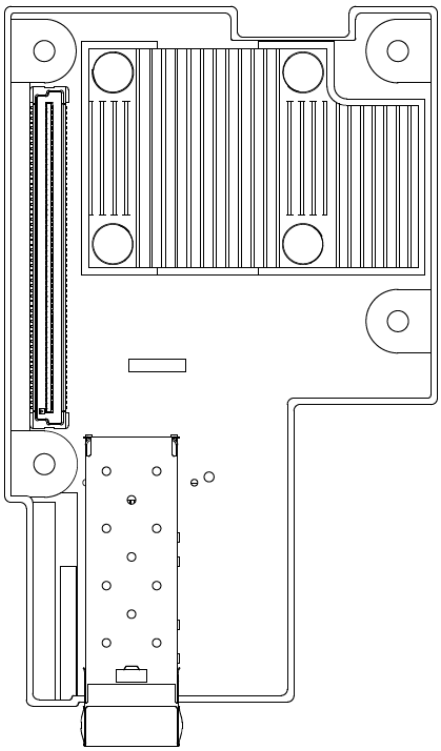


Figure 1: Single-Port Form Factor (Top)

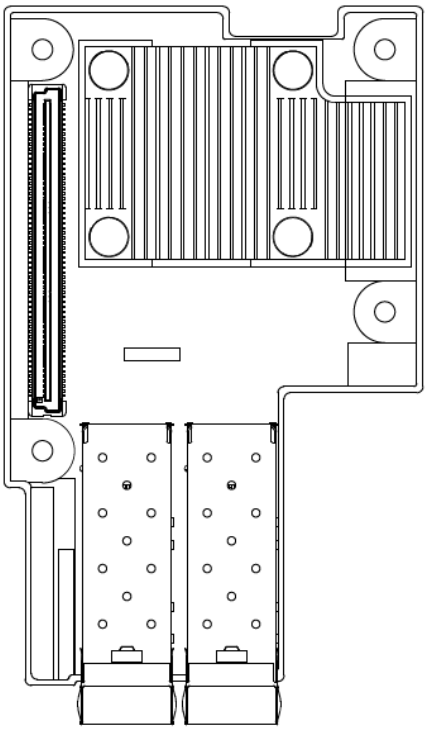


Figure 2: Dual-Port Card Form Factor (Top)

