Open Accelerator Infrastructure (OAI) - OCP Accelerator Module (OAM) Base Specification r2.0 v0.75

Author: OCP OAI Workstreams

Table of Contents

1 License

Contributions to this Specification are made under the terms and conditions outlined in Open Web Foundation Contributor License Agreement ("OWF CLA 1.0") ("Contribution License") by:

AMD, Baidu, Broadcom, Graphcore, H3C, Inspur, Intel, Intel Habana, Luxshare, Meta, Microchip, Molex, Nallasway, Supermicro, TE, Wiwynn, ZT Systems

Usage of this Specification is governed by the terms and conditions outlined in the Open Web Foundation Final Specification Agreement ("OWFa 1.0").

Note: The following clarifications, which distinguish technology licensed in the Contribution License and/or Specification License from those technologies merely referenced (but not licensed), were accepted by the Incubation Committee of the OCP:

None.

NOTWITHSTANDING THE FOREGOING LICENSES, THIS SPECIFICATION IS PROVIDED BY OCP "AS-IS", AND OCP EXPRESSLY DISCLAIMS ANY WARRANTIES (EXPRESS, IMPLIED, OR OTHERWISE), INCLUDING IMPLIED WARRANTIES OF MERCHANTABILITY NON-INFRINGEMENT, FITNESS FOR A PARTICULAR PURPOSE, OR TITLE, RELATED TO THE SPECIFICATION. NOTICE IS HEREBY GIVEN THAT OTHER RIGHTS NOT GRANTED AS SET FORTH ABOVE, INCLUDING WITHOUT LIMITATION, RIGHTS OF THIRD PARTIES WHO DID NOT EXECUTE THE ABOVE LICENSES, MAY BE IMPLICATED BY THE IMPLEMENTATION OF OR COMPLIANCE WITH THIS SPECIFICATION. OCP IS NOT RESPONSIBLE FOR IDENTIFYING RIGHTS FOR WHICH A LICENSE MAY BE REQUIRED TO IMPLEMENT THIS SPECIFICATION. THE ENTIRE RISK AS TO IMPLEMENTING OR OTHERWISE USING THE SPECIFICATION IS ASSUMED BY YOU. IN NO EVENT WILL OCP BE LIABLE TO YOU FOR ANY MONETARY DAMAGES WITH RESPECT TO ANY CLAIMS RELATED TO, OR ARISING OUT OF YOUR USE OF THIS SPECIFICATION, INCLUDING BUT NOT LIMITED TO ANY LIABILITY FOR LOST PROFITS OR ANY CONSEQUENTIAL, INCIDENTAL, INDIRECT, SPECIAL OR PUNITIVE DAMAGES OF ANY CHARACTER FROM ANY CAUSES OF ACTION OF ANY KIND WITH RESPECT TO THIS SPECIFICATION, WHETHER BASED ON BREACH OF CONTRACT, TORT (INCLUDING NEGLIGENCE), OR OTHERWISE, AND EVEN IF OCP HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

2 Compliance with OCP Tenets

2.1 Openness

OAM Base Specification is an open standard with joint contribution across Hyperscalers, ODMs, component suppliers, and connector vendors.

2.2 Impact

OAM Base Specification defines common form factor for AI/HPC accelerator modules and greatly cut down design cycle time and resources for companies adopting the standard.

OAM has been widely adopted by Hyperscalers, CSPs, and OEMs.

2.3 Scale

OAM Base Specification defines GPU/ASIC power consumption up to 1000W with form factor and connector selection to support future advanced ASIC package technology and power delivery solution.

2.4 Sustainability

OAM Base Specification defines liquid cooling implementation to improve data center facility PUE and WUE.

OAM Base Specification defines power delivery solution with dual power entries and multiple stack height connector to improve power delivery efficiency for AI/HPC GPU or ASIC at 1KW TDP.

3 Acknowledgements

We want to acknowledge all the OCP OAI Workstream members for their contributions to this specification: The incredible collaboration between customers, accelerator manufacturers, system developers, and industry partners shows how Open Compute develops industry-standard form factors and specifications that benefit all its members.

We would especially like to thank **AMD, Boyd, CoolerMaster, H3C, Intel, Intel Habana, Meta, Molex, and Wiwynn** for their extra efforts in putting this specification together.

4 Overview

Artificial Intelligence (AI) applications are rapidly evolving and producing an explosion of new types of hardware accelerators for Machine Learning (ML), Deep Learning (DL), and High-performance Computing (HPC).

Different implementations target similar requirements for power/cooling, robustness, serviceability, configuration, programming, management, debug, inter-module communication to scale-up, and input/output bandwidth to scale out.

To take advantage of the available industry-standard form factors to reduce the required time and effort in producing suitable solutions, various implementations have selected the PCIe CEM form factor as a quick market entry.

Such solutions are not optimized for the upcoming AI workloads, which require ever-growing bandwidth and interconnect flexibility for data/model parallelism.

The state-of-the-art applications require multiple cards in a system with various inter-card links running at highspeed interconnect bandwidth between cards.

Using the PCIe CEM form factor to meet such interconnect requirements poses several challenges. These include excessive signal insertion loss from ASIC to PCIe connectors and baseboard, inter-card cabling complexity reducing robustness and serviceability, and limiting the supported inter-ASIC topologies.

This base specification outlines an interoperable, modular hierarchy based on the OAM form factor (OCP Accelerator Module), an interconnect Baseboard, a Tray, and a Chassis. It enables flexible high-speed interconnect topologies for multi-ASIC solutions,

- OAM (various accelerators)
- Universal Baseboard (interconnecting topologies between accelerators, hosts, and other IO devices for scale up and scale out)
- OAI Tray (a means for ease of field replacement and serviceability)
- OAI Chassis (an outline for a collection of Trays and input/output resources to scale out)

Based on this base specification, various design and product implementations may maintain interoperability while offering enhancements in each hierarchy level.

We invite open contributions in the following areas:

1. Base specification (OCP Accelerator Infrastructure Sub-Project Specification)
2. Design specification (This document, detailed description of alternative, interoperable components which meet the base specification)
3. Products (schematic, layout, mechanical/thermal solutions, and firmware/software to realize the above designs)

4.1 Scope

The OAM base specification defines the form factor, standard specifications for a compute accelerator module, and a compliant baseboard design enabling interoperability across multiple ASIC or GPU-based mezzanine modules and baseboard design interfaces.

The OAM form factor facilitates scalability across accelerators by simplifying the system solution when interconnecting communication links among modules compared to a PCIe Add-in card form factor.

4.2 Acronyms

Acronym	Definition
ASIC	Application Specific Integrated Circuit
OAM	OCP Accelerator Module
BGA	Ball Grid Array
BMC	Baseboard Management Controller
TDP	Thermal Design Power
EDP	Excursion Design Power
GPU	Graphic Processing Unit
MPN	Manufacturing Part Number
DXF	Drawing eXchange Format
PCBA	Printed Circuit Board Assembly
UBB	Universal Baseboard
OAI	Open Accelerator Infrastructure

5 High-Level Specification for the OCP Acceleration Module

Module Dimension	102mm x 170mm
Module Power/Input Voltage	Supports up to 1000W, using 44V-59.5V DC as input power
Connectors	2* Molex Mirror Mezz Pro Connectors (MPN: 218910-1115 or 218916-1115) Stack height 5mm ± 0.15 or 8 mm ± 0.15 Differential pair Impedance: 90ohm ± 5%
Host Interface	One or two x16 host links. E.g., PCIe Gen5/6 x16, or alternate protocols.
Module to Module Interconnect Links	Up to 7 Links per module, each link has up to x16 lanes Each link may be able to be configured into sub-links.
Bottom stiffener height (including Mylar)	5.15mm (MAX) or 8.15mm (MAX)

6 OAM Mechanical Specifications

This section describes the OAM form factor. It uses a single accelerator ASIC on the module as an example to illustrate the mechanical specifications. The top and bottom stiffeners may be different if the modules have multiple accelerator ASICs.

Please refer to 2D DXF and 3D files for further details. 2D DXF and 3D files are in the contribution package, with relevant reference drawings to mechanical components. Please note that some OAM features call out as required, but others are for reference.

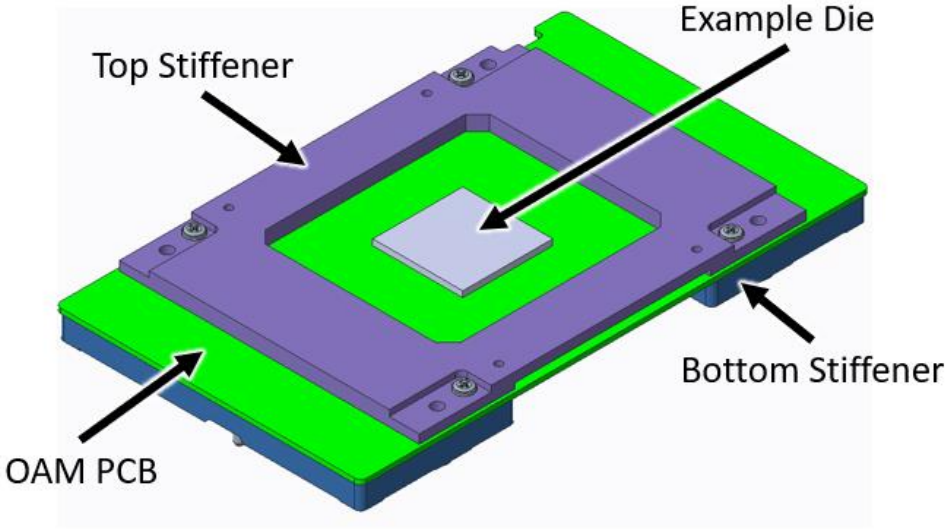


Figure 1 OAM isometric view

6.1 Module PCBA Form Factor

This section covers the required and recommended dimensions of the module PCBA and its parts. Figures 2 and 3 illustrate the OAM form factor and dimensions, with all the measurements in Figure 2 required. It is a 102mm x 170mm PCB size, Mezzanine Connectors on the bottom side, and the accelerator on the top side. The connector to connector pitch is 107mm. Four NPTH mounting holes attach the module to a corresponding bolster plate secured below the system PCB. These mounting holes should provide clearance for an M3.5 screw with enough thread length to secure the bottom stiffener. There is a notch located near the southwest corner of the board, adjacent to Connector 1. For connector orientation, see Figure 4 Top and bottom views of the OAM Assembly.

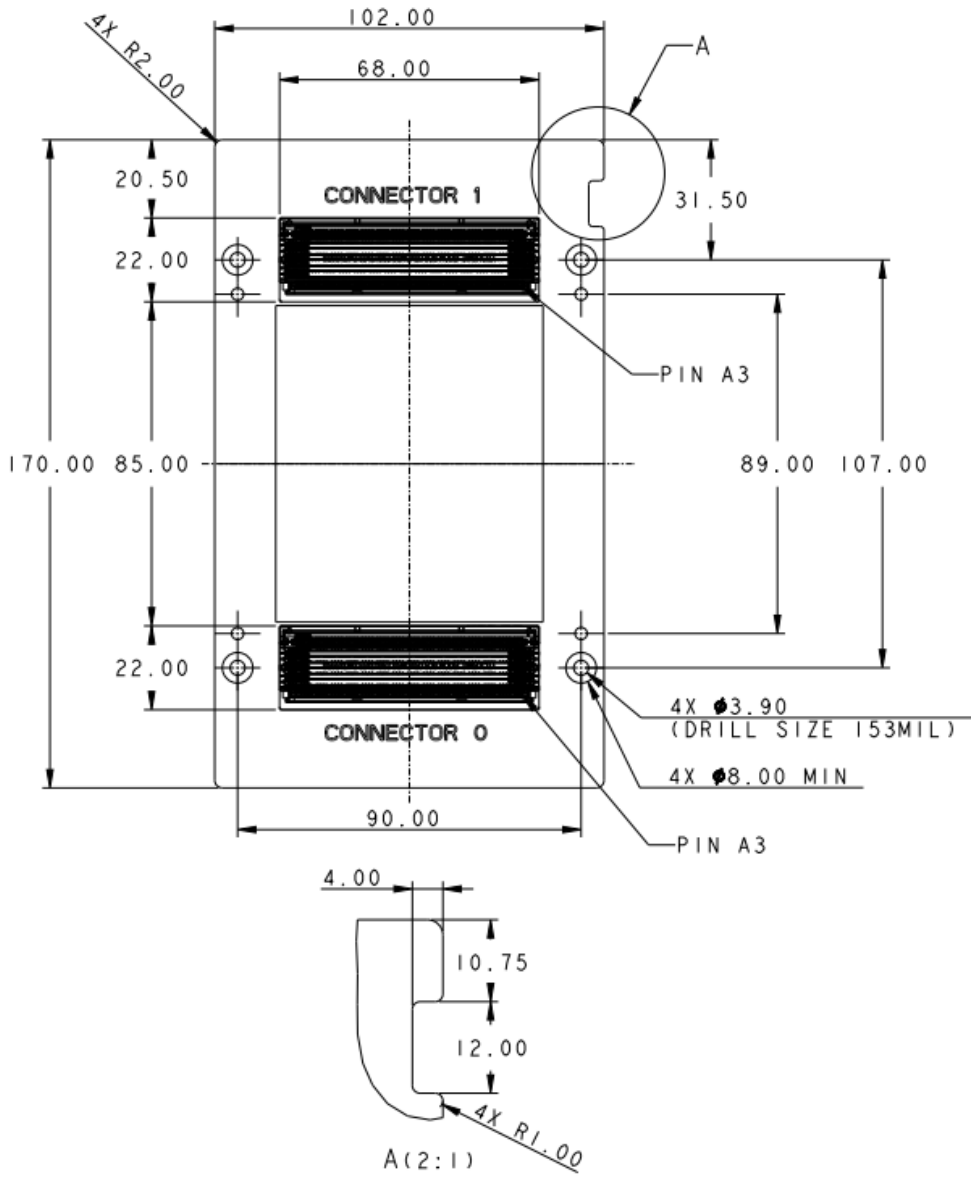


Figure 2 OAM Form Factor Dimensions, Bottom View

Figure 3 OAM Form Factor, Side View with System Baseboard

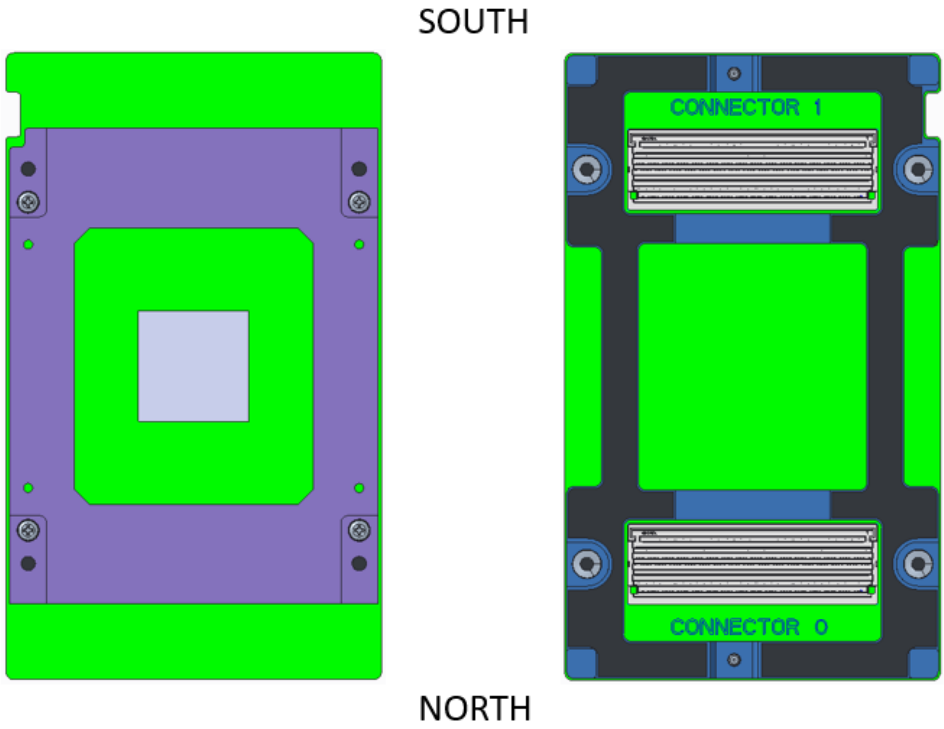


Figure 4 Top and bottom views of the OAM Assembly

6.2 Mezzanine Connector

There are two versions of Molex Mirror Mezz Pro, MPN 218910-1115 and 218916-1115. Molex Mirror Mezz Pro is the PCB to PCB interconnect solution supported by the OAM form factor. Mirror Mezz Pro is a highspeed differential pair-based mezzanine connector in a footprint-identical genderless plug and receptacle part for module and baseboard. Figure 5 Mirror Mezz Pro 218910-1115 and 218916-1115 are provided courtesy of Molex.

- Stack height: 5mm (218910-1115 engages with 218910-1115 on UBB).
- Mating Force: 0.50N/pin Max, total 344.0N MAX. Data on mate forces of the 218910-1115 connector is in Table 1.
- Unmating force: 0.045N/pin MIN, total 31.0N MIN. Data on the unmate force of the 218910-1115 connector is in Table 1.
- 172 Total Differential Pairs, of which 161 are fully ground shielded (non-orphan).

The specification of another version of Mirror Mezz Pro 218916-1115 are listed below.

- Stack height: 8mm (218916-1115 engages with 218910-1115 on UBB).
- Mating Force: 0.50N/pin Max, total 344.0N MAX. Data on mate forces of the 218916-1115 connector is in Table 1.
- Unmating force: 0.045N/pin MIN, total 31.0N MIN. Data on the unmate force of the 218916-1115 connector is in Table 1.
- 172 Total Differential Pairs, of which 161 are fully ground shielded (non-orphan).

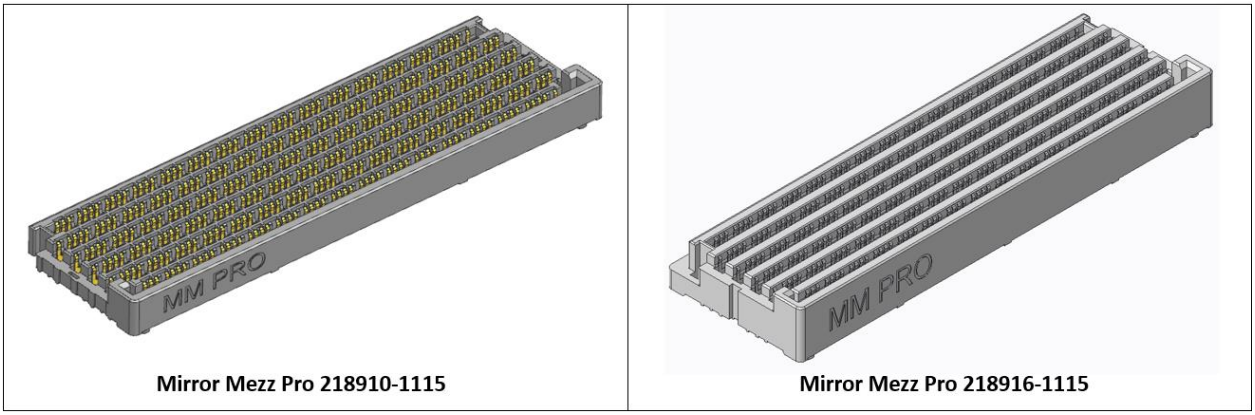


Figure 5 Mirror Mezz Pro 218910-1115 and 218916-1115

6.2.1 Mate/Unmate Force Data

The mating connectors will be vertically inverted when mated.

