Co-Packaged Optic Assembly Guidance Document

Table of Contents

Table of Figures

Table of Tables

1. Introduction

1.1. Collaboration Agreement

Before providing feedback to the CPO JDF, please ensure that a representative from your company has reviewed, signed and returned the collaboration agreement document.

1.2. Scope

This document provides guidance on the requirements for co-packaged optic assemblies designed for high-radix, network switch applications with 100Gb/s electrical interfaces.

1.3. Introduction

The CPO JDF plans to release three documents focused on different elements of Co-Packaged Optics (CPO): the optical module, the External Light Source (ELS), and the CPO assembly (covered here).

This collection of documents is intended to provide guidance to vendors pursuing Co-Packaged Optics (CPO). The first revisions are intended to facilitate structured conversations about the different elements of CPO. The documents will be revised periodically as input are collected from collaborators.

1.4. Common Terms

The following terms and acronyms are used in this document.

- CPO – Co-Packaged Optics
- 400GE – 400 Gigabit Ethernet
- 400G FR4 – 400GE optical standard utilizing 4 wavelengths on a 1310nm CWDM grid with each wavelength transmitting 106.25Gb/s using 53.13Gbaud PAM4 modulation
- 400G DR4 – 400GE optical standard utilizing ribbon fiber with 4x106.25Gb/s transmit and 4x106.25Gb/s receive lanes/fibers using 53.13Gbaud PAM4 modulation. Single wavelength range, centered at 1310nm, is specified for all lanes
- 200G FR4 - 200GE optical standard utilizing 4 wavelengths on a 1310nm CWDM grid with each wavelength transmitting 53.13Gb/s using 26.56GBaud PAM4 modulation
- CWDM 4 – 100Gb/s optical standard utilizing 4 wavelengths on a 1310nm CWDM grid with each wavelength transmitting 25.78Gb/s using NRZ modulation
- 100G PSM4 MSA – 100GE optical standard utilizing ribbon fiber with 4x25.78Gb/s transmit and 4x25.78Gb/s receive lanes/fibers using 25.78Gb/s NRZ modulation. Single wavelength range, centered at 1310nm, is specified for all lanes
- QSFP-DD –Quad Small Form-factor Pluggable Double Density used for 400GE
- BER – Bit Error Rate
- CEI-112G-XSR-PAM4 – Common Electrical Interface specification operating at a serial data rate of 112Gb/s over short reaches common in multi-chip modules
- CMIS – Common Management Interface Specification
- ESD – Electro-Static Discharge
- SPI - Serial Peripheral Interface
- LD – Laser Diode used as a CW source for the co-packaged optical module
- ELS – External Laser Source

1.5. References

The following specifications are referenced in this document.

- IEEE 802.3bs – Media Access Control Parameters, Physical Layers and Management Parameters for 200 Gb/s and 400 Gb/s Operation
- IEEE 802.3cu – Media Access Control Parameters, Physical Layers and Management Parameters for 100 Gb/s and 400 Gb/s over SMF at 100 Gb/s per Wavelength (currently in draft)
- 100G CWDM4 MSA Technical Specifications Rev 1.1
- OIF-CEI-x.x (CEI-112G-XSR currently in draft)
- CMIS revision 4.0 – Common Management Interface Specification

2. Absolute Maximum Ratings

The values in Table 1 below are for guidance and subject to change.

Parameter	Symbol	Min	Typical	Max	Unit	Note
Storage temperature	T_st	-40		85	°C	
Relative humidity, storage and transportation	RH	5		95	%	Noncondensing
ESD, electrical pins	ESD	-1		1	kV	Human body model
Power supply voltage		-0.3		Vtyp + 10%	V	

Table 1: Absolute maximum ratings

3. Operating Conditions

See Table 2 below for a description of the operating conditions for the CPO assembly.

Parameter	Symbol	Min	Typical	Max	Unit	Note
Power supply voltage[1]	Vcc1		TBD		V	Host will supply multiple voltage rails to optical module. Min/max +/- 5%
	Vcc2		TBD			
	Vcc3		TBD			
	Vcc4		TBD			
	Vcc5[2]	11.5	12	12.5		
Optical module operating case temperature	Tcase	20		70	°C	With integrated LDs
		20		80		With external LDs
Switch IC operating junction temperature	Tj	20		105	°C	
Power consumption of optical modules	Poptic			1024	W	51T capacity, 8W/400G
				960		Same as above using ELS
Power consumption of switch IC	Pswitch			900	W	Preliminary estimate for 51T switch, XSR serdes
Steady state current	Icc1			TBD	A	See note 1 below for Icc5
	Icc2			TBD		
	Icc3			TBD		
	Icc4			TBD		
	Icc5			150[3]		
Relative humidity	RH	5		85	%	Non-condensing

Table 2: Operating conditions

Notes:

1. Number of voltage supply rails and rail voltages shown are TBD pending input from collaborators. The intent is to provide multiple voltage rails to minimize the need for voltage regulation on the optical module.
2. 12V rail is for the main digital supply and will be bucked-down and regulated on the CPO assembly. We are seeking guidance on the preferred DC level for this rail.
3. Assume main digital supply for switch and optical modules (~0.7V) will be supplied to the CPO assembly at 12V and regulated down as needed on the CPO assembly. This is required to manage the total current through the CPO assembly's DC electrical connector. 150A estimate is based on 1.6kW max power on main digital rail and 90% min regulator efficiency.