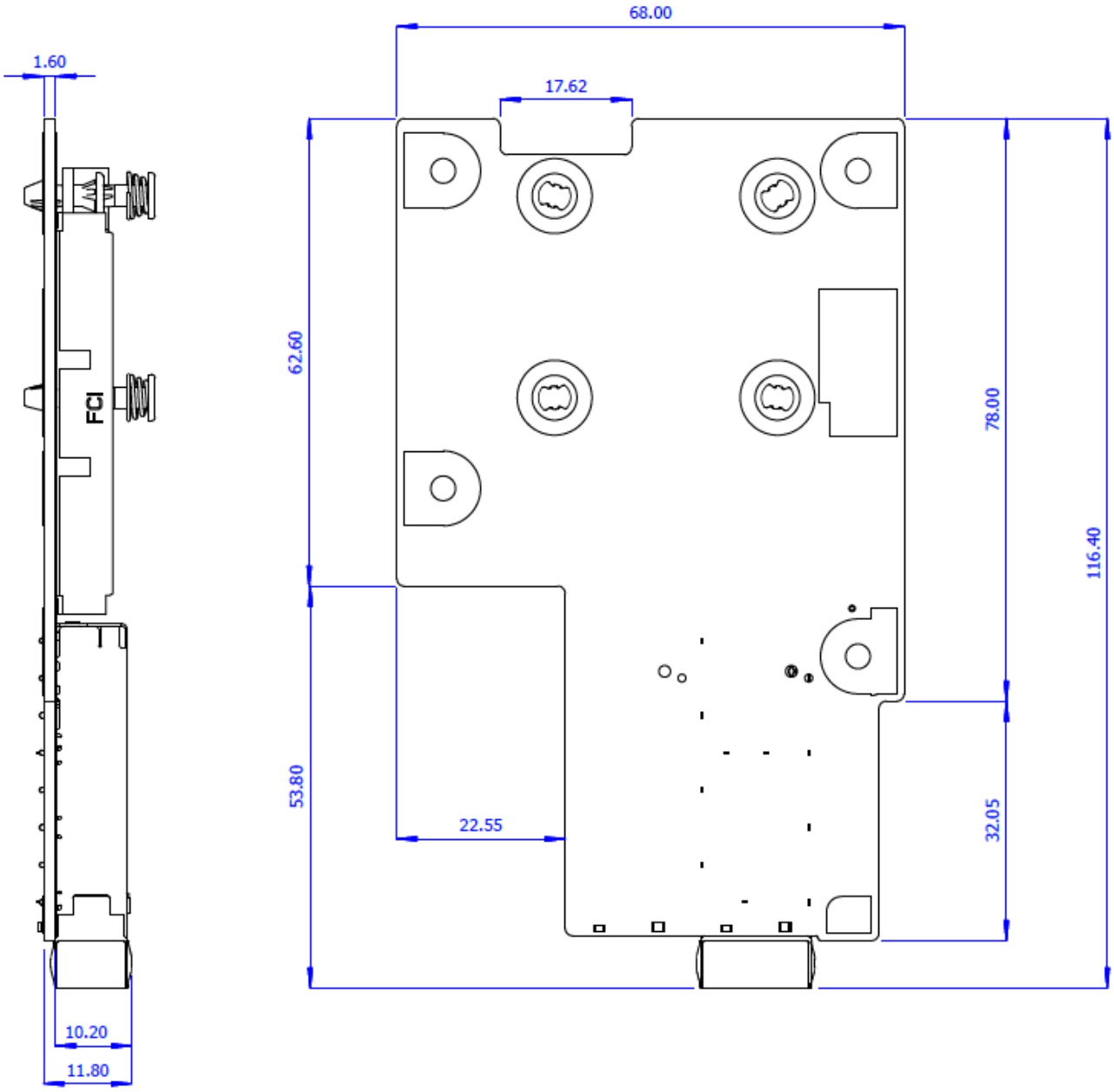


Figure 3: Single-Port Card Dimensions

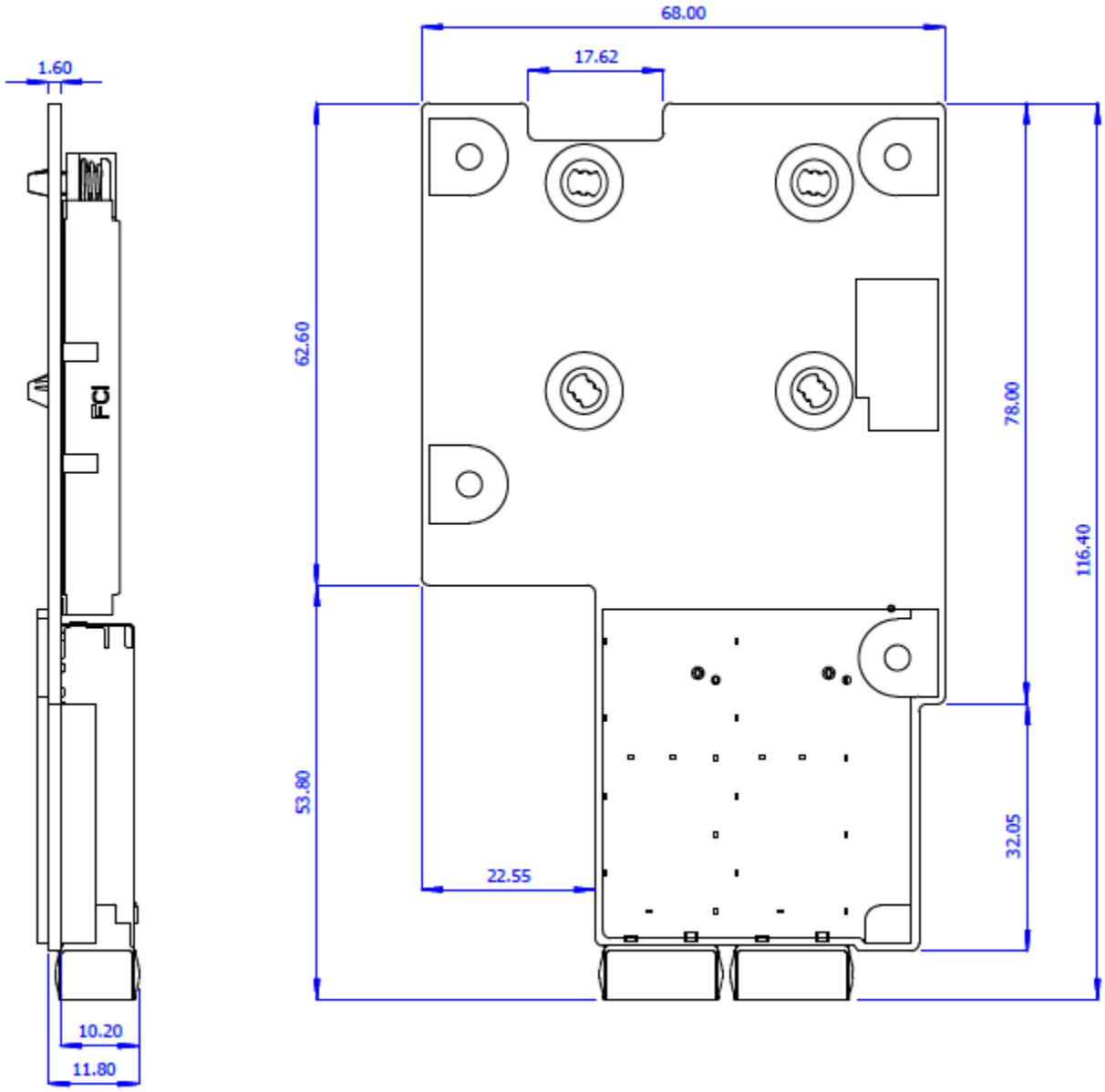


Figure 4: Dual-Port Card Dimensions

3.2 Major Components

- SFP28
- Heat sink
- LEDs for link indication
- DC-2-DC converters
- FCI connector to the baseboard
- 156Mhz Oscillator