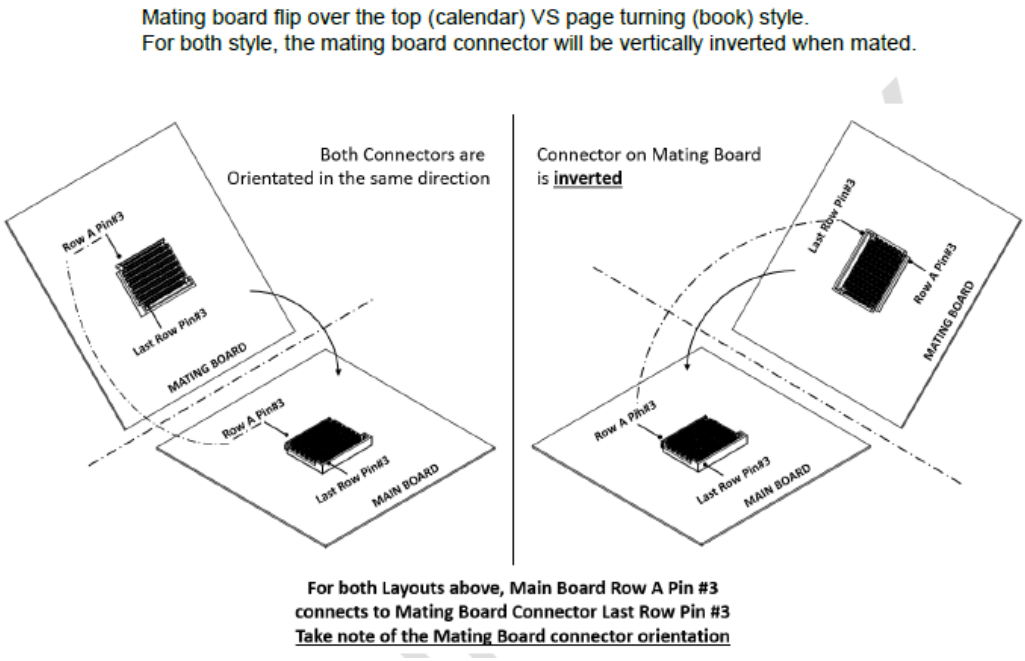


Figure 6 Mirror Mezz Pro Connector Mating

The mate and unmate forces provided in the product specification are conservative. The specific 218910-1115 and 218916-1115 connectors that the OAM uses has mate/unmate forces more in line with those found in Table 1 and in Figure 7. The mating force per pin trends upwards for the initial 5 cycles before settling back towards the average of 0.41N/pin.

Table 1 Mate/Unmate Averaged Data for both Mirror Mezz Pro 218910-1115 and 218916-1115

Mate and Un-mate Force (unit:N)

Unit:N	Cycle	1st	2nd	3rd	4th	5th	Max	Min	Ave
Mating Force	Sample1	267.5	279.4	292.8	300.4	309.5	309.5	250.9	280.7
	Sample2	253.5	263.1	278.6	286.9	298.5			
	Sample3	263.4	268.5	286.5	290.3	303.0			
	Sample4	266.1	275.7	292.6	294.4	294.9			
	Sample5	264.0	277.6	290.3	294.9	299.5			
	Sample6	260.2	262.8	279.5	277.4	277.9			
	Sample7	250.9	263.1	270.2	279.5	282.3			
	Sample8	263.2	264.4	273.5	287.7	289.5			
	Sample9	261.4	268.1	285.4	289.8	293.4			
	Sample10	274.0	285.5	294.5	300.8	306.4			
Un-mating Force	Sample1	151.0	161.5	173.6	184.5	185.4	185.6	151.0	176.1
	Sample2	162.8	176.3	181.1	181.4	180.9			
	Sample3	162.7	179.1	182.5	184.8	185.6			
	Sample4	159.7	176.6	182.1	182.6	184.9			
	Sample5	169.0	179.4	184.0	184.7	183.9			
	Sample6	159.8	170.1	172.5	173.6	176.3			
	Sample7	160.0	173.1	176.6	178.1	178.6			
	Sample8	161.8	177.0	177.3	180.5	181.6			
	Sample9	162.1	175.2	179.4	181.1	181.8			
	Sample10	168.3	182.3	185.3	185.3	184.8			

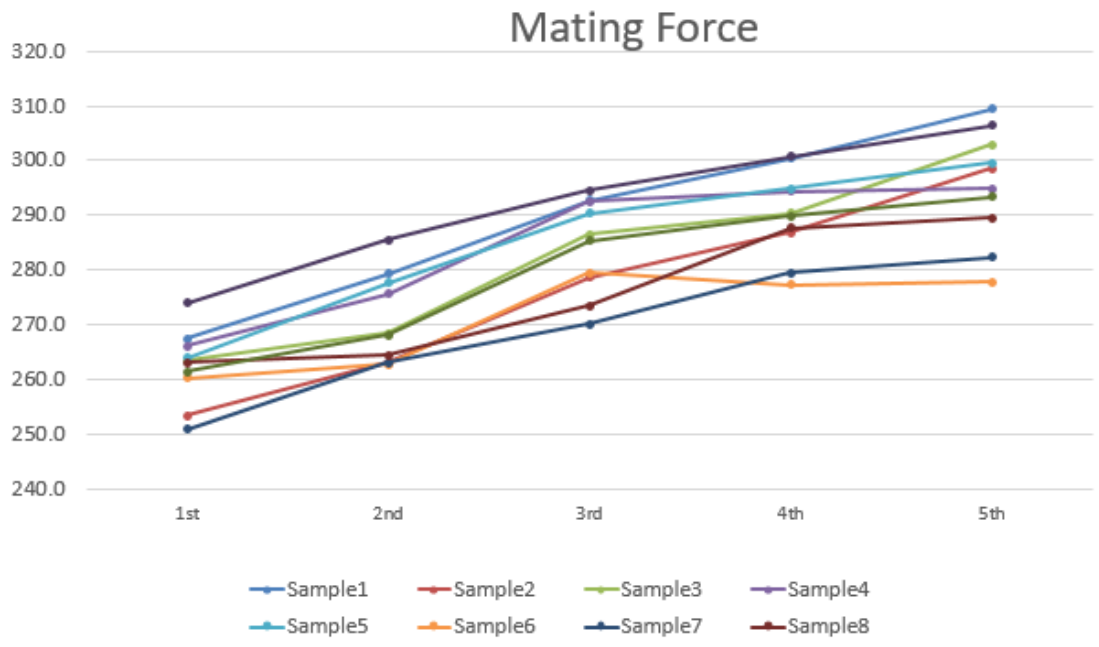


Figure 7 Measured Mate Force per Pin for both Mirror Mezz Pro 218910-1115 and 218916-1115

6.3 OAM Top Stiffener

The reference model of OAM top stiffener is for reference only, and dimensions may be changed or adjusted to accommodate the specific application and board layout of the OAM design.

6.4 OAM Bottom Stiffener

The reference model of OAM bottom stiffener in Figure 8 Bottom Stiffener Reference Design and dimensions are shown in Figure 9 Bottom Stiffener Dimensions. The bottom stiffener must accommodate the SMT nuts sizes shown in Figure 16 SMT Receiving Nut for Baseboard. Other features and dimensions of the bottom stiffener shown in the reference model are optional and can be adjusted based on the needs of the OAM design.

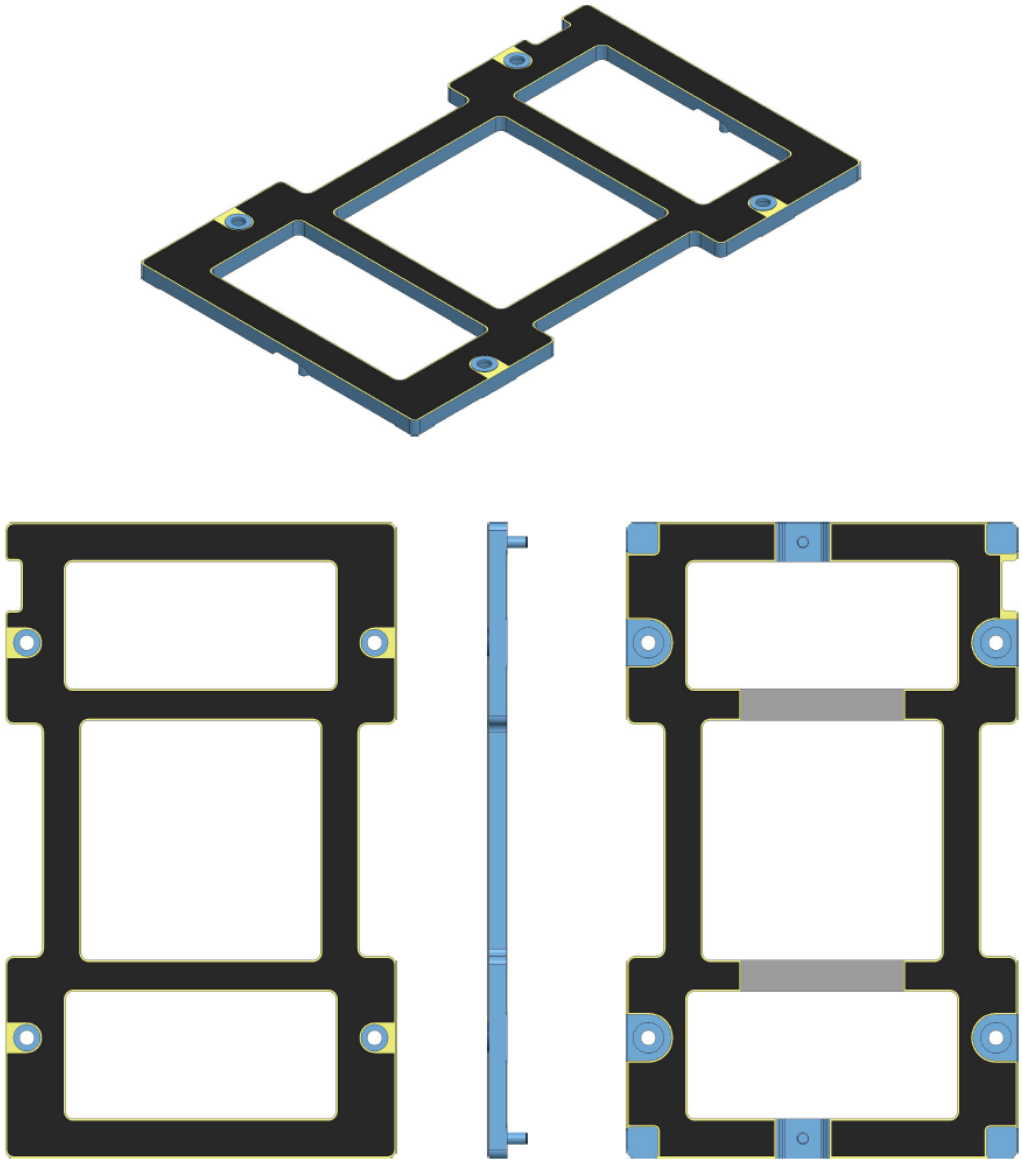


Figure 8 Bottom Stiffener Reference Design

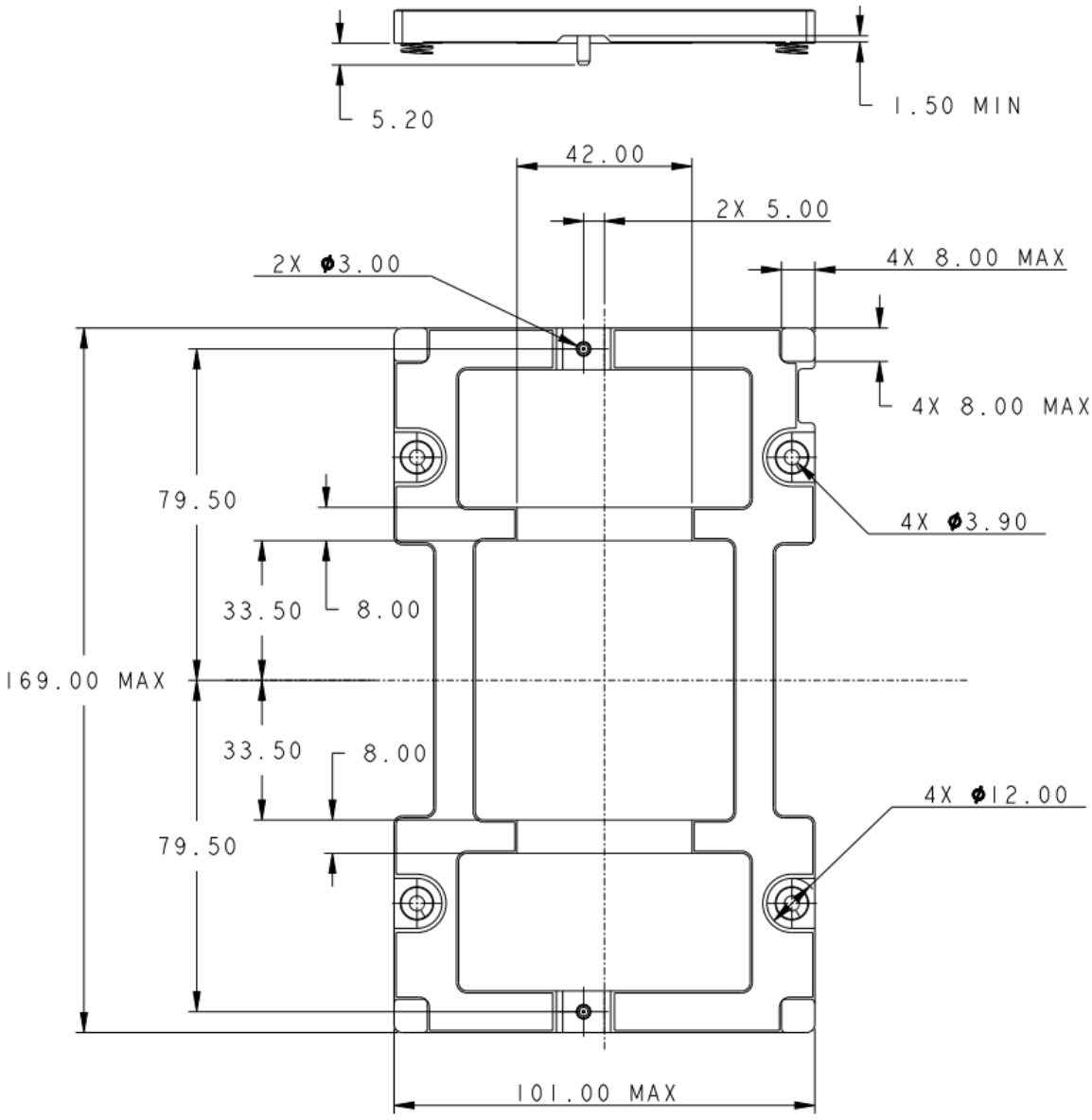


Figure 9 Bottom Stiffener Dimensions

6.4.1 Tolerance Stack-up of Bottom Stiffener

Standoff height as recommended by Molex for the Mirror Mezz Pro Connector is 5mm ± 0.15mm or 8mm ± 0.15mm, depending on the version of Mirror Mezz Pro. This tolerance may be difficult to attain using an insulator-adhesive-stiffener-adhesive-insulator stack, so it is highly recommended that pockets be machined into the stiffener to account for the tolerances of the insulator and adhesive (see reference design CAD for further details). With a stiffener only stack, 0.15mm should be easily attainable.

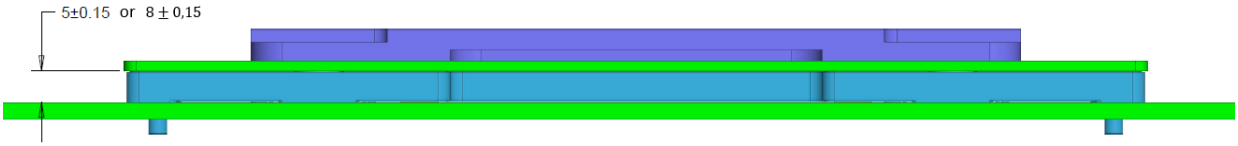


Figure 10 Tolerance Stack-up of Bottom Stiffener

6.4.2 Alignment Pins

There are two alignment pins required on the bottom stiffener of the OAM, intended as guidance features and an additional keying feature for the module (see Section 6.5 for more details). The pins are 3mm diameter. For stacking height 5mm, the alignment pin is with a length of 10mm measured from the bottom of the OAM PCB. For stacking height 8mm, the alignment pin is with a length of 13mm measured from the bottom of the OAM PCB. Note that since there may be components on the bottom side of the PCB, if the stiffener pocket in this area, the total length of the pin will be shorter. The minimum thickness of the stiffener is 1mm in these areas, as recommended. Figure 11 and 12 show examples of the possible alignment pins. Note that the length will vary depending on the specifically chosen geometry of the bottom stiffener.

MPN: PEM TPS-3mm-10 or equivalent

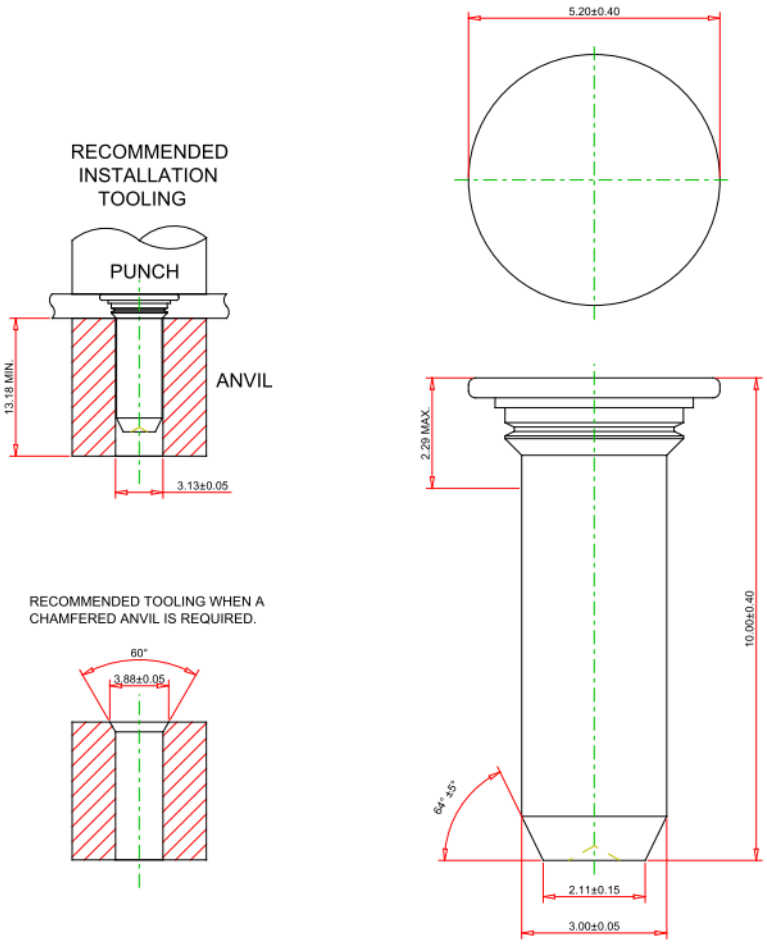


Figure 11 Alignment pin example of 5mm Stacking Height

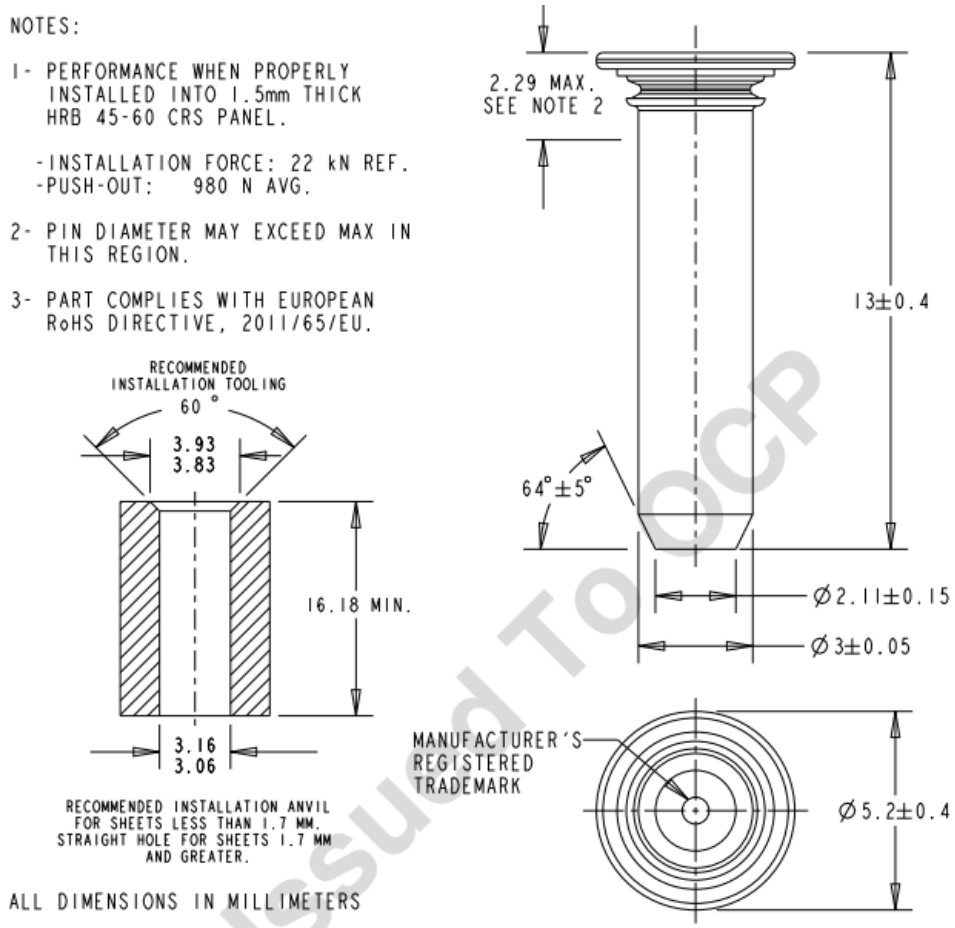


Figure 12 Alignment pin example of 8mm Stacking Height

6.4.3 EMI Gaskets / Pads

The bottom stiffener has two defined areas of 8x42mm size, reserved for the placement of fabric-over-foam gaskets. This area is designed to have a 0.5mm depth, and the gasket defined should have a 6x40mm footprint, with a 1mm height. It provides a 50% nominal compression and solid grounding to the baseboard (equivalently designed ground pad).

MPN: Laird 4Y03PC51H00158 or equivalent

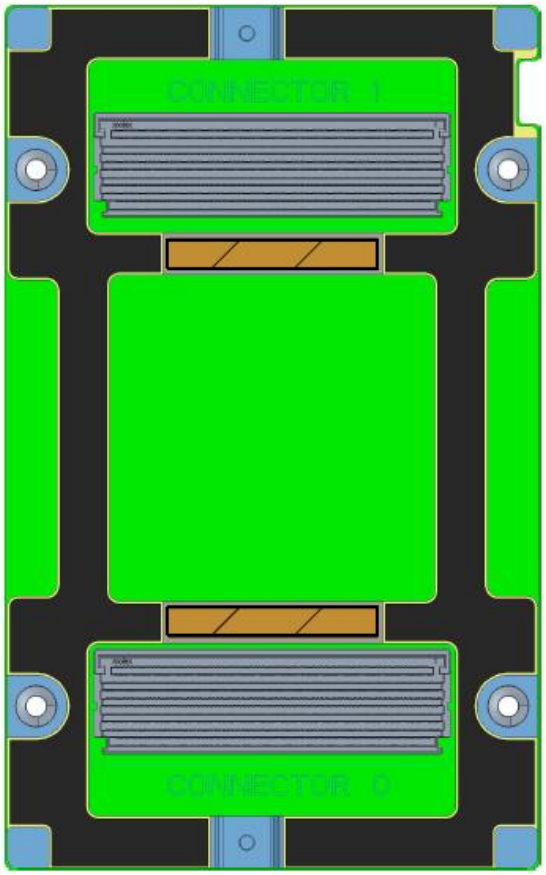


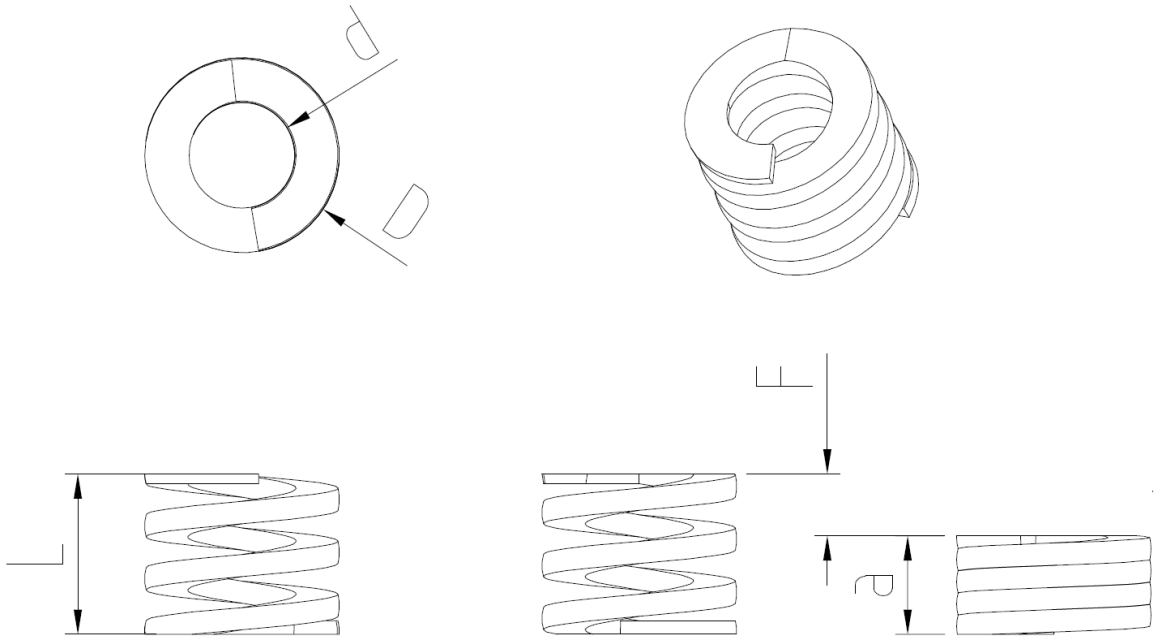
Figure 13 Fabric-over-foam Gasket Locations (brown)

6.4.4 Die Springs

Due to the mate and unmate forces of Mirror Mezz Pro connectors are high (see Section 6.2.1). Die springs are to be used to assist with the de-mate. It is strongly suggested for OAM vendors to use this reference spring in their design.

However, an equivalent spring shall have a spring constant of at least 70N/mm and at least 2.5mm compression. The inner diameter shall be 4.2mm, and the outer diameter shall be 7.8mm. These springs fit into 8mm diameter counterbores of 4mm depth in the bottom stiffener. The installation method uses glue (3M DP810 or equivalent), applied with a maximum thickness of 0.1mm.

MPN: Timson WG774265 or equivalent



D (mm)	d (mm)	L (mm)	a (mm)	F (mm)	K (N/mm)
$7.8^{+0.05}_{-0.20}$	$4.2^{+0.1}_{-0.15}$	$6.5^{+0.2}_{-0.2}$	$3.6^{+0.15}_{-0.15}$	$2.9^{+0.35}_{-0.35}$	78.6±10%

Figure 14 Die Spring dimensions and drawing

Table 2 Spring constant and free length of die springs, shown compared to cycle count

	Sample 1		Sample 2		Sample 3		Sample 4		Sample 5	
	L (mm)	K (N/mm)	L (mm)	K (N/mm)	L (mm)	K (N/mm)	L (mm)	K (N/mm)	L (mm)	K (N/mm)
1	6.48	81.87	6.44	80.88	6.48	82.12	6.45	80.20	6.46	80.70
2	6.47	81.23	6.43	79.80	6.47	81.86	6.44	79.98	6.46	80.58
3	6.47	81.01	6.42	79.84	6.47	81.33	6.44	79.96	6.45	80.48
4	6.46	80.95	6.42	79.70	6.46	81.20	6.44	79.80	6.45	80.46
5	6.46	80.95	6.41	79.37	6.46	81.17	6.43	79.68	6.44	80.28
6	6.46	80.90	6.41	79.22	6.46	81.13	6.43	79.40	6.44	80.29
7	6.46	80.79	6.40	79.31	6.45	80.97	6.42	79.52	6.43	80.12
8	6.45	80.77	6.40	79.16	6.45	81.17	6.42	79.48	6.43	80.02
9	6.45	80.76	6.39	79.11	6.43	80.98	6.41	79.47	6.42	79.90
10	6.44	80.68	6.39	79.02	6.43	80.91	6.41	79.38	6.42	79.93
Avg	**6.46**	**80.99**	**6.41**	**79.54**	**6.46**	**81.28**	**6.43**	**79.69**	**6.44**	**80.28**

6.5 Baseboard Keep out Zone & Grounding Pads

The below figure shows the baseboard outline (top side view) required to accommodate this module. All cross-hatched areas need to be grounded except for the four corners 10x10 square holes. The 10x10 square holes in the corners are highly recommended for being grounded. Additionally, it is recommended to route high-speed traces away from mounting hole areas due to large compression forces from the die spring.

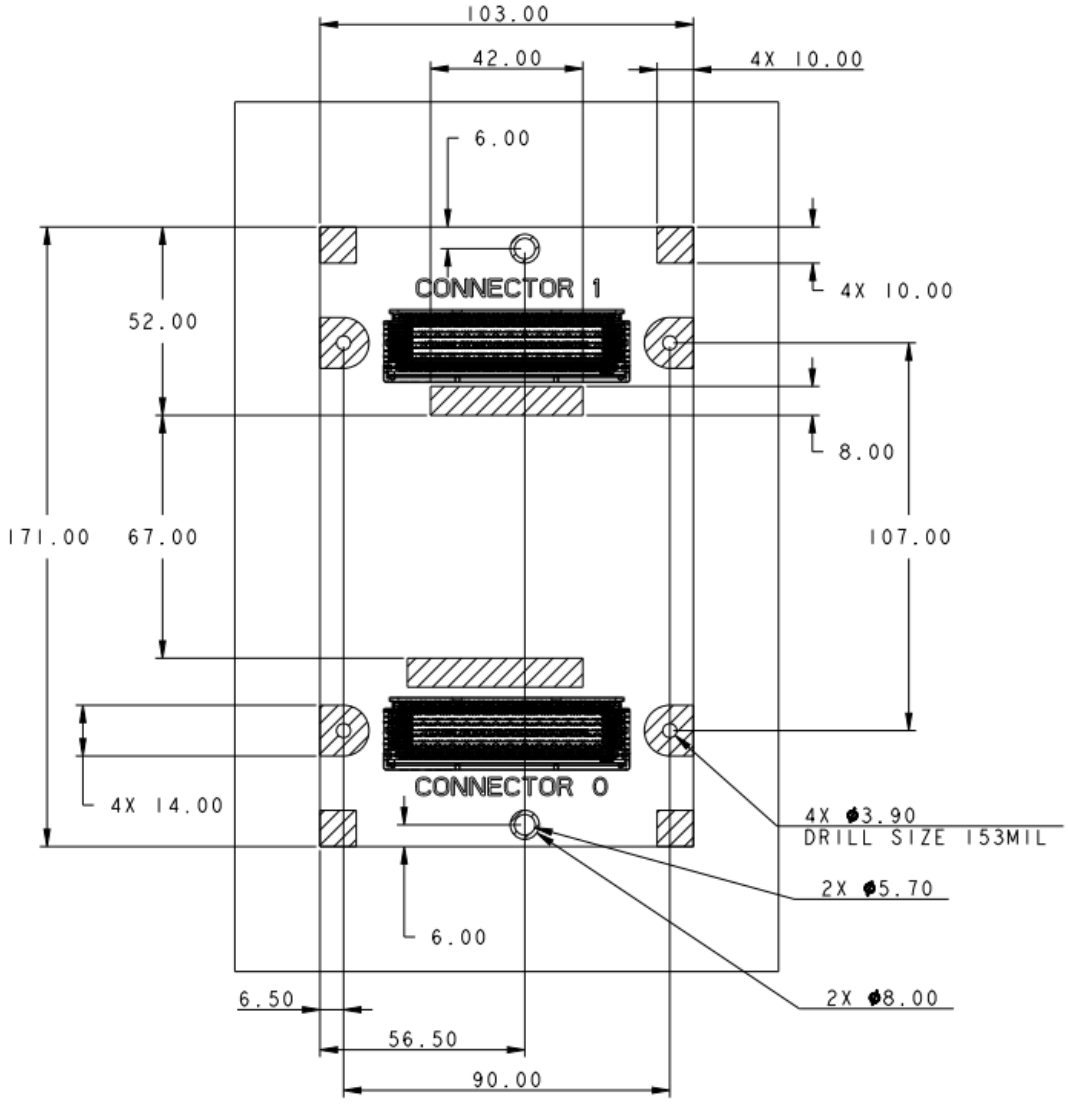


Figure 15 Baseboard KOZ and Grounding Pad Dimensions

6.5.1 SMT Nut

The SMT nuts with the dimensions shown in Figure 16 are to be soldered to the baseboard with 5.7mm diameter holes. These nuts provide the mating features to the alignment pins on the bottom stiffener of

the OAM. Clearance of the 3mm pins in the 3.6mm nuts means that the module will come within 0.3mm of its final position.

MPN: Ray Home 1000401319 or equivalent

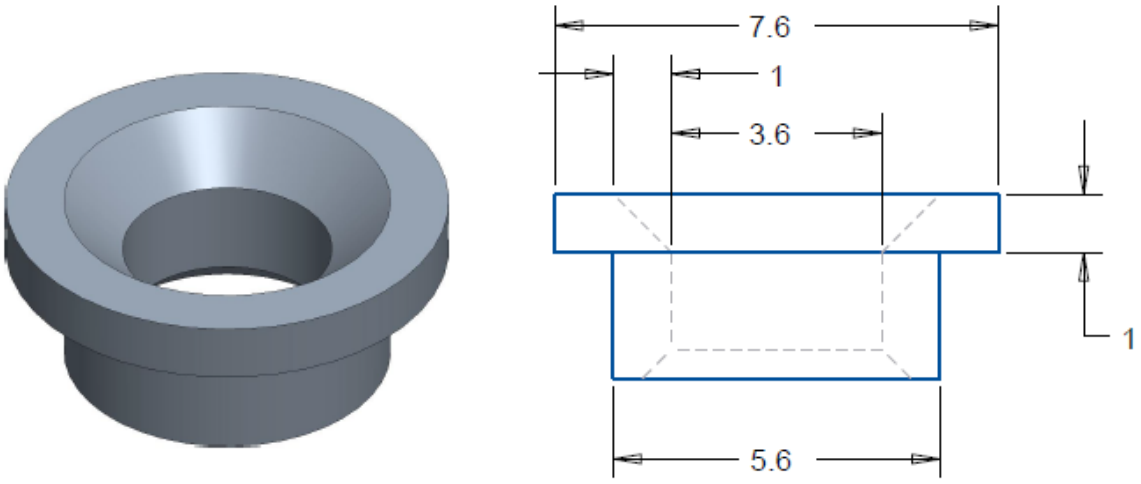


Figure 16 SMT Receiving Nut for Baseboard

6.5.2 Component Keep-out Zone

The baseboard has a component keep-out zone of 103x171mm, as shown in Figure 15 Baseboard KOZ and Grounding Pad Dimensions.

6.5.3 Grounding Pads

As with the bottom stiffener, the baseboard has two grounding pads of size 8x42mm, for the EMI fabric-over-foam gaskets on the stiffener to provide good contact. Refer to Section 6.4.3 for gasket MPN and description.

6.6 Recommended Alignment Features

There are three stages of engagement when installing the OAM to the system.

Stage 1: Notch in top of heatsink providing visual guidance and orientation reference. The reference design shows 2mm clearances.

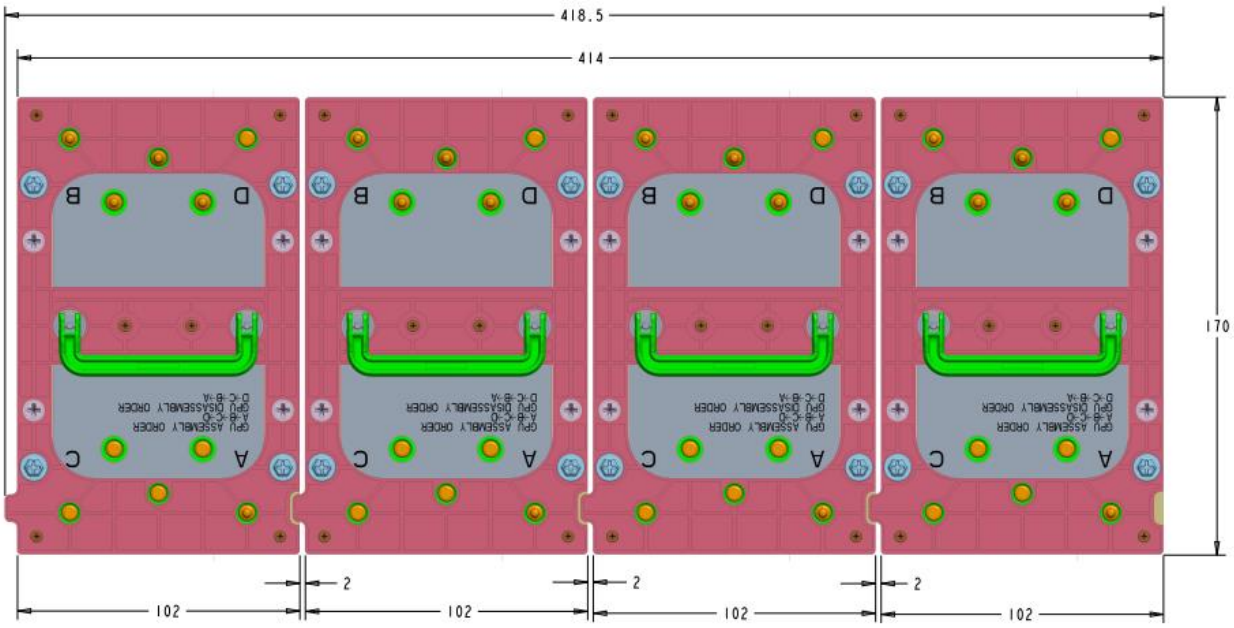


Figure 17 Top view of four adjacent OAM with heatsinks

Stage 2: Alignment pins, two 3mm pins from the OAM into two 3.6mm SMT nuts on the baseboard.

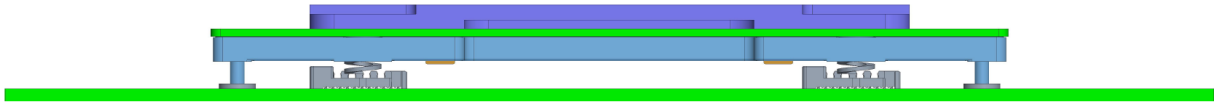


Figure 18 Side view (exploded) showing alignment pins being received by 1mm tall SMT nuts

Stage 3: Connector housing built-in engagement (Molex Mirror Mezz Pro gather ability: 0.76mm).

Figure 19 Side view (exploded) showing mezzanine connectors doing final alignment

6.7 Reference Heatsink Design

An air-cooled solution recommends TDP equal or less than 600W modules. For modules that are over 600W, other cooling solution is recommended such as liquid cooling.

Below figure shows the reference model of the heatsink with OAM assembly.

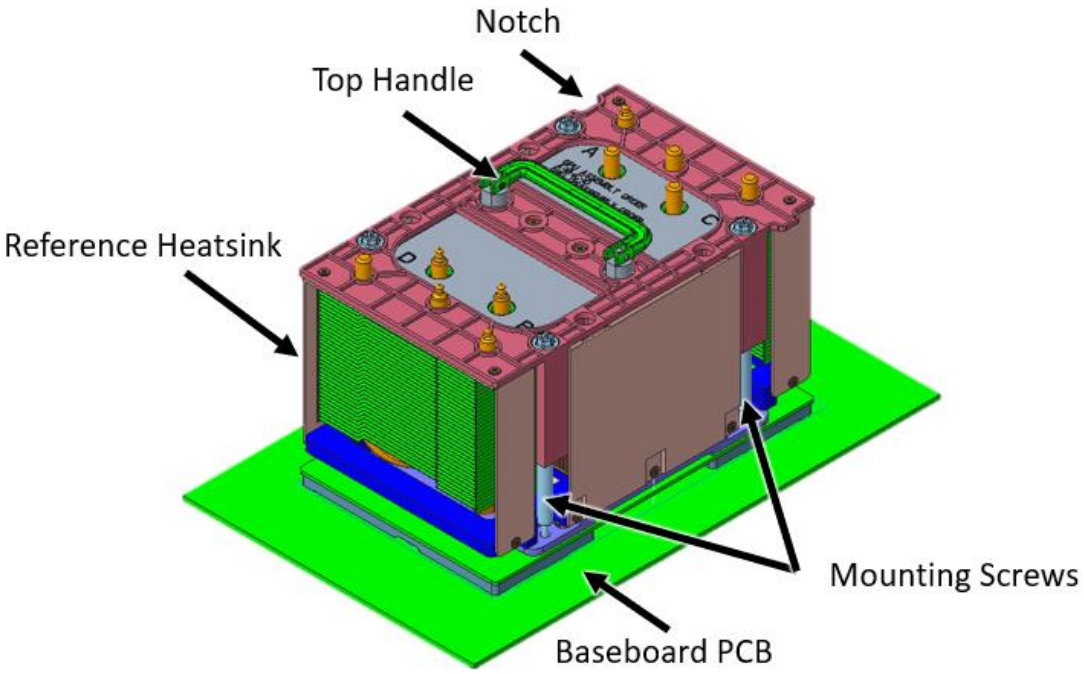


Figure 20 Reference Model of Heatsink with OAM Assembly

6.7.1 Top Handle

Due to the size and bulk of the heatsink and module assembly, a handle is recommended. The reference design uses a folding handle. This handle screws into a sheet metal panel attached to the heat sink base with six M2.5 flathead screws. This attachment method allows the load to transfer through the more rugged base instead of the delicate heat sink fins.