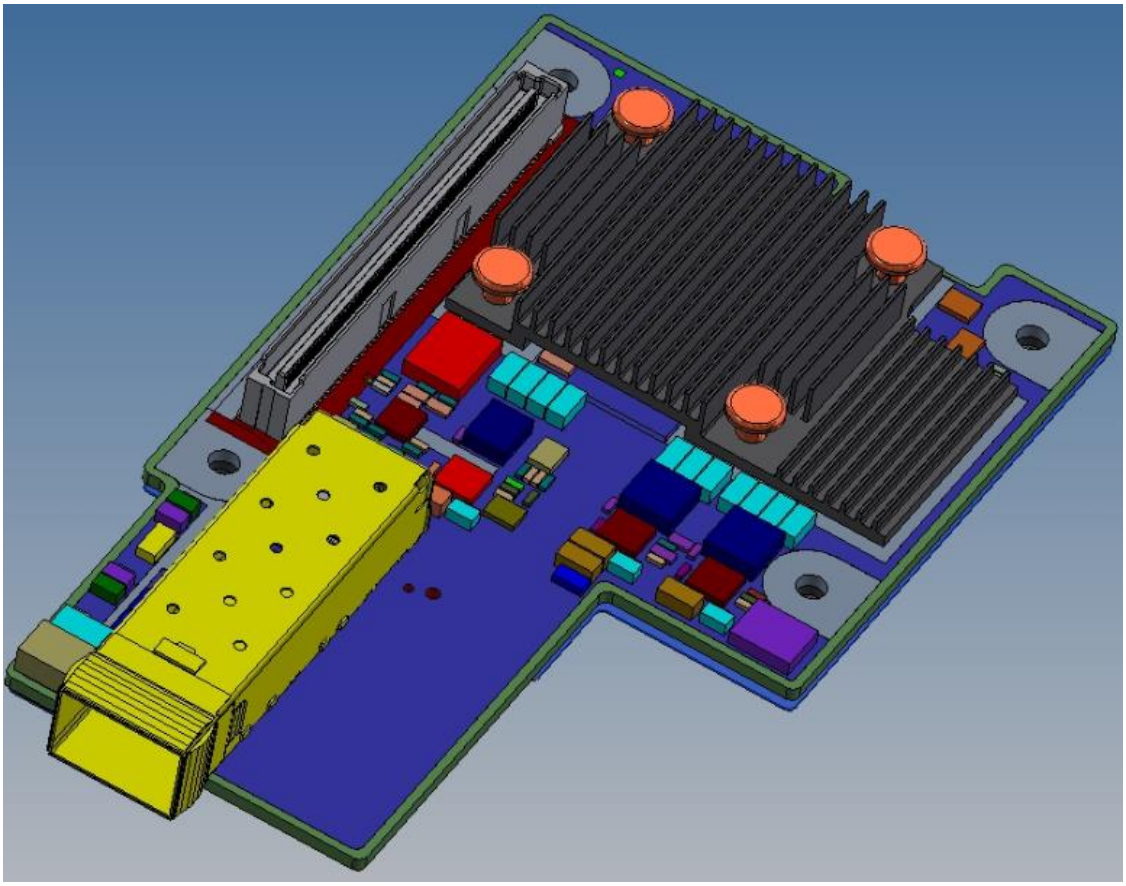


Figure 5: Single-Port Top View

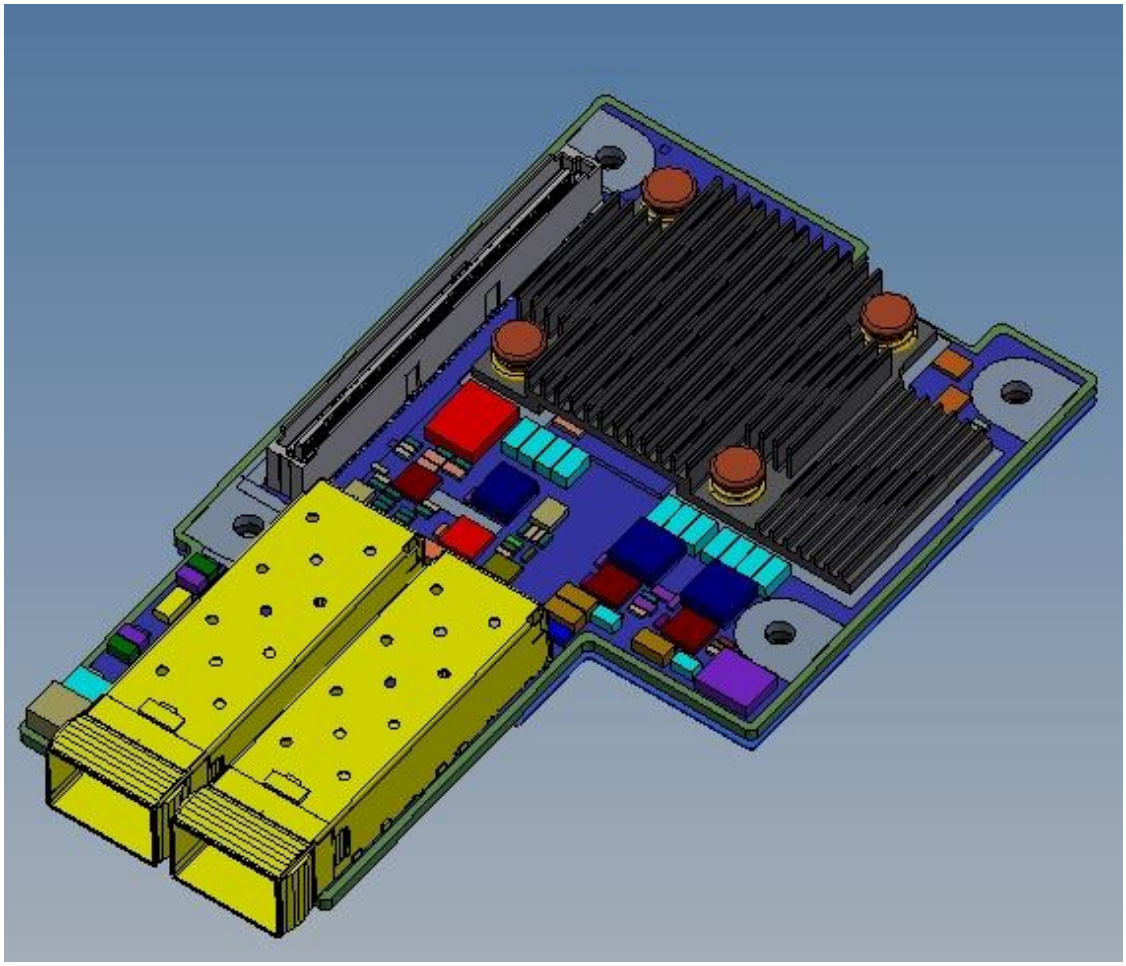


Figure 6: Dual-Port Top View

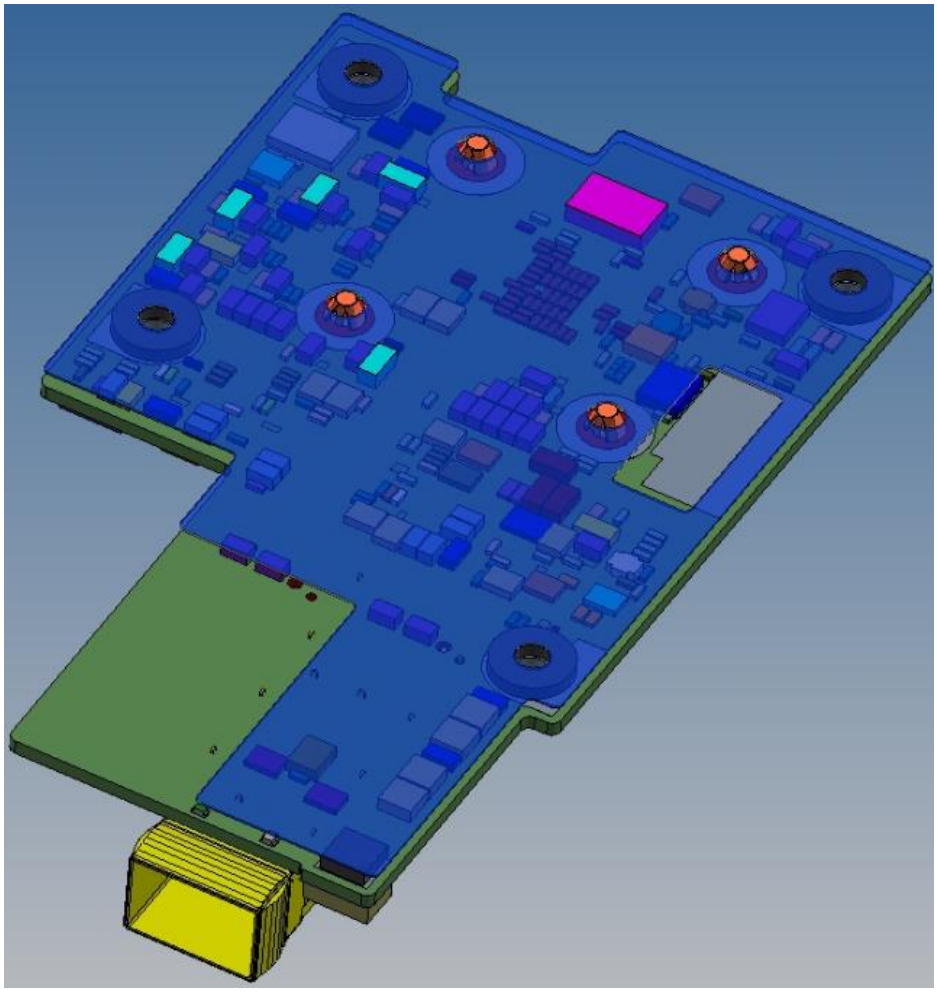


Figure 7: Single-Port Bottom View

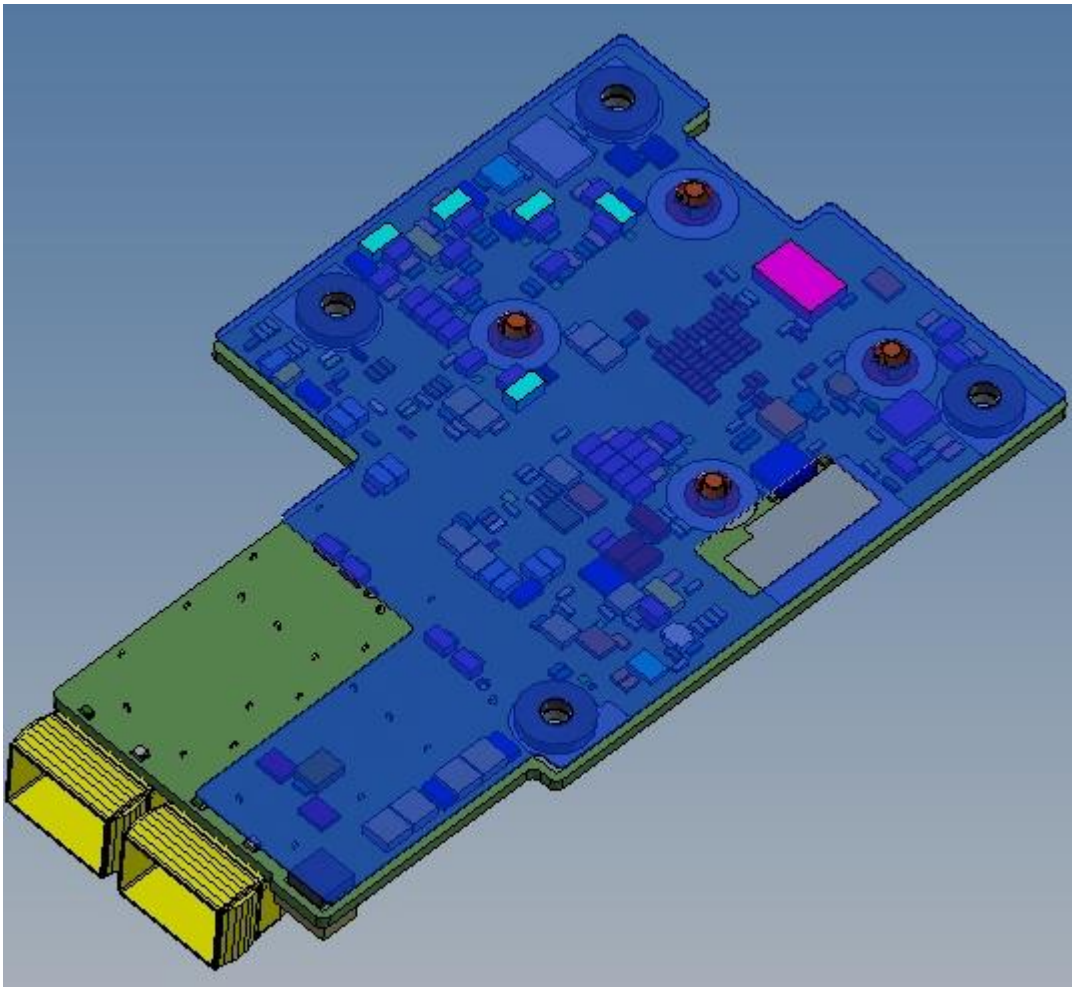


Figure 8: Dual-Port Bottom View

3.3 Connector

3.3.1 Cards following OCP Type1 Stacking Height

An FCI/61083-124402LF connector is mounted on the mezzanine card to mate with the connector mounted on the motherboard. The Connector is a 120 pin connector and provides x8 PCI-e lanes.

3.3.2 Cards following OCP Type4 Stacking Height

An FCI/61083-121402LF connector is mounted on the mezzanine card to mate with the connector mounted on the motherboard. The Connector is a 120 pin connector and provides x8 PCI-e lanes.

3.4 Pin Definition

The mezzanine card pin definition is as follows. The directions of the signals are from the perspective of the motherboard.