MPN: Fivetech 62-57P-064-7-02-5 or equivalent

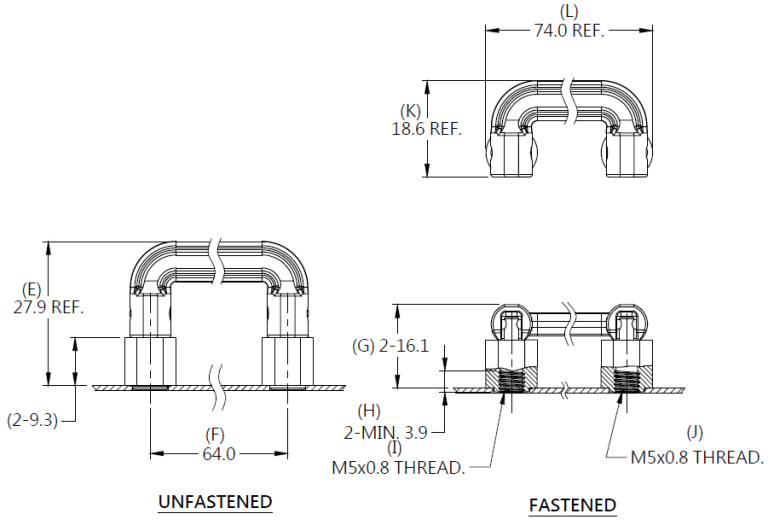


Figure 21 Dimensioned Smart Folding Handle from Reference Design

6.7.2 Long Screw Attachment

A set of four M3.5, spring-loaded, Phillips head long screws are used to attach the module to the baseboard. Note that the reference screw provided is simply a reference and that L3 and L4 will need to be adjusted based on the thicknesses of the baseboard and bolster plates. However, these mounting screw locations are fixed per the requirements of the OAM board layout and the baseboard layout. Each screw clears the top stiffener, mezzanine PCB, bottom stiffener (including the die spring), and the baseboard to screw directly into the bolster plate below the baseboard PCB. It is recommended that the OAM be attached to the baseboard by torquing the screws in a diagonal pattern.

MPN of Long Screw: Wujiang Screw MDCM0359733N or similar
MPN of Spring for Long Screw: Surpassing Hardware Spring FDJG7004010 or similar

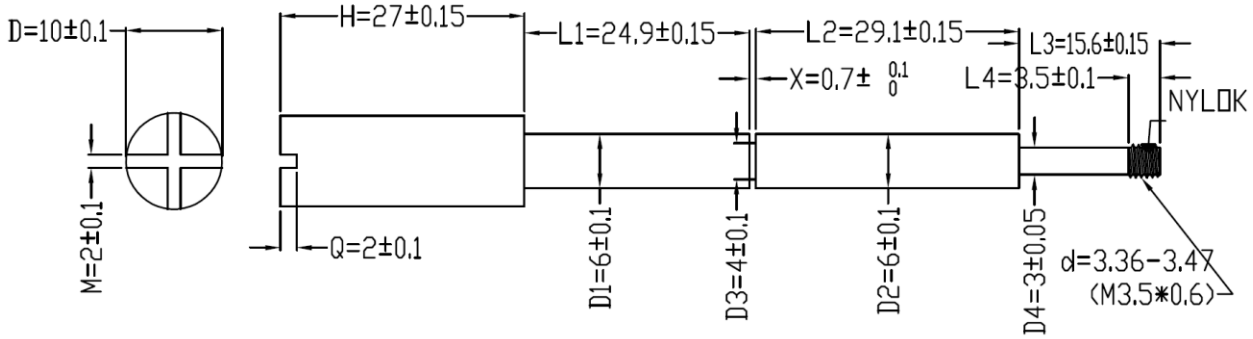


Figure 22 Drawing of Wujiang Screw MDCM0359733N

7 Thermal Specification

The thermal requirements are applicable to air-cooled conditions by default, except where a different cooling approach is specifically called out.

7.1 Environmental Conditions

The thermal and cooling solution should dissipate heat from the components when the module operates at its thermal design power to meet the thermal reliability requirement. The module should be able to work in the following environmental conditions without any throttling or thermal issues:

- Ambient temperature: 5°C to 35 °C
- Approach temperature: 10°C to 48 °C, considering shadowing other components
- Altitude: sea level to 3000 ft*, without temperature deration
- Relative Humidity: 20% to 90%
- Cold boot temperature: the module should be able to boot and operate at an initial temperature of 10°C

*An extended altitude range of up to 6000ft is recommended.

In addition, the module should remain unaffected at a non-operational storage temperature range of -20°C to 85°C.

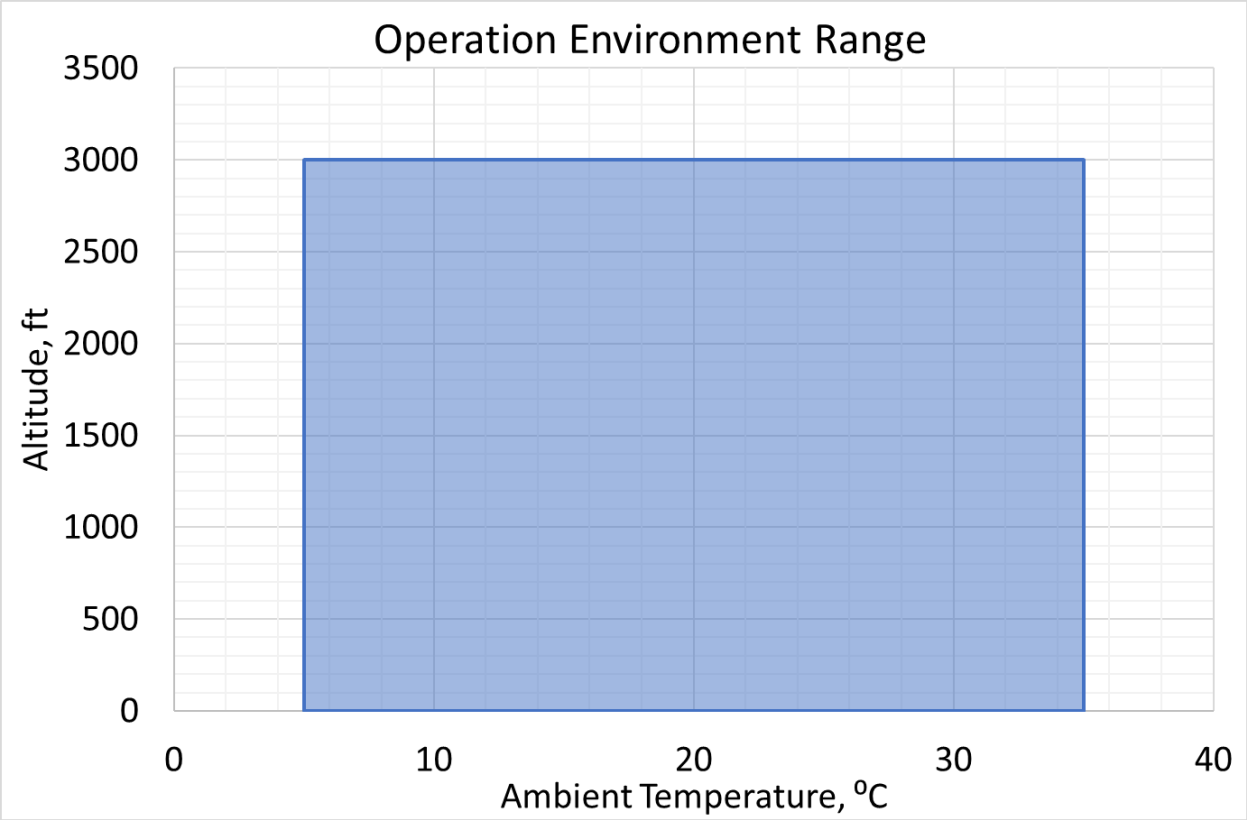


Figure 23 Module Operation Ambient Temperature

7.2 Temperature Report

7.2.1 Temperature Sensors

The module reports readings of ASIC temperature sensor and HBM temperature sensor to support software or hardware throttling, shutdown, and drive fan speed through BMC. The sensors should be located or calibrated to:

- Always report the hottest junction temperature in the component
- Keep accuracy within ±3°C

Lower temperature limit, non-critical temperature limit, and critical temperature limit should be defined for those temperature sensors to support throttling or shutdown features.

7.2.2 Remaining Components

For the remaining components that are not monitored by temperature sensors or not included in fan speed control (FSC), their cooling solutions should be appropriately designed such that:

- Before ASIC or Memory temperature readings reach throttling thresholds, they will be maintained below the temperature limits.
- When any ASIC or Memory temperature reading reaches a throttling threshold but not the hardware shutdown limit, these components will remain functional to support the reduced functionality of the module.

7.3 Thermal module information

To enable the module with appropriate cooling solutions, the supplier will provide the following thermal info for each product model:

- ASIC & Memory (HBM or DRAM) junction temperature limit.
- ASIC & Memory (HBM or DRAM) junction to surface/case temperature correlations.
- Connector surface temperature limit.
- ASIC & Memory (HBM or DRAM) junction temperature range at nominal operation conditions.

7.4 Heatsink Assembly

The module will meet these requirements to minimize the complexity of assembly, servicing, and risk of failure:

- Only one replaceable heatsink assembly (primary heatsink) is needed for the module, which can be swapped in the field.
- The other heatsink parts (i.e., secondary heatsinks) and thermal interface materials will come with the module and do not need replacement over the module lifetime.

Reliability test reports will be provided to validate the lifetime of the thermal interface materials. Shock and Vibration test reports will be provided to validate the robustness of the module assembly.

7.5 Thermal Recommendation

7.5.1 Airflow Budget

Considering OAM module meets the limits on the air delivery/removal capabilities of typical infrastructures, it is recommended that the OAM module operates with a complete performance at or below an airflow/power ratio of 0.158 CFM/W, with ambient temperature up to 35°C at sea level, equivalent to an inlet/outlet air temperature increase of 20°F.

- For operation at altitude, the same air temperature difference of 22°F is recommended.
- For a single OAM that is shadowed by other components, the airflow/power ratio calculate with airflow through its heatsink and the module power
- For an OAM shadowing other components or multiple OAMs in serial, this calculation uses the airflow through the flow channel and the sum of the power of OAM modules and upstream components.
- For OAM modules with a power lower than 300W, an airflow/power ratio of 0.1 CFM/W or lower is usually achievable and recommended.

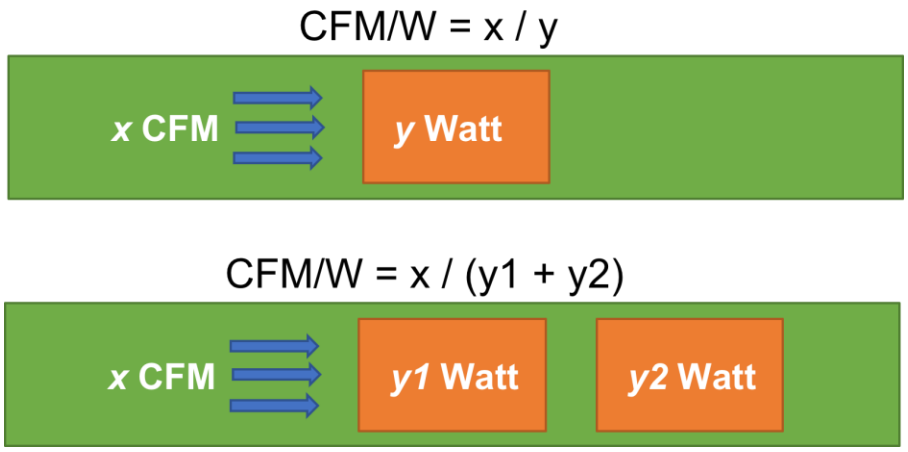


Figure 24 CFM per Watt

7.5.2 Reference Heatsink Design

Refer to Figure 20 Reference Model of Heatsink with OAM Assembly. A reference heatsink design with 3D-VC technology is provided as in Figure 25.

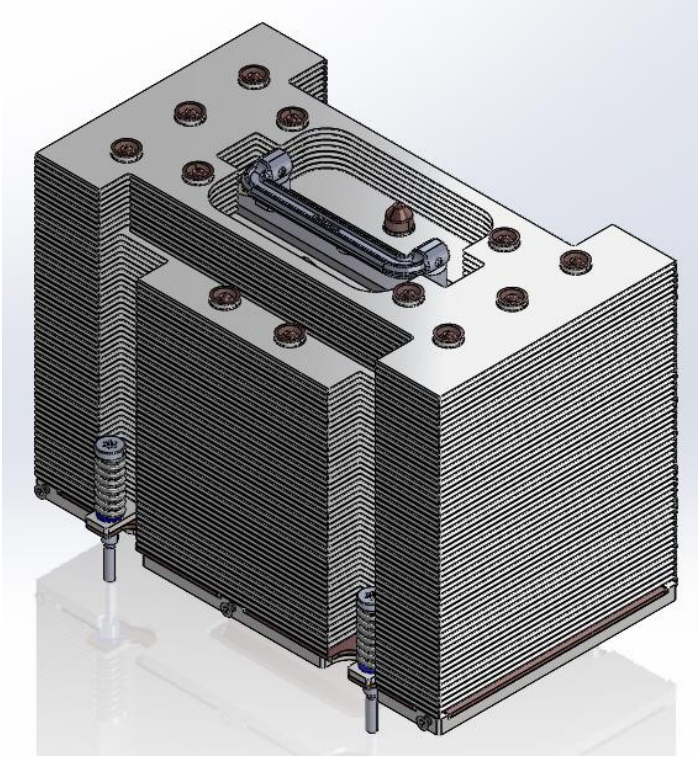


Figure 25 OAM 2.0 Reference Heatsink Solution, 4U version

Performance of the reference heatsink is provided in Figure 26, the thermal resistance of which calculate based on:

$$R_{ca} = \frac{T_{case} - T_{LA}}{P_{die}},$$

Where T_{case} is the surface center temp of the heater, T_{LA} is the approaching temperature, and P_{die} is the heater's power, indicating the die instead of the total module power.

Die size and power density play an essential role in the thermal performance of the OAM module. The chart below provides expected thermal resistance and pressure drop of the heatsink with multiple assumptions applied. Those assumptions include nominal estimation of package size, power distribution and thermal interface material to the best of our knowledge, and performance on every real OAM2.0 product could vary due to different characteristics on any of the metrics mentioned. Results demonstrated in Figure 26 were based on TIM resistance of 0.2°C·cm^2/W.

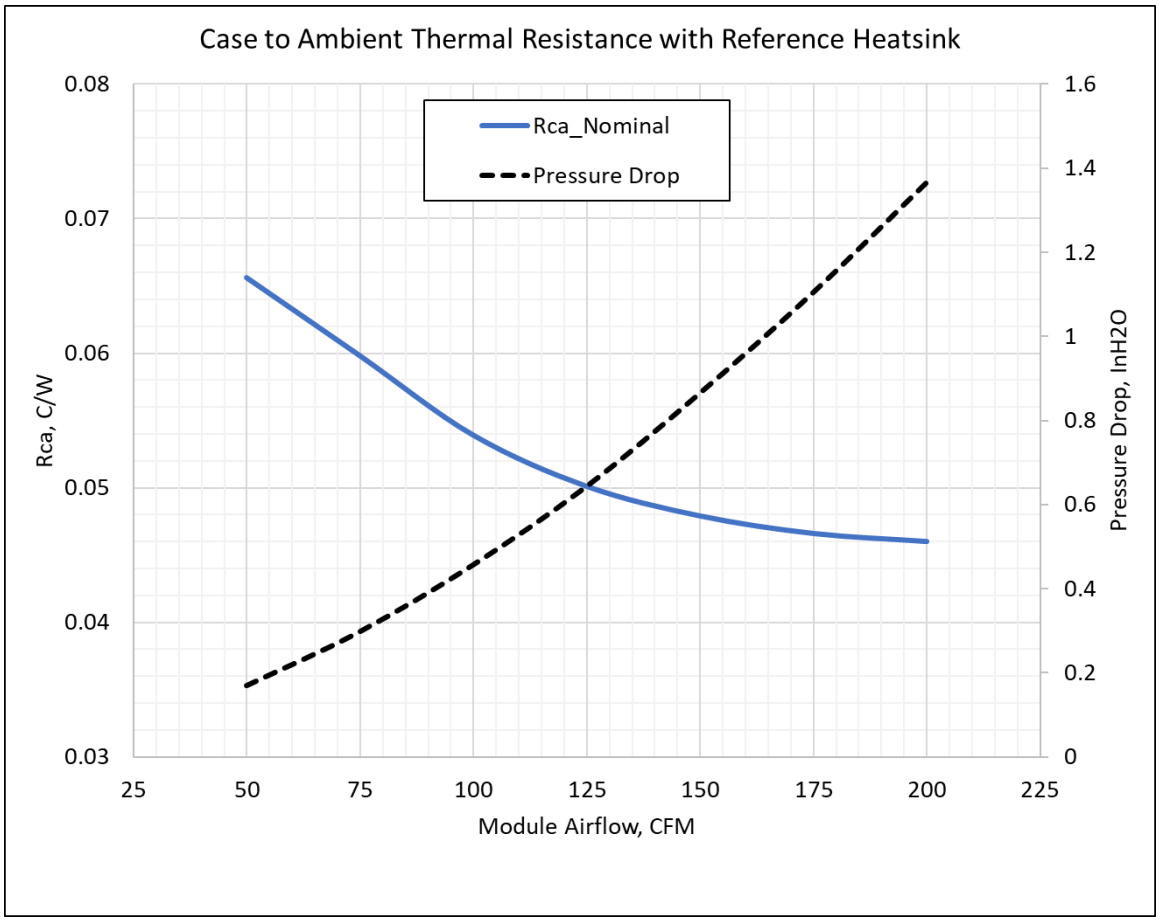


Figure 26 Thermal resistance and pressure drop of reference heatsink; assumptions apply.

Depending on model and application, the OAM may operate at a variety of power levels. However, air-cooled heatsinks may hit their performance limit due to the constraints on heat spreading and airflow delivery. Beyond a certain chassis height, fin size, and airflow rate, the improvement on the thermal resistance of the air-cooled heatsink becomes minimal.

For a reference OAM in a typical platform with 8x OAMs, shadowing layout, it is observed that the maximum module power that air cooling can support is approximately 600W. Beyond this power limit, advanced cooling solutions are recommended to keep its operation at the hotter part of the operational boundary condition range. These advanced cooling solutions would also be recommended for extended environment boundary conditions. This limit may vary for different products, depending on die size, power distribution, and junction temperature limits.

7.5.3 Thermal Interface Material

The thermal interface material between the die (ASIC/HBM) and the primary heatsink should maintain a thermal conductivity of at least 3W/m×K through the end of its life.

Depending on a variety of parameters, TIM could easily contribute 0.1 $^{o}C \cdot cm^2/W$ ~ 0.2 $^{o}C \cdot cm^2/W$, up to 50% of total thermal resistance.

The potentially achievable performance with better thermal interface resistance at 0.1$°C \cdot cm^2/W$ is demonstrated in Figure 27. This may be achieved by using more suitable TIM choices, higher mounting pressure, or better warpage control of the package. Such resistance is usually observed on well-controlled lidded packages, where tradeoff applies due to junction-to-case temperature gradient.

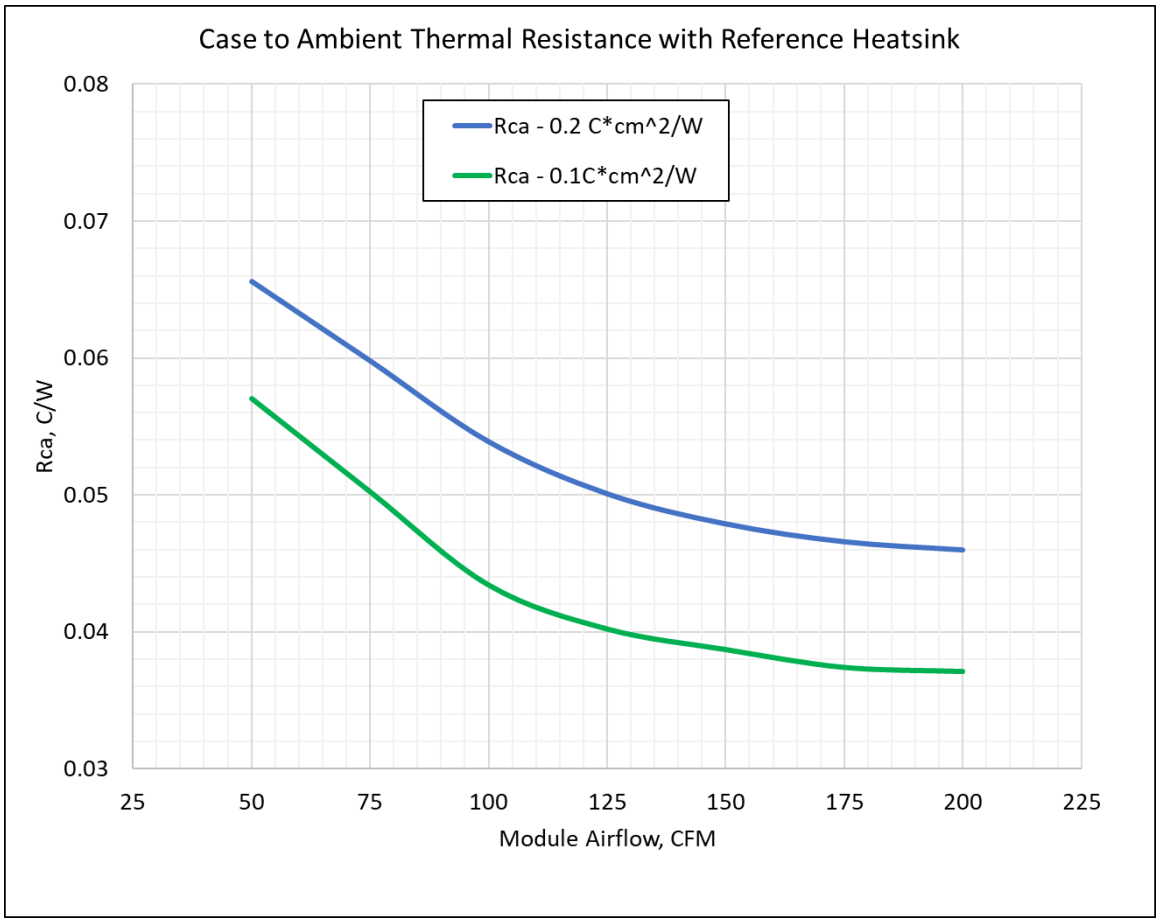


Figure 27 Thermal resistance improvement due to better thermal interface between OAM package and heatsink

7.5.4 Heatsink Installation

Many OAM modules use bare die design, which may be fragile and susceptible to an imbalance of pressure on its surface. The system integrator should contact the OAM supplier for the maximum static and dynamic pressure for the die to guide the installation of the primary heatsink to the module. The static mounting pressure should also be high enough to enable optimum performance of the TIM material.

We suggest installing the heatsink with multistage diagonal tightening sequences, for better balancing of the stress during the installation process. Customized jigs enabling simultaneous tightening of multiple screws may provide better yield rate.

The mounting pressure of the heatsink is determined by:

- Max pressure the package can sustain
- Min pressure the TIM needs to deliver enough performance

The mounting pressure may vary across the range of 20 ~ 60 psi for lidless OAM products, where lidded products could potentially sustain higher mounting pressure. We recommend an initial mounting pressure of 15~30 psi for engineering samples without assembly yield rate learnings.

7.6 Liquid Cooling Requirements

7.6.1 Boundary Conditions

The OAM coldplate should be able to operate with coolant supply temperature across the spectrum of 15~50°C. Higher coolant temperatures within the spectrum are usually preferred for better hardware and facility efficiency, however the tradeoff would be cooling performance reduction. Accelerated test at higher temperatures are recommended to validate the design robustness and long term reliability. The coolants applicable to OAM cold plates are water with additives (treated water), or glycol-based liquids (PG25, etc.).

7.6.2 Reliability Expectations

The OAM coldplate products should pass the reliability tests described in OCP Coldplate Development and Qualification, including:

- Hydrostatic Pressure Test
- Corrosion Test
- Dynamics Shock & Vibration test
- Temperature Cycling Test

It is recommended to have OAM coldplate products' annual failure rate below 0.3%.

7.6.3 Reference Coldplate Design

An OAM 2.0 reference coldplate is demonstrated in Figure 28. With ideal package layout, coldplate design, quality control and thermal interface resistance of 0.1 °C·cm^2/W, such coldplate is possible to deliver a case-to-inlet resistance of 0.02 °C/W in the best case, as shown in Figure 29.

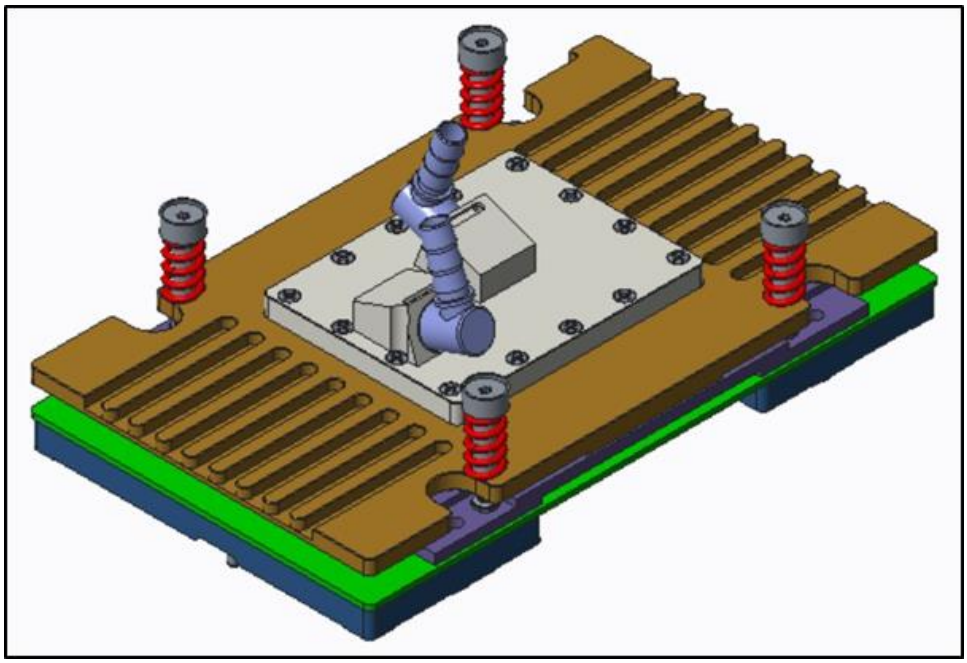


Figure 28 An Example of Open Loop Liquid Cooling setup concept for OAM

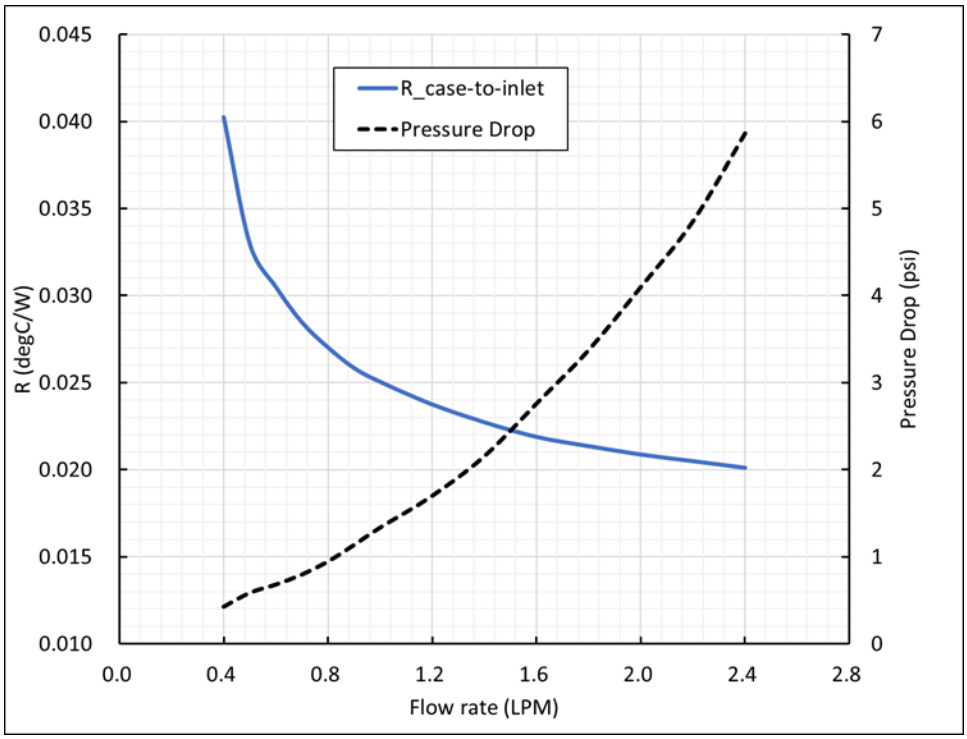


Figure 29 Performance Curve of Reference Coldplate under ideal assumptions

7.6.4 More Recommendations

More guidelines around OAI liquid cooling are provided in this publication:

- OAI Liquid Cooling Guidelines:
    - https://www.opencompute.org/documents/oai-system-liquid-cooling-guidelines-in-ocp-template-dec-7-2022-1-pdf

From OCP CE Coldpate group, there are also multiple guidelines published to help coldplate-based liquid cooling solution development for general hardware products:

- OCP ACS Cold Plate Leak Detection and Intervention
    - https://www.opencompute.org/documents/acs-cold-plate-leak-detection-and-intervention-white-paper-pdf-1
- OCP ACS Liquid Cooling Cold Plate Requirements
    - https://www.opencompute.org/documents/ocp-acs-liquid-cooling-cold-plate-requirements-pdf
- OCP Cold Plate Development and Qualification
    - https://www.opencompute.org/documents/ocp-cold-plate-development-and-qualification-with-integrated-comments-pdf

7.7 Consideration for Immersion Cooling

Immersion cooling is another advanced cooling technology that could bring performance benefit. However the impact on various aspects are still yet to be examined, including material compatibility, signal integrity, hardware design philosophy, reliability and serviceability, facility design, etc. A series of guidelines have been released by OCP CE Immersion group, to help development of immersion-based hardware products on multiple, but not all areas yet:

- Material Compatibility in Immersion Cooling
    - https://www.opencompute.org/documents/material-compatibility-in-immersion-cooling-document-version-1-0-nov-28-2022-1-pdf
- Base spec of immersion fluid
    - https://www.opencompute.org/documents/ocp-base-specification-for-immersion-fluids-20221201-pdf
- Design Guidelines for Immersion-Cooled IT Equipment
    - https://www.opencompute.org/documents/design-guidelines-for-immersion-cooled-it-equipment-revision-1-01-pdf

8 OAM Electrical Specification

8.1 Electrical Connector

The module utilizes two 688pin Molex Mirror Mezz Pro connectors. It is a BGA attached connector. It supports bit rates up to 56Gbps NRZ or 112 Gbps PAM4 in a 90 Ohms nominal impedance ± 5% tolerance, making it compatible with support typical 85 Ohms based interfaces such as PCIe Gen 5/6 as well as other 100 Ohms based high-speed interfaces. All power and I/O signals are routed through the two connectors down to the system baseboard. The system baseboard should connect these signals to the appropriate circuitry depending on the required feature sets. The below table lists the electrical requirements for the module connectors.

Table 3 Electrical Requirements for Molex Mirror Mezz Pro

Items	Mirror Mezz Pro
Data Rate Support	Up to 56G NRZ, 112G PAM4
Connector Impedance	90ohm ± 5%
Differential pairs per two connectors	172 pairs
Pin Pitch	0.9mm and 1.3mm
Current Rating per pin @85C (inclusive of 30C T-rise) ambient temp, 1.5oz copper	1A/pin after 20% derating
Max Voltage Application	30V AC (OAM supports 60V)
Connector insertion cycles	100cycles
Withstand voltage	500V min
Low-Level Contact Resistance (max initial):	30mΩ for 5mm stack height
Insulation resistance	1-MΩ min
Intra-pair skew	<=5 ps

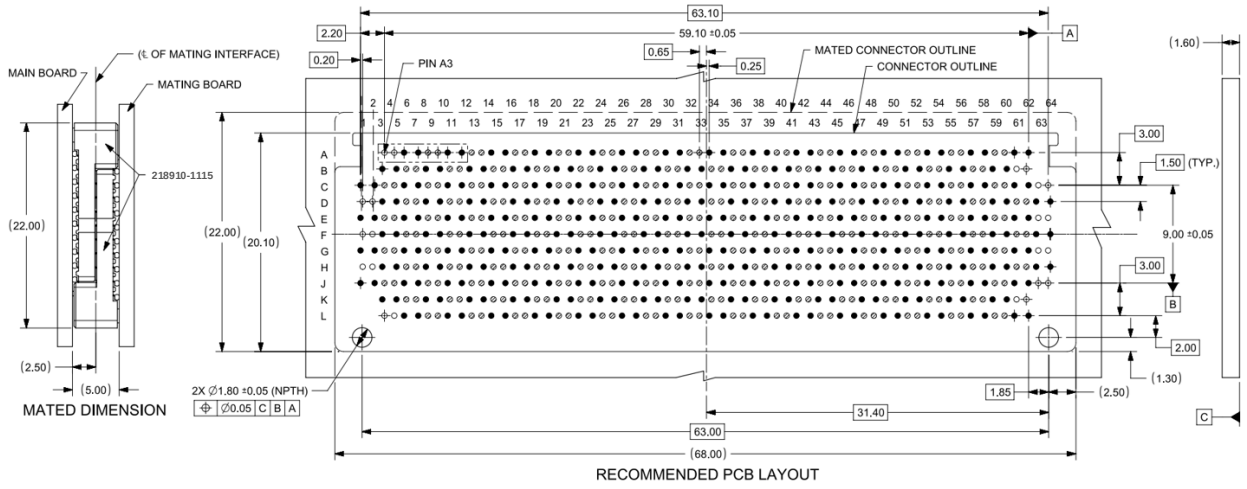


Figure 30 Mirror Mezz Pro Connector Footprint

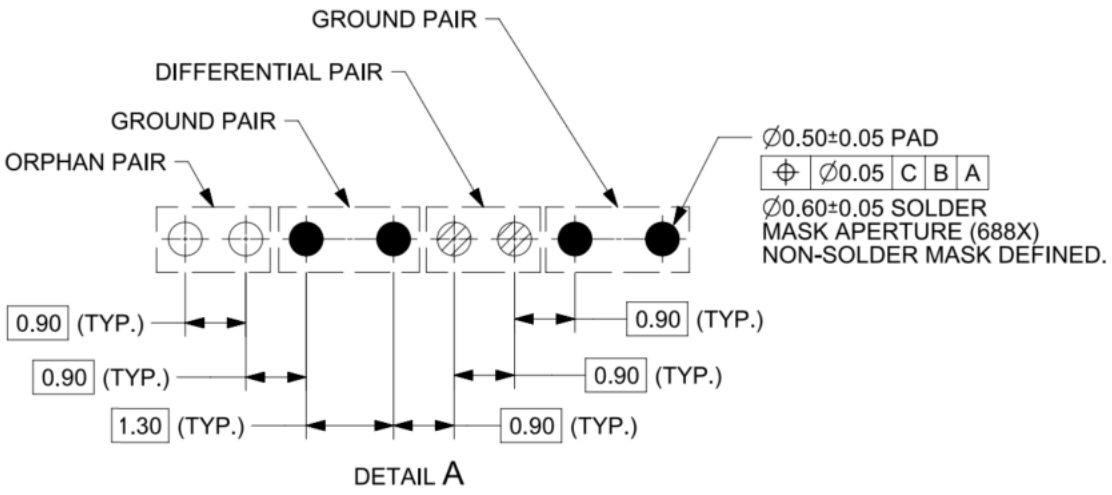


Figure 31 Mirror Mezz Pro Connector Pin to Pin Pitch

8.2 OAM Connector Pinout Quadrants

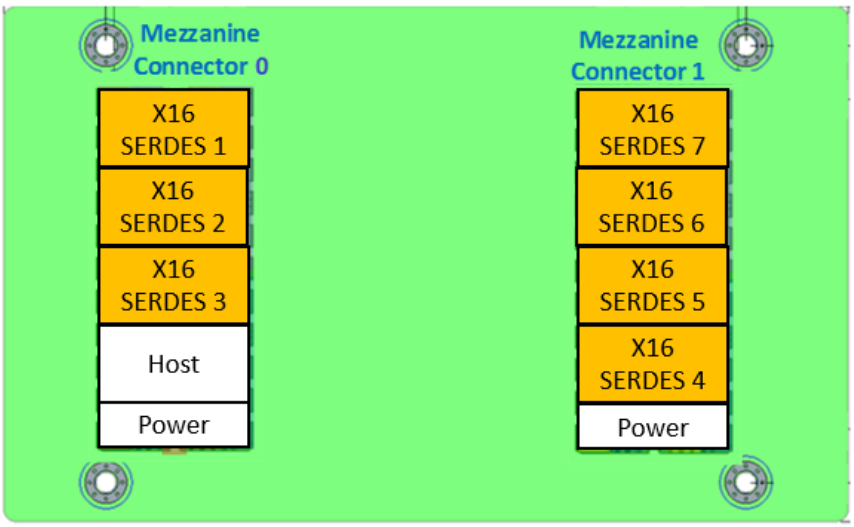


Figure 32 Mezzanine Connectors Pinout Quadrants

The OAM Connector 0 has the following interfaces:

- 54V/48V and 3.3V input power
- x16 SerDes to connect to host
- 3x16 SerDes for the accelerator to accelerator communication
    - x16 may split into sub-links like 2* x8s or 4* x4s or 16*x1s.
    - If the ASIC or ASICs on the module only support x8 or x4 per SerDes, it should start from Lane0 from the SerDes, e.g., lane [7:0] or lane [3:0].

    - We do not recommend lane reversal support on the baseboard due to modules having the option to be 1x16 links or 2x8 or 4x4 links.
- Other single-ended signals, like PRESNT#, SMBus, GPIOs etc.

The OAM Connector 1 has the following interfaces:

- 54V/48V and 12V (optional) input power
- Other single-ended signals like JTAG, GPIOs etc.
- Up to four SerDes for the accelerator to accelerator communication or other purposes:
    - SerDes 4, 5, 6, and 7 are up to x16 lanes which can split into x8s, x4s, or x1s.
- SerDes 7 may be defined for different use cases:
    - This link could be the 7th SerDes for some cases to have a fully connected interconnect between the modules
    - It could be the 2nd link to host for the ASIC(s) on the module, e.g., a full x16 link, 2x8, or 4x4 links.
    - Or it could be a unique defined link by some ASICs. E.g., it could be a downstream port for the ASIC on the module.

8.3 OAM Pinout Description

The detailed pin mapping to connectors is in the separated spreadsheet (OAM_Pinlist_Pinmap_r2.0_v0.75.xlsx). This section only shows the pin list and description.