Note: Table 1 applies both to FCI/61083-124402LF and FCI/61083-121402LF connectors.

Table 1: Connector Pin Definition

Pin	Signal	Description	Pin	Signal	Description
A1	MEZZ_PRSNTA1_N/BASEBOARD_A_ID	Present pin1, short to pin120 on Mezz card	A61	P12V_AUX/P12V	AuxPower
A2	P5V_AUX	Aux Power	A62	P12V_AUX/P12V	AuxPower
A3	P5V_AUX	Aux Power	A63	P12V_AUX/P12V	AuxPower
A4	P5V_AUX	Aux Power	A64	GND	Ground
A5	GND	Ground	A65	GND	Ground
A6	GND	Ground	A66	P3V3_AUX	Aux Power
A7	P3V3_AUX	Aux Power	A67	GND	Ground
A8	GND	Ground	A68	GND	Ground
A9	GND	Ground	A69	P3V3	Power (Not Connected)
A10	P3V3	Power (Not Connected)	A70	P3V3	Power (Not Connected)
A11	P3V3	Power (Not Connected)	A71	P3V3	Power (Not Connected)
A12	P3V3	Power (Not Connected)	A72	P3V3	Power (Not Connected)
A13	P3V3	Power (Not Connected)	A73	GND	Ground
A14	NCSI_CRSDV	SE_NCSI_CRS_DV_CON (Carrier Sense/Receive Data Valid)	A74	LAN_3V3STB_ALERT_N	SMBus Alert
A15	NCSI_RCLK	SE_NCSI_REF_CLK (NCSI CLK)	A75	SMB_LAN_3V3STB_CLK	SMBus Clock

A16	NCSI_TXEN	SE_NCSI_TX_EN (Transmit Enable)	A76	SMB_LAN_3V3STB_DAT	SMBus Data
A17	RSVD	S_PERST0_L	A77	RSVD	PCIE wake up
A18	MEZZ_SMCLK	I2C SMBus Clock	A78	NCSI_RXER	SE_NCSI_RX_ERR_CON(Receive error, not supported)
A19	MEZZ_SMDATA	I2C SMBus Data	A79	GND	Ground
A20	GND	Ground	A80	NCSI_TXD0	SE_NCSI_TX_D0 (Transmit Data)
A21	GND	Ground	A81	NCSI_TXD1	SE_NCSI_TX_D1 (Transmit Data)
A22	NCSI_RXD0	SE_NCSI_RX_D0_CON (Receive Data)	A82	GND	Ground
A23	NCSI_RXD1	SE_NCSI_RX_D1_CON (Receive Data)	A83	GND	Ground
A24	GND	Ground	A84	RSVD	2nd set of 100MHz PCIE Clock
A25	GND	Ground	A85	RSVD	2nd set of 100MHz PCIE Clock
A26	RSVD	set of 100MHz PCIE Clock (Not Connected)	A86	GND	Ground
A27	RSVD	set of 100MHz PCIE Clock (Not Connected)	A87	GND	Ground
A28	GND	Ground	A88	KR_TX_DP<8>	PCIE TX signal
A29	GND	Ground	A89	KR_TX_DN_<8>	PCIE TX signal
A30	KR_RX_DP<8>	PCIE RX signal	A90	GND	Ground
A31	KR_RX_DN<8>	PCIE RX signal	A91	GND	Ground
A32	GND	Ground	A92	KR_TX_DP<9>	PCIE TX signal

A33	GND	Ground	A93	KR_TX_DN_<9>	PCIE TX signal
A34	KR_RX_DP<9>	PCIE RX signal	A94	GND	Ground
A35	KR_RX_DN<9>	PCIE RX signal	A95	GND	Ground
A36	GND	Ground	A96	KR_TX_DP<10>	PCIE TX signal
A37	GND	Ground	A97	KR_TX_DN_<10>	PCIE TX signal
A38	KR_RX_DP<10>	PCIE RX signal	A98	GND	Ground
A39	KR_RX_DN<10>	PCIE RX signal	A99	GND	Ground
A40	GND	Ground	A100	KR_TX_DP<11>	PCIE TX signal
A41	GND	Ground	A101	KR_TX_DN_<11>	PCIE TX signal
A42	KR_RX_DP<11>	PCIE RX signal	A102	GND	Ground
A43	KR_RX_DN<11>	PCIE RX signal	A103	GND	Ground
A44	GND	Ground	A104	KR_TX_DP<12>	PCIE TX signal
A45	GND	Ground	A105	KR_TX_DN_<12>	PCIE TX signal
A46	KR_RX_DP<12>	PCIE RX signal	A106	GND	Ground
A47	KR_RX_DN<12>	PCIE RX signal	A107	GND	Ground
A48	GND	Ground	A108	KR_TX_DP<13>	PCIE TX signal
A49	GND	Ground	A109	KR_TX_DN_<13>	PCIE TX signal

A50	KR_RX_DP<13>	PCIE RX signal	A110	GND	Ground
A51	KR_RX_DN<13>	PCIE RX signal	A111	GND	Ground
A52	GND	Ground	A112	KR_TX_DP<14>	PCIE TX signal
A53	GND	Ground	A113	KR_TX_DN_<14>	PCIE TX signal
A54	KR_RX_DP<14>	PCIE RX signal	A114	GND	Ground
A55	KR_RX_DN<14>	PCIE RX signal	A115	GND	Ground
A56	GND	Ground	A116	KR_TX_DP<15>	PCIE TX signal
A57	GND	Ground	A117	KR_TX_DN_<15>	PCIE TX signal
A58	KR_RX_DP<15>	PCIE RX signal	A118	GND	Ground
A59	KR_RX_DN<15>	PCIE RX signal	A119	GND	Ground
A60	GND	Ground	A120	MEZZ_PRSNT2_N	present pin2, short to pin1 on mezz card

3.5 Power Capability and Status on Connector

Baseboard supplies power to power pins on Mezzanine card connectors. The current capability and power status is as provided Table 2. Normal power is available on state S0 only. Auxiliary power is available at all power states including hibernate state S4 or off state S5.

Note: 25GbE is fed from 5VAUX and 3.3VAUX power rails only.