Table 4 OAM Pinouts

Signal	IO Type (Mezz Module Direction POV)	Description	Voltage	Total Diff Pair	Total SE pins	Conn0 or 1	Note
P48V	Power Input	44V-59.5V main voltage for 1000W. 40V min for lower power (< 450W) Split power pins to Conn0 and Conn1. Check with OAM suppliers with 40V support	44V-59.5V		100	Conn0/ Conn1	
P12V	Power Input	12V Board Infrastructure Power. Up to 50W (Optional)	12V		8	Conn1	Optional
P3V3	Power Input	3.3V Main Power. Up to 10W	3.3V		6	Conn0	

PVREF	Power Output	Low voltage output for GPU/ASIC sideband I/O reference on motherboard components. Module should provision maximum 0.5A to be provided to motherboard. Vref can be any value between 1.2V to 3.3V set as per GPU/ASIC sideband I/O voltage. Each OAM supplier will specify the actual Vref.	Vref		2	Conn0	
PETp/n [15:0]	Output	PCIe or equivalent host link Transmit differential pairs. Module Transmit, Host Receive. Note: AC coupling caps must be placed on the motherboard side.		16		Conn0	
PERp/n [15:0]	Input	PCIe or equivalent host link Receive differential pairs. Module Receive, Host Transmit. Note: AC coupling caps must be placed on the motherboard side.		16		Conn0	
SERDES_1Tp/n [15:0]	Output	SerDes link 1 Transmit differential pairs. AC caps must be placed on Module/die (if required)		16		Conn0	
SERDES_1Rp/n [15:0]	Input	SerDes link 1 Receive differential pairs.		16		Conn0	
SERDES_2Tp/n [15:0]	Output	SerDes link 2 Transmit differential pairs. AC caps must be placed on Module/die (if required)		16		Conn0	
SERDES_2Rp/n [15:0]	Input	SerDes link 2 Receive differential pairs.		16		Conn0	
SERDES_3Tp/n[15:0]	Output	SerDes link 3 Transmit differential pairs. AC caps must be placed on Module/die (if required)		16		Conn0	
SERDES_3Rp/n [15:0]	Input	SerDes link 3 Receive differential pairs.		16		Conn0	
SERDES_4Tp/n[15:0]	Output	SerDes link 4Transmit differential pairs. AC caps must be placed on Module/die (if required)		16		Conn1	

SERDES_4Rp/n [15:0]	Input	SerDes link 4 Receive differential pairs.		16		Conn1	
SERDES_5Tp/n [15:0]	Output	SerDes link 5 Transmit differential pairs. AC caps must be placed on Module/die (if required)		16		Conn1	
SERDES_5Rp/n [15:0]	Input	SerDes link 5 Receive differential pairs.		16		Conn1	
SERDES_6Tp/n [15:0]	Output	SerDes link 6 Transmit differential pairs. AC caps must be placed on Module/die (if required)		16		Conn1	
SERDES_6Rp/n [15:0]	Input	SerDes link 6 Receive differential pairs.		16		Conn1	
SERDES_7Tp/n [15:0]	Output	SerDes link 7 Transmit differential pairs. AC caps must be placed on Module/die (if required)		16		Conn1	
SERDES_7Rp/n [15:0]	Input	SerDes link 7 Receive differential pairs.		16		Conn1	
PE_REFCLKp/n	Input	PCIe Reference Clock. 100MHz PCIe Gen 5 compliant.		1		Conn0	
AUX_100M_REFCLKp/n	Input	100MHz PCIe Gen 5 compliant Auxiliary Reference Clock.		1		Conn1	
AUX_156M_REFCLKp/n	Input	156.25MHz Auxiliary Reference Clock (Optional)		1		Conn1	
DWN_REFCLKp/n	Output	Downstream Reference Clock. Vendor specific.		1		Conn1	
PERST#	Input	CEM Compliant PCIe Reset	3.3V		1	Conn0	
WARMRST#	Input	Warm Reset	Vref		1	Conn0	
DWN_PERST#	Output	Down device PCIe Reset. Vendor specific.	3.3V		1	Conn1	
HOST_PWRGD	Input	Host power good. Active high when P48V, P12V, P3V3 voltages are stable and within specifications. This is considered the “Power Enable” signal for the module.	3.3V		1	Conn0	
MODULE_PWRGD	Output	Module power good. Active high when the module has completed its own power up sequence and is ready for PERST# de-assertion	3.3V		1	Conn0	

PWRBRK#	Input	Emergency power reduction. CEM Compliant Power Break	3.3V		1	Conn0	
PWRRDT#[1:0]	Input	Power Reduction GPIO to instruct OAM to go certain stage to reduce power 11 - default state L0, normal power 10 - L1, 1st level power reduction. 01 - L2, 2nd level power reduction. 00 - L3, max power reduction. Details defined by specific OAM product specification.	3.3V		2	Conn1	
THERMTRIP#	Output	Catastrophic thermal event for module components. Active low and latched by the Module logic. Released until motherboard power cycles the module input voltages	3.3V		1	Conn0	
MODULE_ID[4:0]	Input	Module node identifier (e.g. Module #0, #1,…#n). Module has weak PU to drive to 1 by default.			5	Conn0	
LINK_CONFIG[4:0]	Input	Mezz Module Host Interface/SerDes Link Configuration and topology. See link config table for details. The module has weak PU to Vref to drive high by default.			5	Conn1	
PE_BIF[1:0]	Output	x16 Host Interface Bifurcation Configuration. 00 = one x16 PCIe host interface 01 = bifurcation into two x8 PCIe host interfaces 10 = bifurcation into four x4 PCIe host interfaces 11 = reserved Vref based when drive to high.			2	Conn1	

PLINK_CAP	Output	"P" Port Module Capability support: '0' = PCIe only support '1' = Alternate protocol supported The host system requests an alternate host link protocol by pulling up LINK_CONFIG[0] and the Module informs the system of protocol support on the "P" link via this pin. If the module only supports PCIe as host, this signal is PD on the module.	Vref		1	Conn1	
SMBus_SLV_D	Bi-directional	Slave SMBus data.	3.3V		1	Conn0	I3C compatible
SMBus_SLV_CLK	Input	Slave SMBus clock	3.3V		1	Conn0	I3C compatible
SLV_ALERT#	Output	Slave alert indication.	3.3V		1	Conn0	
I2C_D	Bi-directional	Master I2C/SMBus data . PU on OAM.	Vref		1	Conn0	I3C compatible
I2C _CLK	Output	Master I2C/SMBus clock. PU on OAM.	Vref		1	Conn0	I3C compatible
UART_TXD	Output	Serial Port Transmit	3.3V		1	Conn0	
UART_RXD	Input	Serial Port Receive	3.3V		1	Conn0	
JTAG0_TRST	Input	Low Voltage ASIC/GPU JTAG Test Reset	Vref		1	Conn0	
JTAG0_TMS	Input	Low Voltage ASIC/GPU JTAG Test Mode Select	Vref		1	Conn0	
JTAG0_TCK	Input	Low Voltage ASIC/GPU JTAG Test Clock	Vref		1	Conn0	
JTAG0_TDO	Output	Low Voltage ASIC/GPU JTAG Test Output	Vref		1	Conn0	

JTAG0 _TDI	Input	Low Voltage ASIC/GPU JTAG Test Input	Vref		1	Conn0	
JTAG1 _TRST	Input	High Voltage JTAG Test Reset	3.3V		1	Conn1	optional
JTAG1 _TMS	Input	High Voltage JTAG Test Mode Select	3.3V		1	Conn1	optional
JTAG1 _TCK	Input	High Voltage JTAG Test Clock	3.3V		1	Conn1	optional
JTAG1 _TDO	Output	High Voltage JTAG Test Output	3.3V		1	Conn1	optional
JTAG1 _TDI	Input	High Voltage JTAG Test Input	3.3V		1	Conn1	optional
PRSNT0#	Output	Module present pin connector 0. Tied to GND through 1K resistor on module side.			1	Conn0	
PRSNT1#	Output	Module present pin connector 1. Tied to GND through 1K resistor on module side			1	Conn1	
SCALE_DEBUG_EN	Output	At-scale debug enable on the module. Isolates any motherboard JTAG debug path when logic High	3.3V		1	Conn0	
DEBUG_PORT_PRSNT#	Input	Presence signal for debug port in baseboard. Notifies logic in the module the debug access is being used by the baseboard debug connector. Debug port on baseboard present when logic low.	Vref		1	Conn0	
MNGMT_LINK0Tp/n	Output	Vendor specific module to module management link port 0 transmit		1		Conn1	Optional
MNGMT_LINK0Rp/n	Input	Vendor specific module to module management link port 0 receive		1		Conn1	Optional
MNGMT_LINK1Tp/n	Output	Vendor specific module to module management link port 1 transmit		1		Conn1	Optional
MNGMT_LINK1Rp/n	Input	Vendor specific module to module management link port 1 receive		1		Conn1	Optional
TEST0	I/O		Vref		1	Conn0	Optional
TEST1	I/O		Vref		1	Conn0	Optional

TEST2	I/O		Vref		1	Conn0	Optional
TEST3	I/O		Vref		1	Conn0	Optional
TEST4	I/O		Vref		1	Conn0	Optional
TEST5	I/O		Vref		1	Conn0	Optional
TEST6	I/O		Vref		1	Conn0	Optional
TEST7	I/O		Vref		1	Conn0	Optional
TEST8	I/O		Vref		1	Conn1	Optional
TEST9	I/O		Vref		1	Conn1	Optional
TEST10	I/O		Vref		1	Conn1	Optional
TEST11	I/O		Vref		1	Conn1	Optional
TEST12	I/O		Vref		1	Conn1	Optional
TEST13	I/O		Vref		1	Conn1	Optional
TEST14	I/O		Vref		1	Conn1	Optional
MANF_MODE#	Input	Manufacturing Mode 1: Normal operation 0: Module enter into manufacturing mode	3.3V		1	Conn0	
FW_RECOVERY#	Input	On board manageability boot recovery mode 1: Normal operation 0: Firmware Recovery boot mode	3.3V		1	Conn0	
TEST_MODE#	Input	Compliance Test Mode 1: Normal operation 0: ASIC/GPU enter into electrical compliance mode	Vref		1	Conn0	
RFU		Reserved for future use			5	Conn0	
RFU		Reserved for future use			6	Conn1	
DO_NOT_USE		48/54V isolation			13	Conn0	
DO_NOT_USE		48/54V isolation			13	Conn1	
GND		Ground			317	Conn0	
GND		Ground			317	Conn1	

Note:

1) **When a catastrophic thermal event (THERMTRIP#) occurs, OAM should shut down all its onboard power rails.**
2) **The baseboard should do the following:**
    a. **Turn off all input clocks**
    b. **Tristate all OAM input GPIO**
    c. **Turn off all input power rails to OAM**

GPIO Recommended Operating Conditions

	Description	MIN	MAX	Unit
V_{3V3}	IO reference voltage	3.135	3.465	V
V_{ref}	2nd IO reference voltage, ranging from 1.2V to 3.3V	0.95*Vref	1.05*Vref	V
V_{IH-3V3}	High level of IO refer to 3.3V	0.7 x V_{3V3}	3.465	V
V_{IL-3V3}	Low level of IO refer to 3.3V	-0.5	0.3 x V_{3V3}	V
$V_{IH-Vref}$	High level of IO refer to Vref	0.7 x V_{ref}	V_{ref}	V
$V_{IL-Vref}$	Low level of IO refer to Vref	-0.5	0.3 x V_{ref}	V

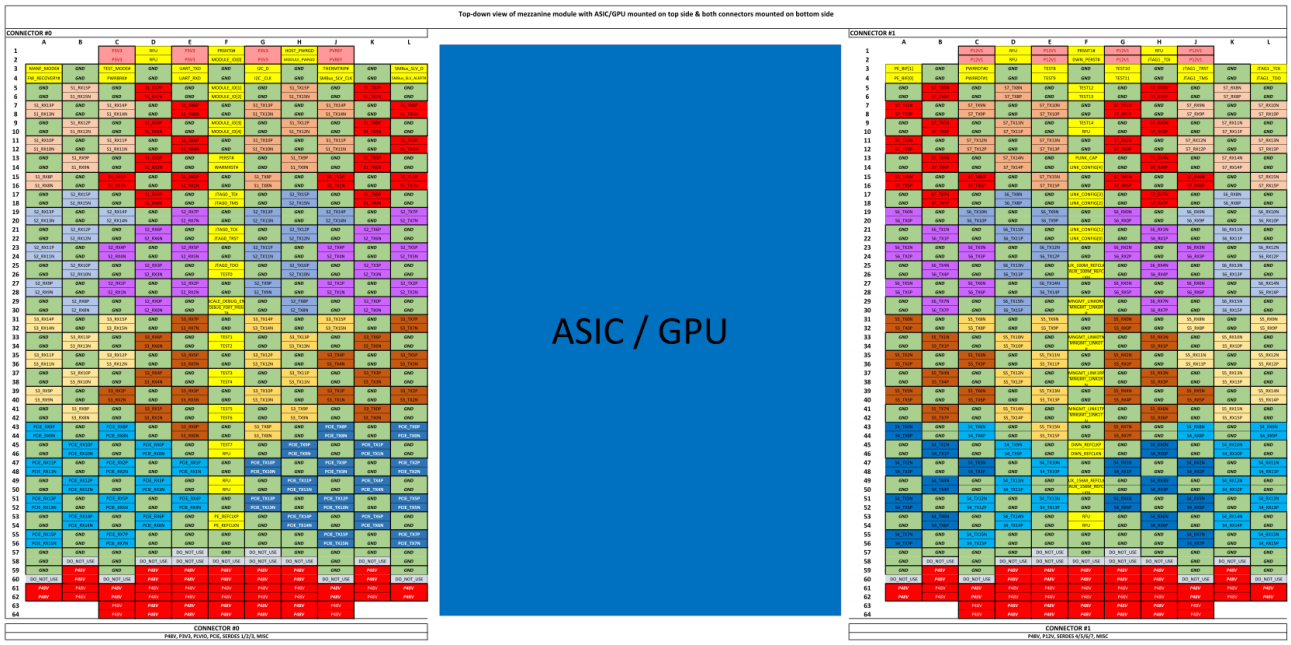


Figure 33 – Mezzanine Connector Pin map

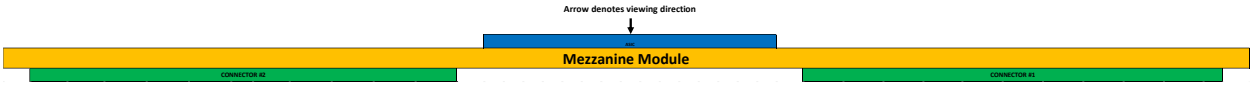


Figure 34 – Mezzanine Connector Pin Out View Reference

8.4 OAM Power Profiles

This section defines the maximum thermal design power (TDP) the module can support and the excursion design power (EDP).

8.4.1 Thermal Design Power TDP

The module supports up to 1000W with input voltage of 44V to 59.5V (<44V for lower power; <450W)The UBB supplies power to the module through both the Mirror Mezz Pro Connector0 and

Connector 1 power pins. The current capability and power status are in the table below. The power is available in state S0 only. Eight P12V power pins are optional. The UBB can supply all three power rails.

Table 5 Power Rails

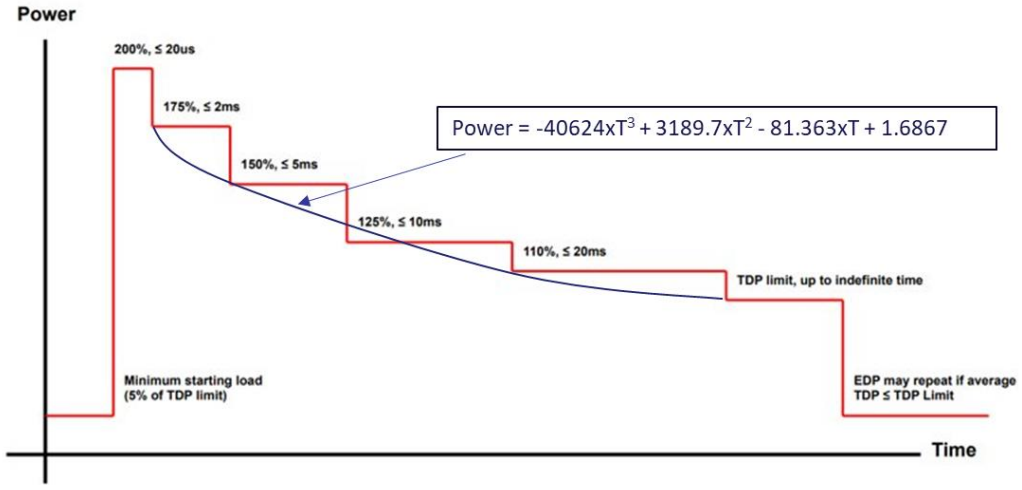
Power Rail	Voltage Tolerance	# of pins	Current Capability	Status
P12V Optional	11V min to 13.2V max	8	8A (when at 11V)	Normal Power
P48V	40V min to 59.5V max	100	100A (when at 44V)	Normal power
P3.3V	3.3V±10%(max)	6	3.3A nominal (6A max)	Normal power

8.4.2 Excursion Design Power EDP

System design should support the OAM's excursion design power (aka EDP). The OAM VR electrical design must handle the instantaneous peak power short period (usually on the order of a µs) with a low duty cycle. The VR's thermal design should be robust enough to handle lower power EDP levels (e.g., 1.1x TDP) for ms level interval without asserting VR HOT over-temperature alert. The system integrator should closely work with the module suppliers to ensure that the baseboard supplies enough power without triggering under-voltage protection.

Table 6 Excursion Design Power Example

EDP	Duration
2x TDP	<= 20µs
1.75x TDP	<=2ms
1.5x TDP	<=5ms
1.25x TDP	<=10ms
1.1x TDP	<=20ms



8.5 System power sequencing

System design should follow below power sequence requirement. It is recommended to check with each specific module specification to ensure the modules work correctly.

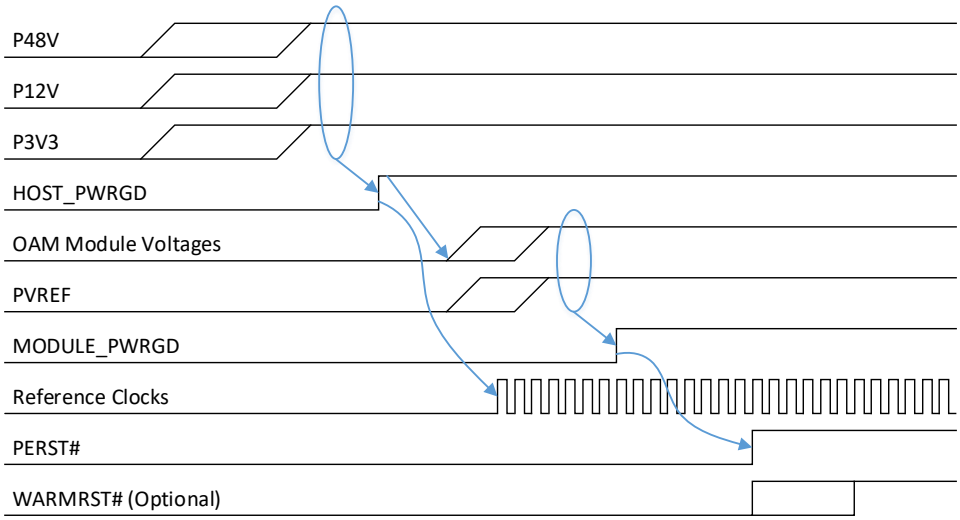


Figure 35 OAM Power Up Sequence

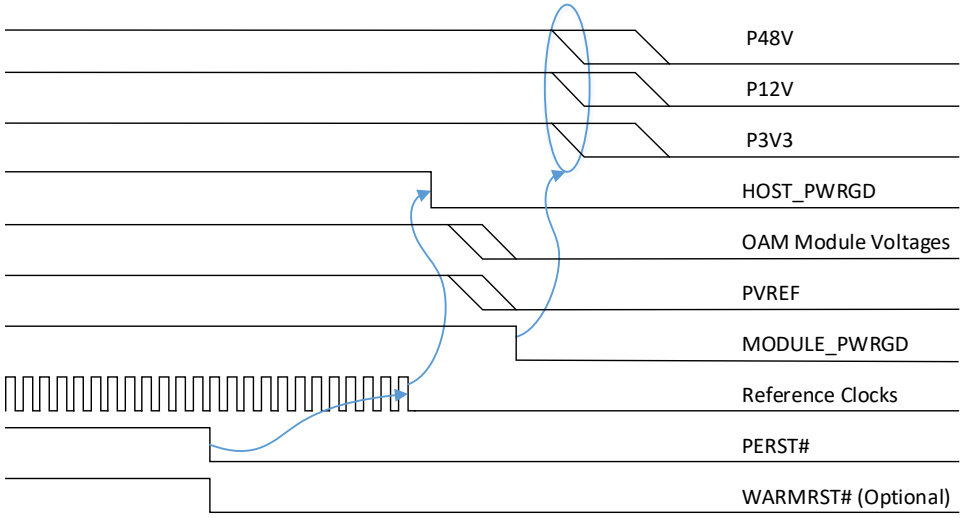


Figure 36 OAM Power-Down Sequence

Notes:

1) If the OAMs with UBB in the disaggregated host system design, the HOST_PWRGD is the UBB power good indication signal.
2) All voltages on the baseboard that OAM plugs into must be within specification before HOST_PWRGD is asserted.
3) HOST PWRGD is the enable signal to the voltage regulators on the OAM.

4) As the voltage planes on the module ramp up, the reference clocks from UBB will begin to run.
5) After all the voltages on the module are within specification, the module asserts MODULE_PWRGD to the baseboard.
6) The baseboard should tri-state all OAMs single-ended input signals before MODULE_PWRGD is asserted. Note that input signals required for the power sequence should be driven accordingly.
7) At least 100ms after MODULE_PWRGD assertion, the baseboard will de-assert the PCIe reset signal(PERST#) to the module.
8) The optional WARMRST# signal de-asserts simultaneously or later than the PERST# signal is de-asserted.

8.6 OAM Insertion Loss

The module interconnection channel total insertion loss from silicon die to mated connector should not exceed -8dB at 28GHz (and PCIe Gen5/6 @16GHz). The system integrator may contact the module supplier for details about the interconnection channel insertion loss and UBB design accordingly.