Table 2: Power Pins for Connector

Power R ail	Voltage Tolerance	# Pins	Current Capability	Status
P12V_AUX/P 12V	±8%(max)	3	2.4A	Normal power or 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

3.6 Installation in Chassis

3D View of shows the 3D view of the mezzanine card installed in an Open Compute chassis.

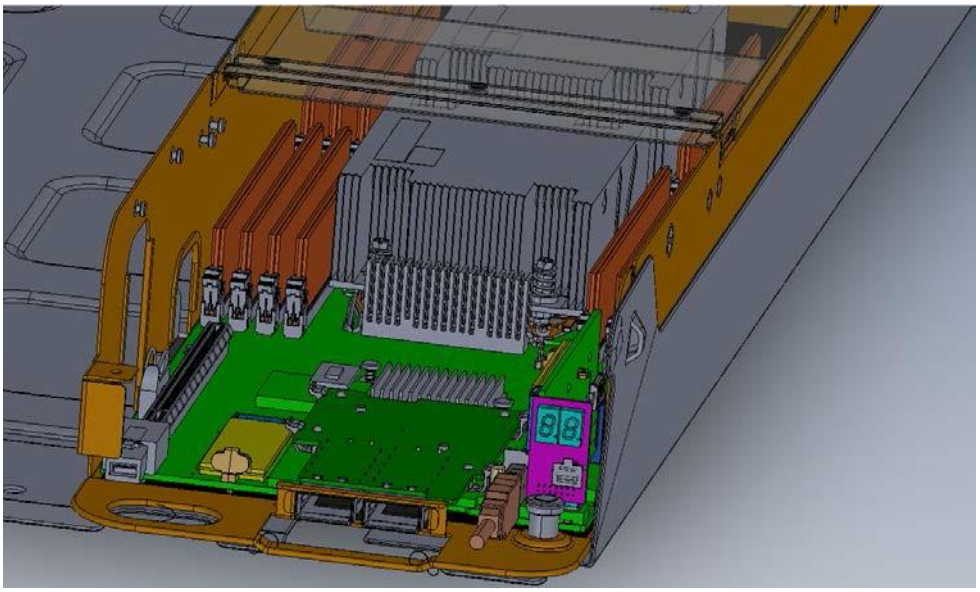


Figure 99: 3D View of Mezzanine Card installed in OCP Chassis – Example

4 Interfaces

4.1 Ethernet Interface

The network ports of the ConnectX®-4 Lx adapter card are compliant with the IEEE 802.3 Ethernet standards. Ethernet traffic is transmitted through the cards' SFP28 connectors.

4.2 PCI Express Interface

The ConnectX®-4 Lx EN network interface cards support PCI Express 3.0 (1.1 and 2.0 compatible). The device can be either a master initiating the PCI Express bus operations or a slave responding to PCI bus operations. The following lists the PCIe interface features:

- PCIe Gen 3.0, compatible with 2.0 and 1.1
- 2.5GT/s, 5.0 GT/s, or 8.0GT/s link rate x8
- Auto-negotiates to x8, x4, or x1

4.3 LED Interface

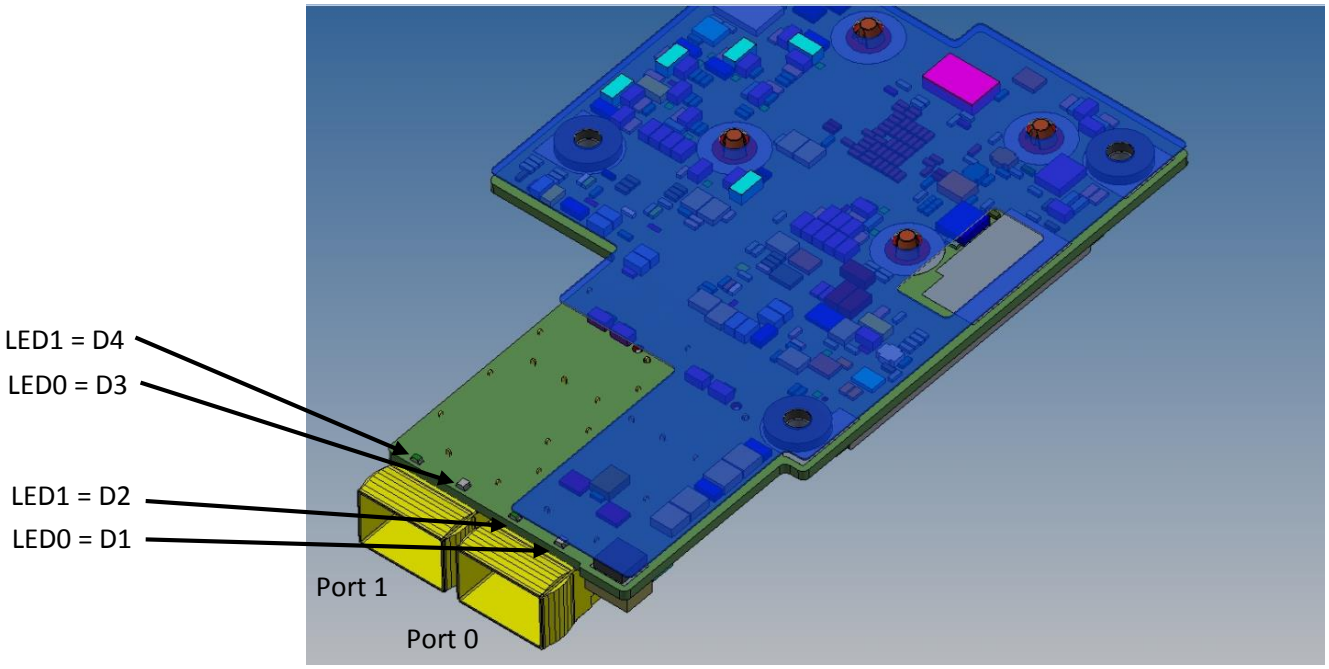


Figure 10: Mezzanine Card LED Location

There are two I/O LED per port to indicate link status and speed. See Table 3.