Total Tx or Rx loss on OAM + mated Mezz Connectors @28GHz (PCIe Gen5/6 @16GHz)	up to 8dB

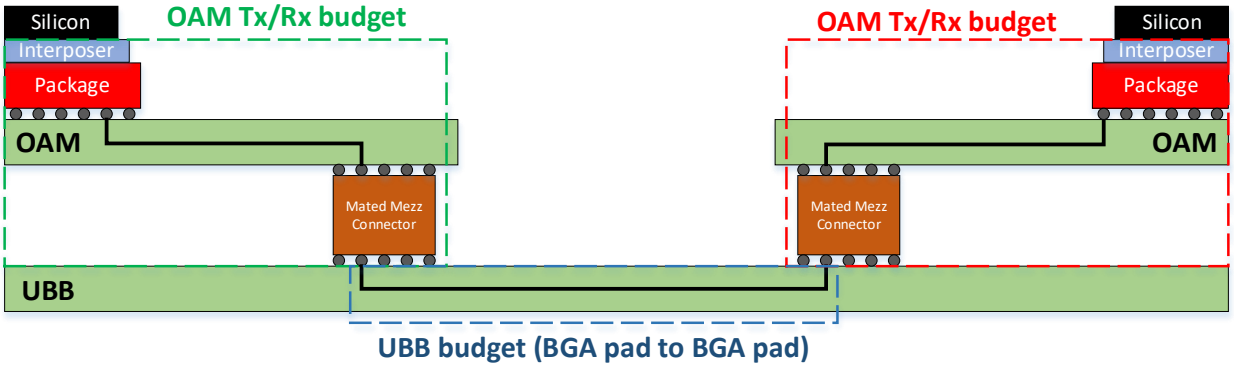


Figure 37 Channel loss Budget (OAM to OAM)

Total channel loss budget = OAM Tx + UBB + OAM Rx

Example:

Total channel loss budget = 30 dB (28GHz)

- **OAM Tx/Rx = 8 dB**
- **OAM Tx/Rx = 8 dB**
- **UBB budget = 30 dB – 8 dB – 8 dB = 14 dB**

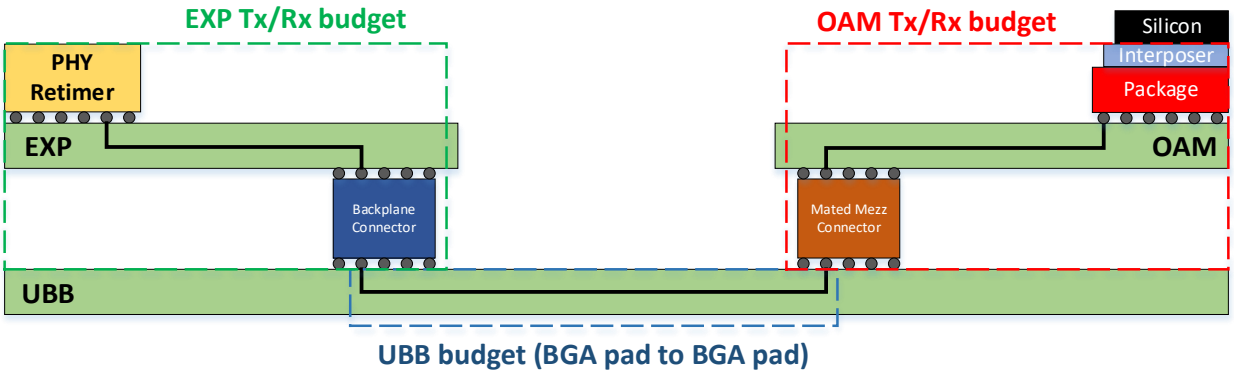


Figure 38 Channel loss budget (OAM to EXP)

Total channel loss budget = OAM Tx + UBB + EXP Rx

Example:

Total channel loss budget = 30 dB (28 GHz)

- **OAM Tx/Rx = 8 dB**
- **EXP Tx/Rx = 6 dB**
- **UBB budget = 30 dB – 8 dB – 6 dB = 16dB**

9 OAM Interconnect Topologies

This section describes the recommended interconnection topology for a system with 4, 8, and 16 OAMs.

9.1 Module ID

The following figure shows the MODULE_ID[4:0] strapping for the physical orientation of modules when 8 interconnected OAMs are used.

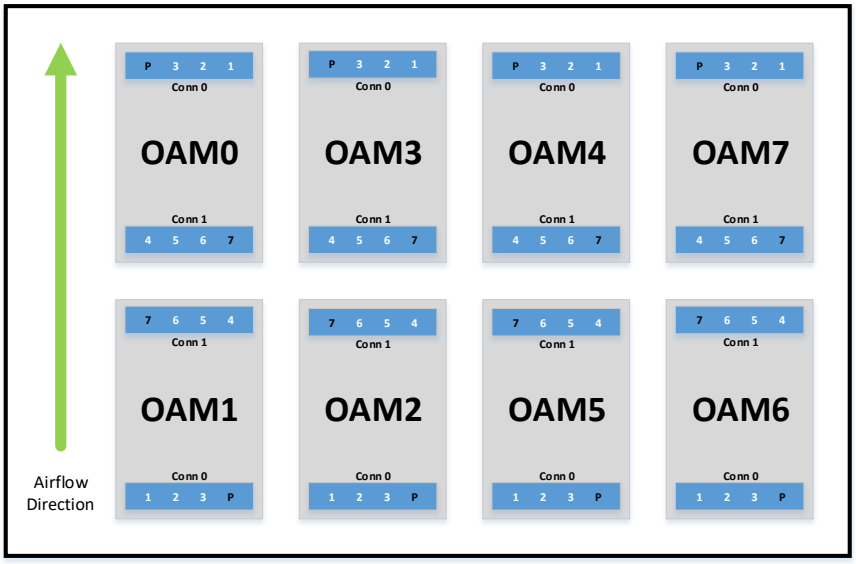


Figure 39 Required MODULE_ID[4:0] assignments for baseboards with 8 interconnected modules

Detail port to port assignment is based on system placement and routing length. Module to module interconnect may decrease to 4 ports if the module only supports 4. Module to module interconnect link may only utilize 8 lanes if the module defines 8 lanes per link.

The following Figure shows the required MODULE_ID[4:0] assignments when only 4 modules are connected as two rows of two.

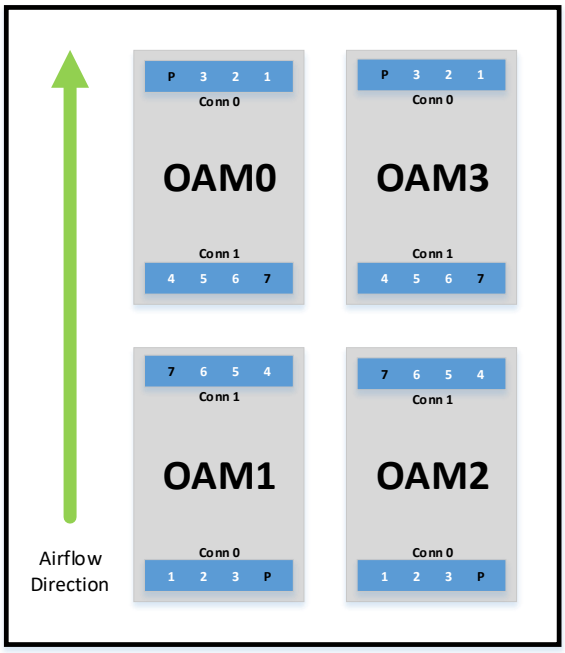


Figure 40 MODULE_ID[4:0] assignments when four only in two rows of two

9.2 Interconnect Topology

This section describes different interconnect topologies and routing guidance for eight OAMs with varying numbers of the port. If all seven ports are configured and routed as x16, there is no additional port(s) for expansion. To reserve expansion port(s), we suggest limiting onboard Interconnect links up to x8. 2nd half of port 1(x8, also referred to as 1H) is reserved for expansion by default. 2nd half of ports 4,6,7 are used in x8 based full connected topology at section 9.2.4.

9.2.1 Combined Fully Connected and 6-port Hybrid Cube Mesh Topology

For fully connect with expansion consideration, the UBB link is routed as x8 (1st x8 of each port, 1L-7L), leaving 2nd x8 of each SerDes port for expansion or embedding other topology. Here is 8-port HCM UBB reference board 7x8 fully connected topology combined with 6x8 hybrid cube mesh topology (x8 FC + x8 6-port HCM):

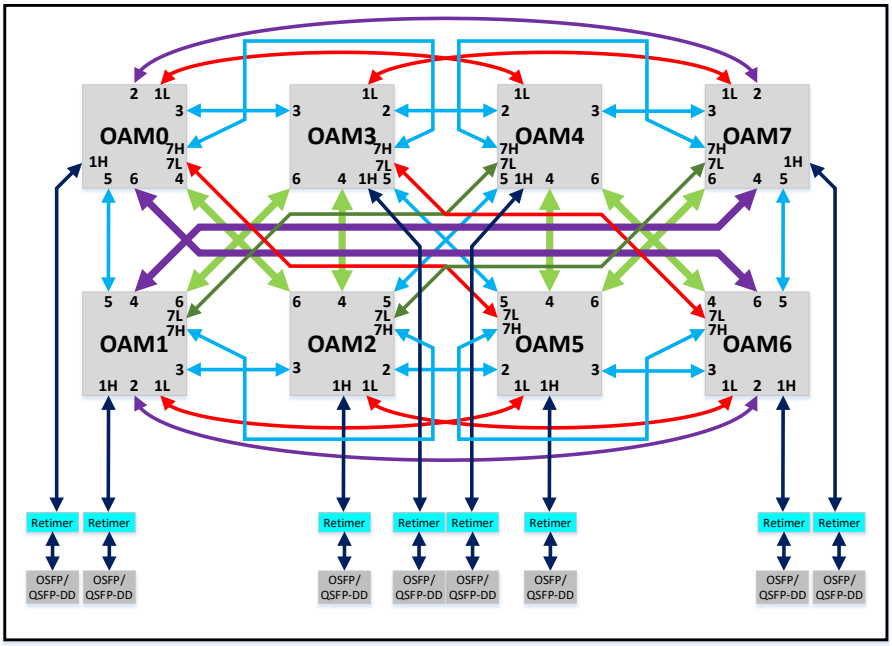


Figure 41 Combined FC/6-Port HCM Topology

Port 1,4,6,7 has a total of 16 lanes, and Port 2,3,5 has a total of 8 lanes in Figure 41:

- Fully connected: 7 x8 links using port 1-7 first x8(1L-7L).
- 2nd half of port 1s(1H) are connected to EXP or on board OSFP/QSFP-DD (scale-out).
- 6 port HCM: all 6 ports are in connector 1 only. x16 link for port 4/6, x8 link(5L) for port 5, and 2nd half of port 7(7H).

Below figure shows the detailed port mapping and routing guide:

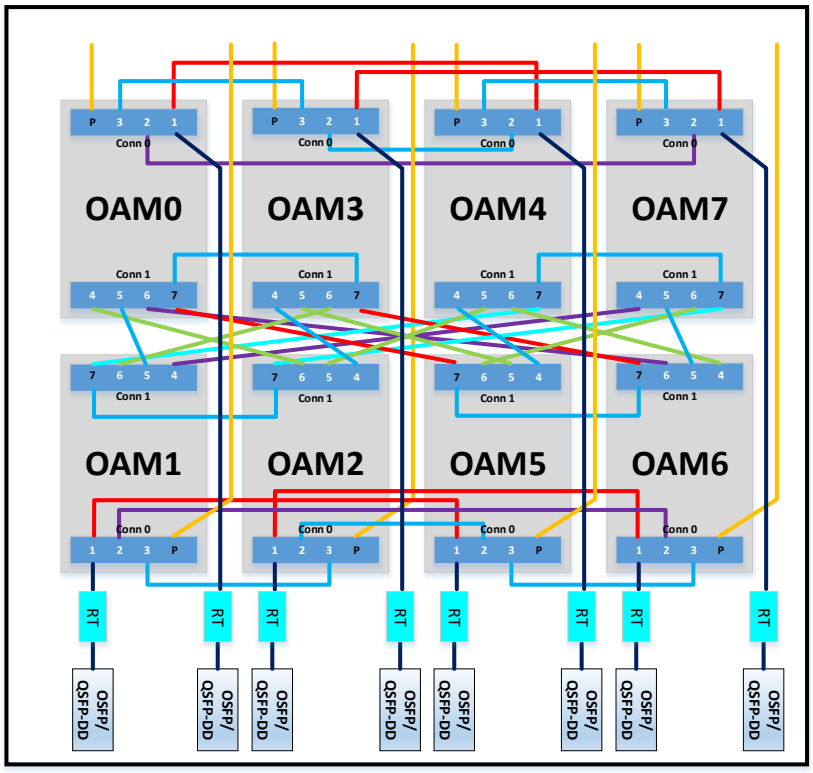


Figure 42 Detail port mapping and routing guidance

Port 4/6(both 4L/6L and 4H/6H), port 5L, 7H are used for 6x8 HCM. This is how it's embedded to this combined topology:

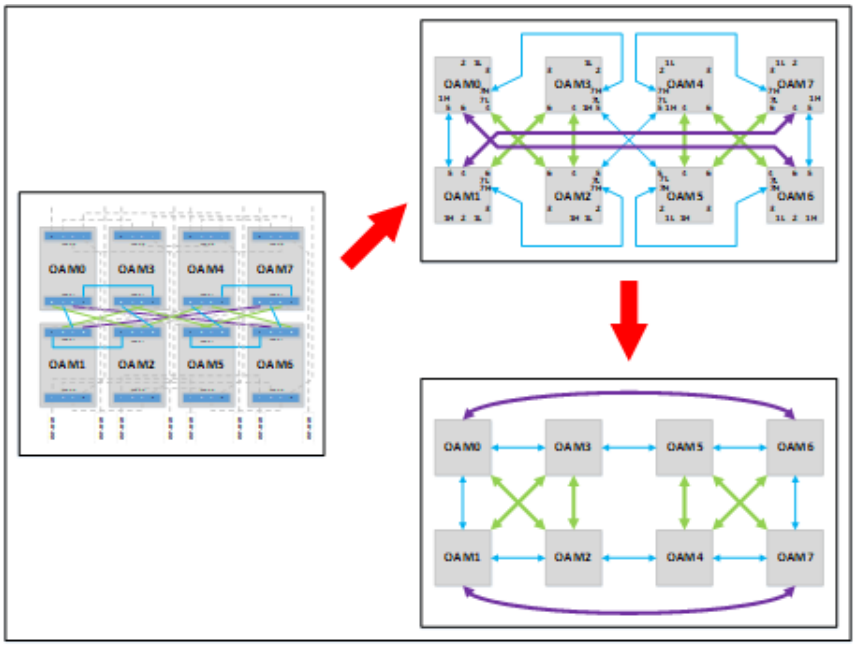


Figure 43 Embedded HCM Topology

9.2.2 8-port Hybrid Cube Mesh Topology

The Figure below shows an 8 port HCM(Hybrid Cube Mesh) topology of 8 modules in a UBB. Please follow port mapping to design OAM to be able to fit in the UBB. Port 4/6 connects through OSFP/QSFP-DD cables for a single 8 module system. These OSFP/QSFP-DD cables are also used for expansion (scale-out).

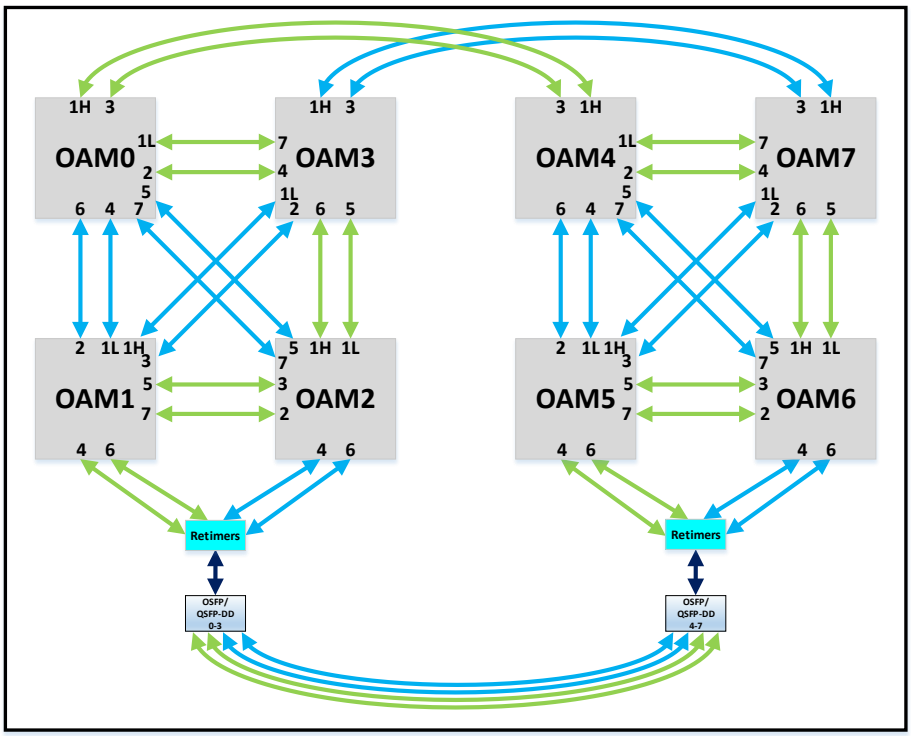


Figure 44 8-port Hybrid Cube Mesh Topology

- links HCM using links: 1, 2, 3, 4, 5, 6, 7.
- SerDes Port 2, 3, 4, 5, 6, 7 are x8 lanes.
- SerDes Port 1 is 2 x8 lanes.
- 1L –Lower 8-bit, 1H –Upper 8-bit.
- Links: 4, 6 (OAM #1, #2, #5, and #6) are used for scale-out, can be connected directly through OSFP/QSFP-DD cables or EXP Module retimers.

Here is the routing suggestion: total 4 layers, two layers for TX, two layers for RX. Port 4/6 are connected through cables.

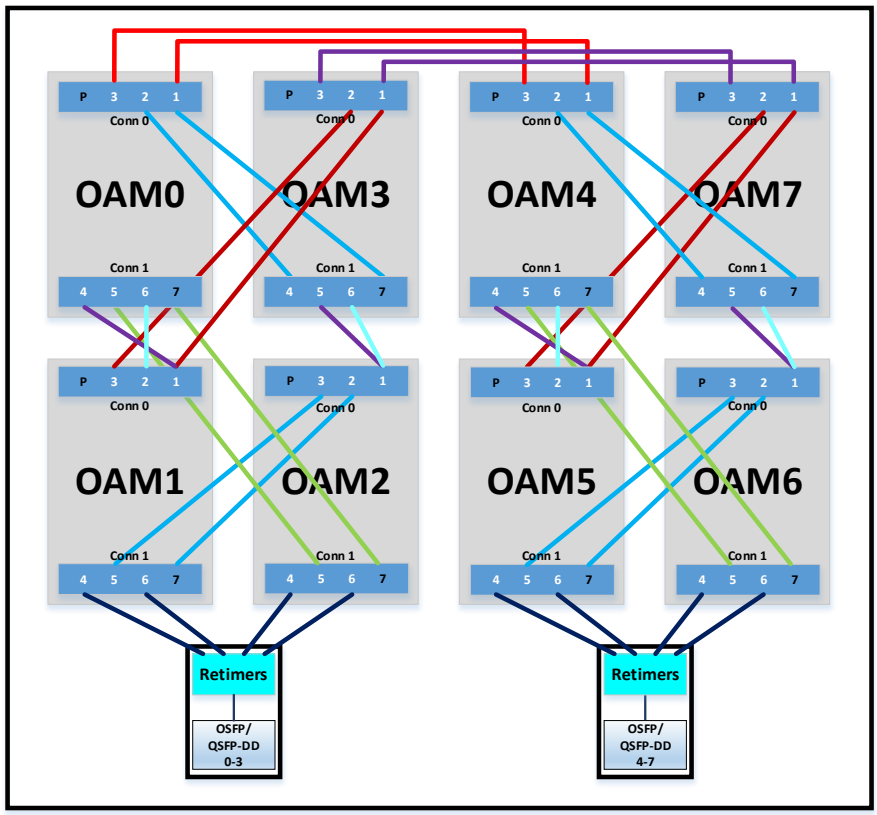


Figure 45 8-port HCM topology routing guide

9.2.3 11-Port (7+4) FC/Retimer Topology (Fully Connected + 4 Scale Out)

Figure 46 below shows an example of 11-port topology with seven SerDes ports for fully connected topology and four SerDes ports for scale out (per OAM) of 8 modules in UBB. The scale out ports are routed to EXP modules to support 32 scale out ports. Please follow port mapping to design OAM to be able to fit in the UBB.