Table 3: Physical and Logical Link Indications

LED	Function	LED Signals
LED0 - Physical link speed	• A constant Green indicates physical link with the highest rated speed. • A constant Yellow indicates physical link with degraded speed. • If LED0 is off, then the physical link has not been established.	Port 0: LED0=D1, LED1=D2 Port 1: LED0=D3, LED1=D4
LED1 - Logical link/activity	• A constant Green indicates a valid logical (data activity) link without data transfer. • A blinking Green indicates a valid logical link with data transfer. • If LED1 is off, then the logical link has not been established.	

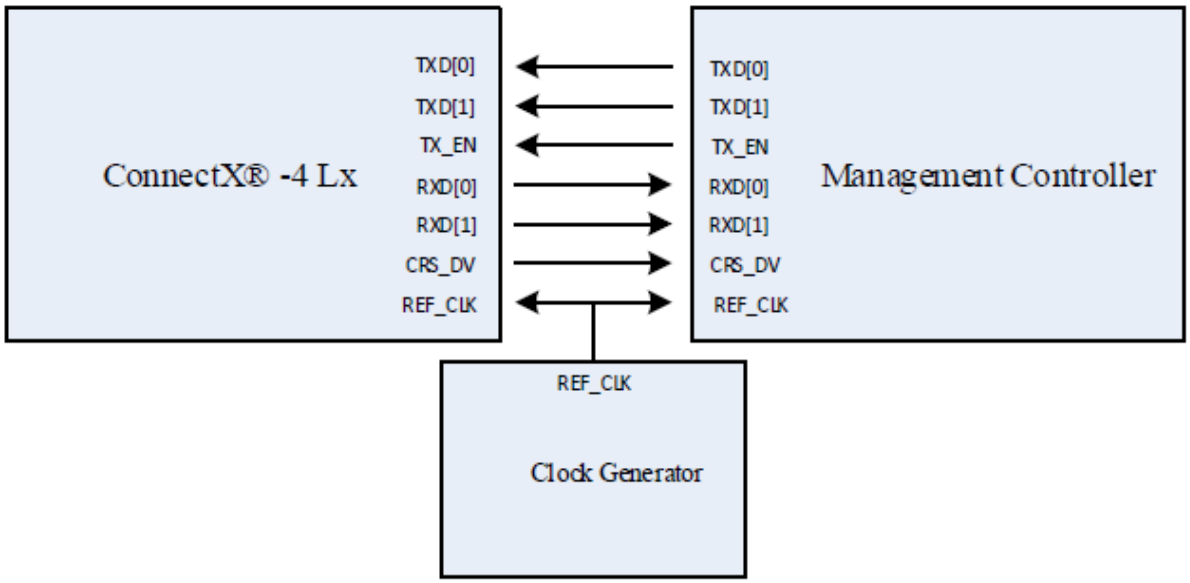
4.4 Management Interfaces

4.4.1 Network Controller Sideband Interface (NC-SI)

Note: Applies only to cards that support Host-Management.

The adapter supports a slave Network Controller Sideband Interface (NC-SI) that can be connected to a BMC. The adapter's NC-SI implementation supports all mandatory NC-SI commands specified in the Network Controller Sideband Interface (NC-SI) Specification, Rev. 1.1.0.

The traffic between the BMC and the network goes through a dedicated data-path, and is agnostic to the host operating state. Additionally, through special configuration, local communication between the BMC and the host can be enabled/disabled. When local-loopback is enabled, additional configuration defines if the traffic between the local host and the BMC is visible on the network port or not.



5 Specifications

5.1 Single-Port Card Specifications

Table 11: Single-Port Card Specifications Table

Protocol Support	**Ethernet:** 25GBASE-R, 20GBASE-KR2, 1000BASE-CX, 1000BASE-KX, 10GBASE-SR, 10GBASE-LR,10GBASE-ER, 10GBASE-CX4, 10GBASE-KX4, 10GBASE-CR, 10GBASE-KR, SGMII
	Data Rate: 1/10/25 Gb/s– Ethernet
	PCI Express Gen-3: SERDES @ 8.0GT/s, 8 lanes (2.0 and 1.1 compatible)
Power and Environmental	**Voltage:** 5VAUX, 3.3VAUX
	Typical Power[1]: Passive Cables – 8.41W
	Maximum Power: Passive Cables – 10.05W 1.5 Active Cables – 11.55W
	Maximum power available through SFP28 port: 1.5W
	Temperature: Operational: 0°C to 35°C Non-operational: -40°C to 70°C
	Humidity: 90% relative humidity[2]
	Air Flow[3]: Passive/Active Cables - 200LFM (port to heat sink)

5.2 Dual-Port Card Specifications

[1] Typical power for ATIS traffic load.

[2] For both operational and non-operational states

[3] Simulated inside air tunnel 15.2 mm high and 73 mm wide. Airflow direction - SFP28 to ConnectX-4 Lx.

Table 5: Dual-Port Card Specifications Table

Protocol Support	**Ethernet:** 25GBASE-R, 20GBASE-KR2, 1000BASE-CX, 1000BASE-KX, 10GBASE-SR, 10GBASE-LR,10GBASE-ER, 10GBASE-CX4, 10GBASE-KX4, 10GBASE-CR, 10GBASE-KR, SGMII
	Data Rate: 1/10/25 Gb/s– Ethernet
	PCI Express Gen-3: SERDES @ 8.0GT/s, 8 lanes (2.0 and 1.1 compatible)
Power and Environmental	**Voltage:** 5VAUX, 3.3VAUX
	Typical Power[4]: Passive Cables - 9.01W
	Maximum Power: Passive Cables - 10.83W 1.5 Active Cables – 13.83W
	Maximum power available through SFP28 port: 1.5W
	Temperature: Operational: 0°C to 35°C Non-operational: -40°C to 70°C
	Humidity: 90% relative humidity[5]
	Air Flow[6]: Passive/Active Cables - 200LFM (port to heat sink)

5.3 MAC Address Label Requirements

5.3.1 Placement Rules

MAC address label(s) must be scannable when 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 12Figure 12.

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.

[4] Typical power for ATIS traffic load.

[5] For both operational and non-operational states.

[6] Simulated inside air tunnel 15.2 mm high and 73 mm wide. Airflow direction - SFP28 to ConnectX-4 Lx

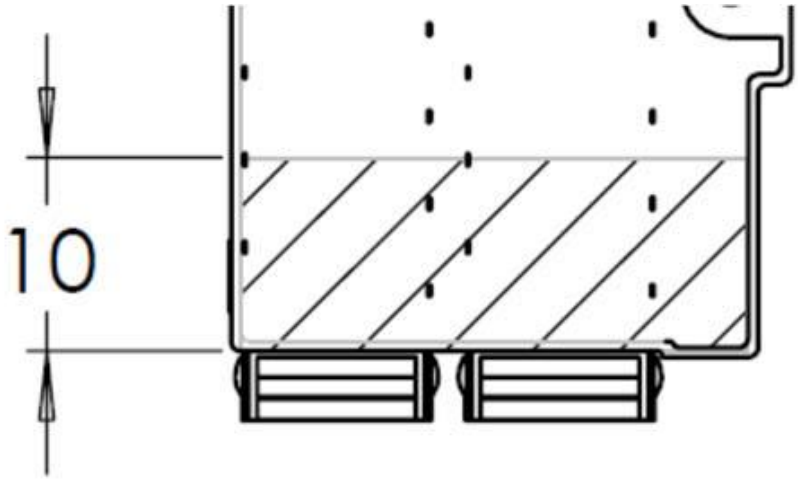


Figure 121: Mac Address Label Placement

The scanned bar code should not include ".". Example: "AA.BB.CC.00.11.20" should scan as "AABBCC001120". Implementation example is shown in Figure 13.

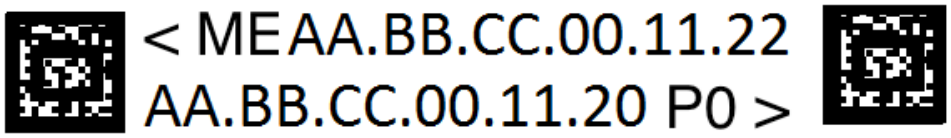


Figure 132: MAC Address Label Implementation Example

5.3.2 Barcode Requirements

A header is required for the label. For example " Po:hh.hh.hh.hh.hh.hh". The font size should be larger that 5 points. If there is more than one MAC address per unit, a human readable text header is required and must differentiate. For example "ME:hh.hh.hh.hh.hh.hh" or "Po:hh.hh.hh.hh.hh.hh" or "P1: hh.hh.hh.hh.hh.hh".

5.3.3 Label Permanence and Technology

The MAC address label must adhere for at least 3-4 years in Open Compute usage and thermal environment. Materials can be PCB silkscreen, polyester, and polyamide with acrylic adhesive.

5.3.4 Label Size

.

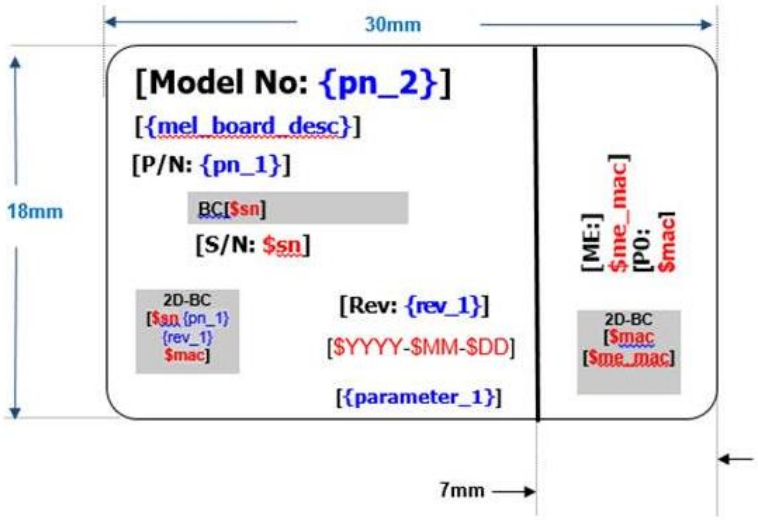


Figure 14: Single-Port Board Label Size

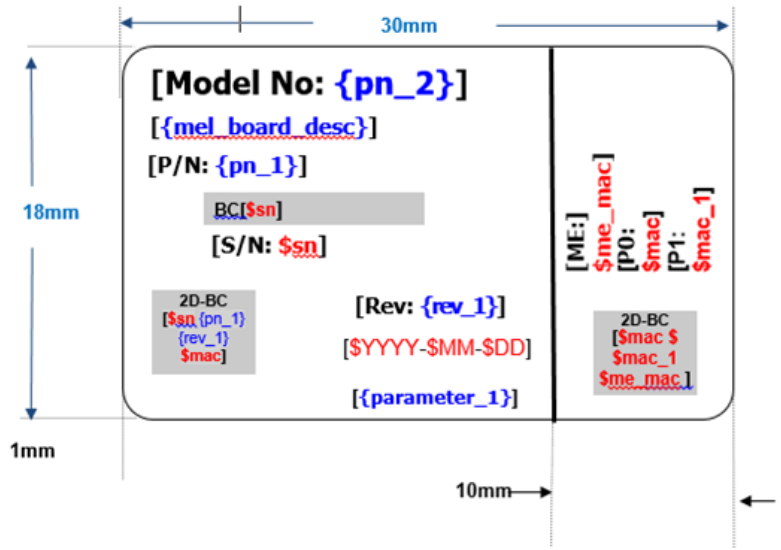


Figure 15: Dual-Port Board Labe Size

6 Environmental

6.1 Environmental Requirements

This Mezzanine card shall meet the same requirements specified in the OCP systems that the Mezzanine cards is in. The OCP system that uses OCP Mezzanine card shall define air flow direction, inlet air temperature, air flow (or speed) to the local area where Mezzanine card is at, and simulation boundary.

Please refer to Section 6: Specifications for Airflow direction, inlet air temperature and air flow (or speed) to the local area where Mezzanine card is at.

6.2 Regulation

This mezzanine card meets KCC, CE, CTUV-US, RCM, ROHS, CB and FCC class A.

7 Revision History

Rev 1.0 – March 2016

- First release