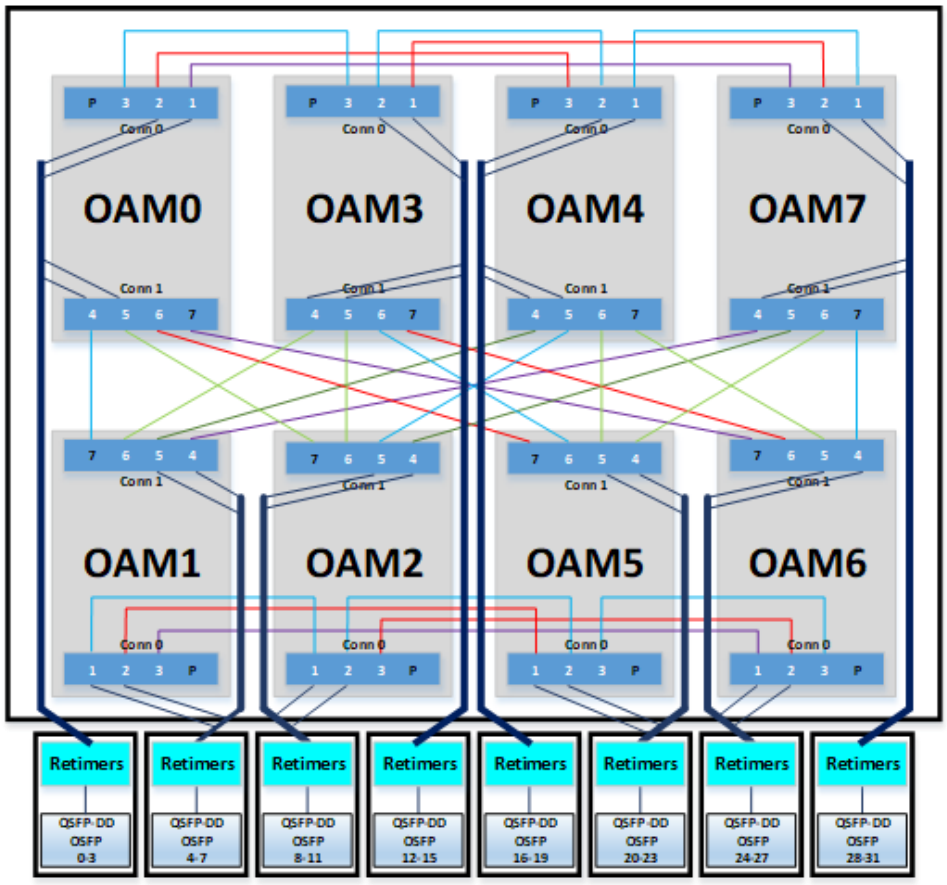


Figure 46 11-Port (7+4) FC/Retimer Topology

9.2.4 8-Port Switch-Based Topology

Figure 47 below shows an example of 8-port switch-based topology with eight SerDes ports of each OAM are connected to eight EXP modules to form fully connected topology through switch devices for scale up and out.

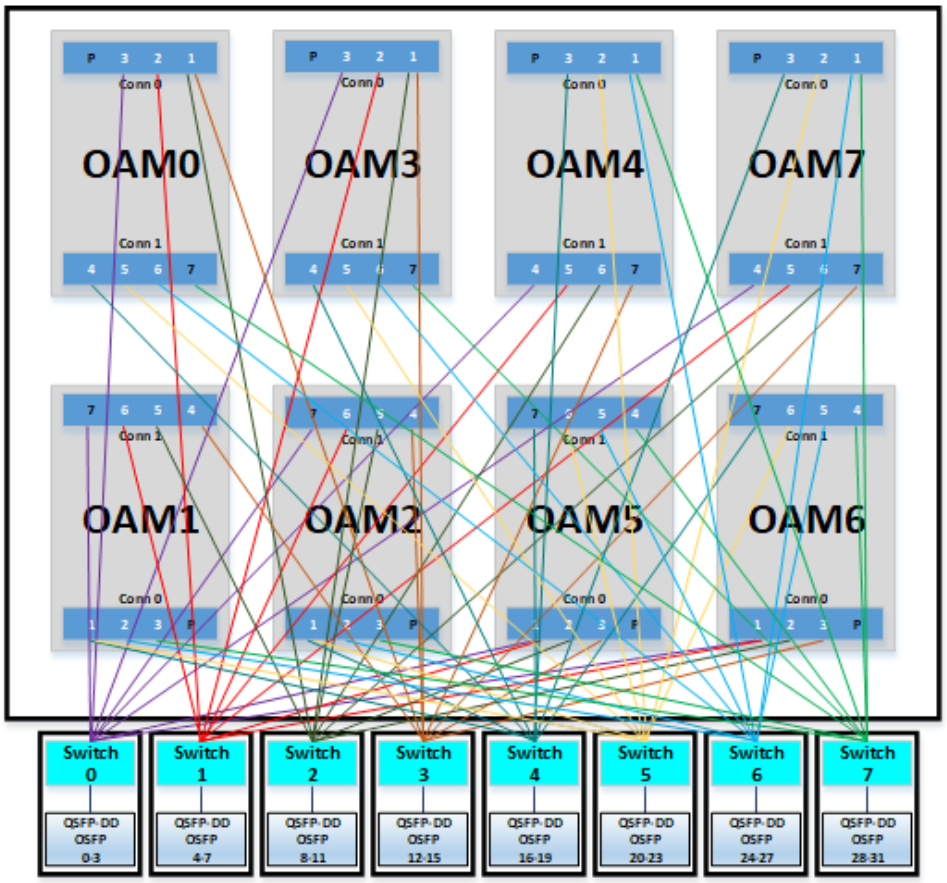


Figure 47 8-Port Switch-Based Topology

9.3 LINK_CONFIG[4:0]

The 5 link configuration strapping bits are pulled up on modules that use them. These bits are strapped to the ground on the baseboard to select logic 0 or left floating to select logic 1. Some OAMs use these LINK_CONFIG[4:0] strapping bits to determine the interconnect topology for the links between modules and determine the protocol of the "P" Link.

Encodings not listed in the table below are currently un-defined.

Table 7 LINK_CONFIG[4:0] Encoding Definitions

LINK_CONFIG[4:0]	Definition
00000	Reserved for OAM. Test use by OAM Vendor.
xxxx0 (except for 00000)	Indicates the "P" link is PCIe.
0001x	Combined FC/6-Port HCM Topology as in Figure 41.
0010x	8-Ports HCM Topology as in Figure 44.
0011x	11-port (7+4) FC/Retimer Topology as in Figure 46.
0100x	8-port Switch-Based Topology as in Figure 47.
0101x – 1xxxx	RSVD
xxxx1 (except for 11111)	Indicates the "P" link is an alternate protocol other than PCIe.
11111	Indicates an alternate means for identifying the link interconnect topology and configuration is used.

9.4 OAM Interconnect PCB Topology

It is recommended to implement blind via or back drill with stub length less than 10mils. System integrators should carefully work with module suppliers to plan the PCB routing and address the signal integrity concern.

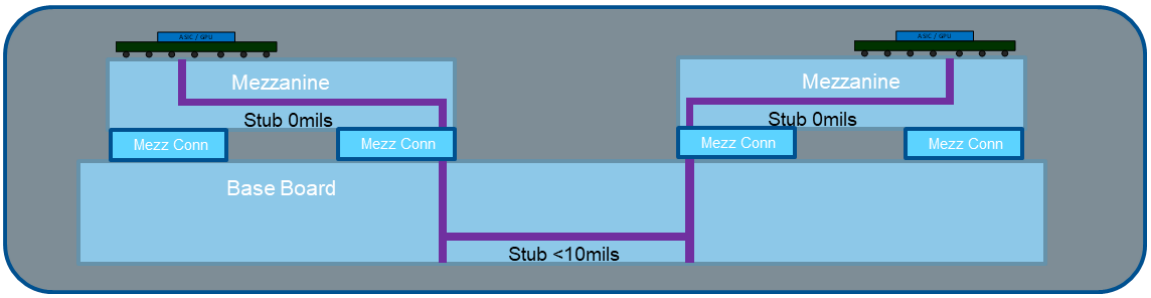


Figure 48 OAM Interconnect PCB Topology

10 OAM reference system design

This section gives a system design concept as a reference. Figure 49 Reference System Design shows 8 OAM modules. The plastic top provides a 0.5mm bumper on each side of the 102mm width OAM, and the 1mm gap between each module assembly provides rough alignment, guidance, and Keying as described in Section 6.6. An air baffle is designed into the 33.8mm space to prevent air bypass in the system. Note that the front and rear rows are oriented 180 degrees opposite, as indicated in Section 9. Recommended Alignment Features

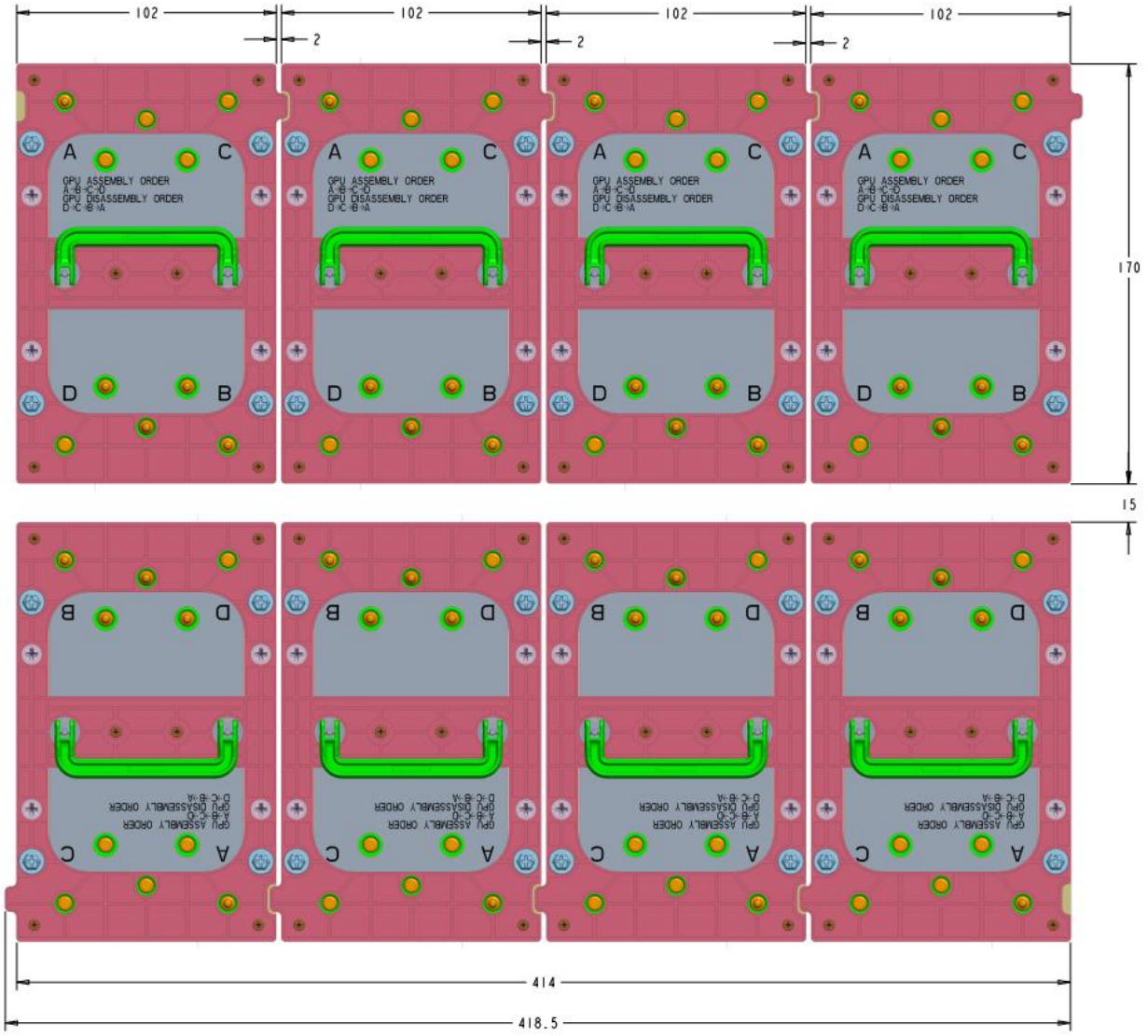


Figure 49 Reference System Design

11 OAM Management and Security Requirements

This section describes a common set of management and security requirements for OAM.

11.1 Management Interface

The OAM sideband management interface is used by a Management Controller (MC) or Baseboard Management Controller (BMC) to communicate with the OAM. Table 8 below summarizes the sideband management interfaces.

Table 8 Sideband Management Interface

Requirement	Voltage Level	Required
I2C/I3C/SMBus 2.0 compliant physical interface (Master/Slave) *	3.3V	Yes
JTAG physical interface(primary)	V_{ref}	Yes
I2C/I3C/SMBus 2.0 compliant physical interface (Master)	V_{ref}	Optional**
JTAG physical interface(secondary)	3.3V	Optional

Note:

*When MCTP over SMBus is used, the BMC shall support both master and slave modes.

**The master I2C/I3C/SMBus physical interface is required for scale-out, baseboard FRU access.

The OAM communicates with the Baseboard Management Controller (BMC) by using:

- SMBus:
  - SMBus supporting 1MHz mode is preferred.
  - Standard Intelligence Platform Management Bus (IPMB) and Intelligent Platform Management Interface (IPMI) commands.
  - SMBus ARP protocol.
  - Management Component Transport Protocol (MCTP) over SMBus binding. (DMTF DSP0237)
- UART (optional)
  - Shall support 115200 Baud Rate.
  - For serial console access.
- JTAG
  - For register dump, memory dump, debug access.
- PCIe (optional)
  - To support MCTP over PCIe binding.

11.2 Sensor Reporting

An OAM module may have several silicon components, including one or more ASICs implementing acceleration functions. It is crucial that connected sensors of these components (voltage,

current/power, temperature, etc.) can be retrieved over sideband interfaces for system management. The sensor reporting interface will only be accessible in the main power mode (S0).

Error! Reference source not found. Summarizes the sensors reporting list: The report from these sensors improves the system monitoring/management and allows the baseboard management device to access the critical components on the module. It is recommended that the voltage/current/power sensor reporting accuracy is within ± 2%, and the temperature reporting accuracy is within ±3°C.

OAM Module shall support Sensor discovery via IPMI or Platform Level Data Model (PLDM) for Platform Monitoring and Control (DMTF DSP0248). MC will follow mechanisms specified in DSP0248 to discover sensors supported by the OAM module and their threshold.

11.3 Error Monitoring/Reporting

System Management Controller (MC) or Baseboard Management Controller (BMC) shall be able to monitor and access the OAM(OAM supplier to specify how to access) through PLDM over MCTP as needed to set thresholds, clear status, determine error counts and syndromes (SW driven interrupt or an Alert pin), and identify error sources/Syndromes, etc.

OAM shall support Platform Level Data Model (PLDM) for Platform Monitoring and Control (DMTF DSP0248) for error reporting, enabling the module to define state sensors and events for health monitoring and reporting.

Table 9 Error state sensors

Sensor List	Remark
Overall OAM health status	It can be vendor-specific
Memory Error event	
Host bus error event	
Device health event	
Interconnect Link events	
Scale-out link Event	for the scale-out bus, if the different buses from the interconnect links
Other custom/Vendor-specific	

11.4 Firmware Update

The OAM should support secure boot. Detail requirements, please refer to section 11.8.1.

OAM supplier shall provide an in-band FW update utility to perform the firmware update.

For out-of-band firmware updates, OAM shall support Platform Level Data Model (PLDM) for firmware update over MCTP as specified in DMTF DSP0267.

Update failure and failure types shall be communicated to BMC as specified in DSP0267.

11.5 Power Capping

Module supplier should provide the utility for in-band power capping.

11.6 FRU Information

System Management Controller (MC) or Baseboard Management Controller (BMC) shall access related internal registers to get module information (see Table 10.

FRU shall be accessed through Platform Level Data Model (PLDM) for FRU Data, following DMTF DSP0257.

Table 10 FRU Information

FRU Info	Remark
Manufacturing Date	
Manufacturer	
Product Name	
Serial	
Part Number	
FRU ID	
Version	
Asset Tag	
Firmware Version	
OAM Spec Version	
Input Power Mode	
OAM TDP	
SerDes Link Speed	
Custom Data 1	
Custom Data 2	
Custom Data 3	
Custom Data 4	
Custom Data 5	
Custom Data 6	

11.7 IO Calibration

The system shall be able to get DDR/PCIe/interconnect training status and margin information. For in-band access, specific tools and/or API shall be provided by the OAM supplier.

11.8OAM Security Requirements

11.8.1 Secure Boot

The device must support a secure boot. There are multiple requirements to implement secure boot. Here is a brief list:

- Hardware-based Root of Trust:
  - Immutable Root-of-Trust (e.g., OTP) required for provisioning asymmetric RoT key and other security-related critical information.
  - ROM code for minimum support needed for Secure Boot
- Boot Time Verification:
  - On every boot, the ROM code should cryptographically verify mutable firmware code using the asymmetric RoT public key.
  - All mutable code should be verified by signature authentication before allowing it to execute.
  - ROM code patching is not permitted in a production device.
  - In case of failures, please refer to the 'Recovery' section for more details.
- TOCTOU Attack Protection:
  - The attacker should not be able to modify the firmware image (Time of Use) after the signature of the firmware is verified (Time Of Check) during the boot process
- Anti-Rollback Protection:
  - The device must support Security Version Number (SVN) in the immutable memory to protect against downgrade attacks
- Key Revocation or Change of Ownership:
  - The device must support Secure revocation of Intermediate key (used for signing of firmware) in case of the Key compromise
  - For a change of ownership, the device should either support revoking the ownership key or rotating the ownership certificate (as suggested in DSP0274)
- Secure Firmware Update:
  - Signature generation for firmware payload using asymmetric RoT private key is necessary for secure boot
  - The device should be able to authenticate the signed FW payloads before booting up using the payload

For further details on implementation and a detailed set of requirements, please refer to the OCP Hardware Secure Boot document for more information.

11.8.2 Recovery

OAM must support a recovery mechanism to restore the mutable firmware code to a state of integrity if any such firmware code or critical data are detected to have been corrupted or forced to recover through an authorized mechanism.

- OAM should also support two mutable firmware (active and recovery) regions where the recovery firmware is the previously known good image.
- In case of failure of redundant copies, OAM should support recovery over the sideband interface.
    - OAM shall support Platform Level Data Model (PLDM) firmware update over MCTP as specified in DMTF DSP0267.
    - Update failure and failure types shall be communicated to BMC as specified in DSP0267.
    - Firmware updated out-of-band should still follow the boot-time verification process of secure boot.

For further details on implementation and a detailed set of requirements, please refer to the OCP Recovery document for more information.

11.8.3 Debug Capabilities

Any intrusive debug capabilities (read/write memory, general-purpose register contents, alter control flow), e.g., JTAG, UART, must be disabled for remote access. If needed, they should only be re-enabled via physical access or using cryptographically authorized tokens.

11.8.4 Attestation

It is critical to verify the firmware running on the device dynamically and the device itself cryptographically; this helps establish trust between the devices. It is recommended to support device attestation for OAM.
Here is a short list of generic requirements to support the attestation for OAM:

- Keys, seeds, and device identifiers
- Provisioning Facility (Initial Provisioning Environment Operations and Equipment)
- Device Ownership Provisioning
- Authentication, Attestation, and Enrollment protocol
- Measurement collection and storage

For further details on implementation and a detailed set of requirements, please refer to the OCP Attestation for System Components v1.0 for more information.

12 Environmental

12.1 Environmental Requirements

The OAM shall meet the following environmental requirements:

- Gaseous Contamination: Severity Level G1 per ANSI/ISA 71.04-1985
- Ambient operating temperature range: 5°C to +35°C
- Operating and Storage relative humidity: 20% to 90% (non-condensing)
- Storage temperature range: -20°C to +70°C
- Transportation temperature range: -55°C to +85°C (short-term storage)
- Operating altitude with no de-ratings: 3048m (10000 feet) – recommended as this is a Meta spec and standard for Telco operation.

12.2 Regulation

The vendor needs to provide CB reports of the OAM, which are required to have rack-level CE. The OAM should be compliant with RoHS and WEEE. The PCB should have a UL 94V-0 certificate.

13 Version History

Author	Description	Version	Date
Song Kok Hang (Meta) Ahmed Abou-Alfotouh (AMD) Cheng Chen (Meta) Anthony Chan (Meta) Xinxin Wang (H3C) Ash Liao (Wiwynn) Nick Wang (Wiwynn)	Initial Release	0.75	03/03/2023