4. Voltage Supplies

4.1. Supply Overview

The intent is to keep the CPO assembly as simple as possible and build the complexity required into the main switch PCB and CPO optical module. With this in mind, it is proposed that the Switch Main Board (SMB) supply the voltage rails/regulation required for the CPO assembly. In order to simplify the design, the number of rails should be kept to the minimum required.

In addition to minimizing the number of voltage levels, it will be important to limit the number of voltage domains within a given voltage level and minimize the amount of filtering/bypassing needed on the CPO substrate. There will be little PCB space for filtering/bypassing, mostly limited to the area on the back-side of the substrate under the switch IC and high-speed routing. Switch IC and optical module vendors should consider this in designs intended for CPO applications.

It is expected that the main digital voltage rail (~0.7V for 16 & 7nm CMOS) will have to be regulated on the CPO assembly, otherwise the current requirements of the CPO assembly DC connector would be excessive. Here we propose supplying the main digital voltage to the CPO assembly at 12V and regulate down as needed on the CPO assembly. This should not be required for other, lower current, voltage rails. We are seeking guidance if 12V is the preferred DC level for the main digital supply voltage to the CPO assembly.

5. Electrical Specifications

5.1. SMB Connector

The CPO assembly will be connected to the Switch Main Board (SMB) using an electrical connector. This connector does not need to be part of the high-speed data path in/out of the switch IC. As such, it is possible to minimize the complexity of this connector and keep the pin count low (~500 positions). If a switch design requires part of the data path to be electrical, and it is desired that the electrical data path passes through the SMB connector, a high-speed connector or interposer with a higher pin count can be used.

The SMB connector will provide the required power supply connections from the SMB to the optical modules and switch. In addition, the electrical connector will be used for control and signaling to the switch (PCIe, I2C & Ethernet), control and signaling to the optical modules (SPI), reference clock delivery (optical module and switch IC), and status/control pins for the switch and optical modules

Figure 1 below shows a sample embodiment of a CPO assembly using two SMB connectors each having a pin count of approximately 260 positions. This figure is provided as an example and is not intended to represent the preferred embodiment.

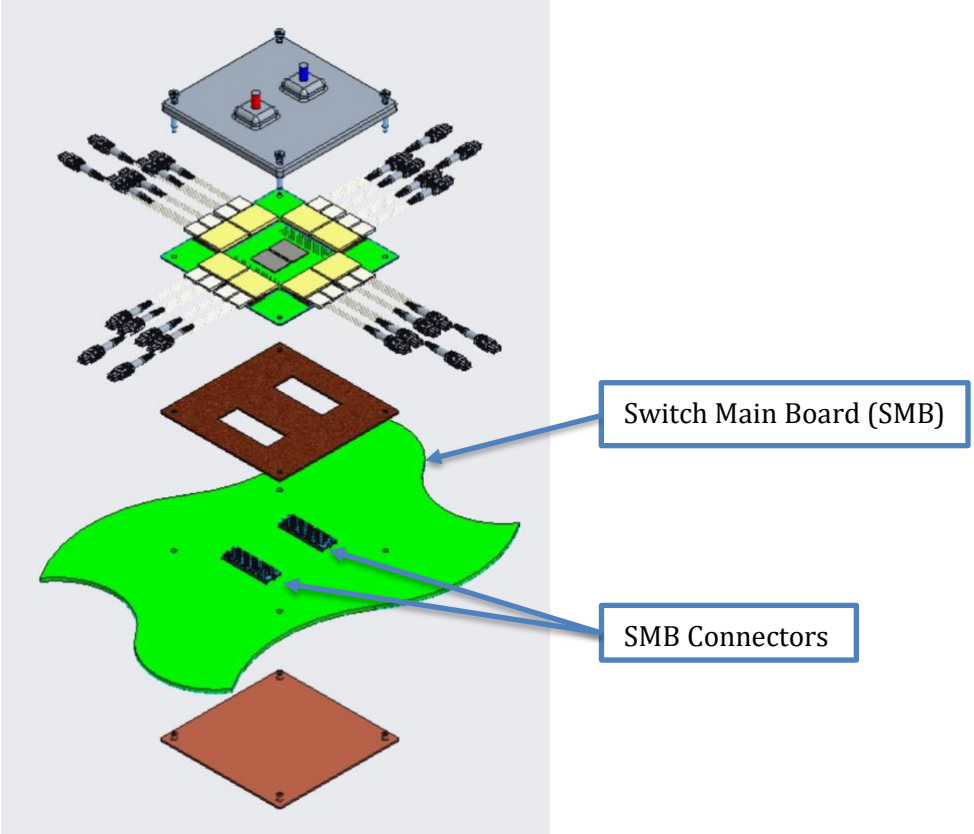

Figure 1: Exploded view of CPO assembly showing bottom-side electrical connectors

5.2. Low-Speed Pins

Low-speed pins used for signaling and control will pass through the SMB connector. See the Optical Module Guidance document for low-speed pins and signaling for the optical modules. The switch IC will also have a small number of low speed pins for control and signaling which will vary by switch IC.

The 12V regulators on the CPO assembly may have enable pins which are accessed through the low speed connector to control enable/disable.

5.3. Switch IC to Optical Module Electrical Interface

The high-speed electrical interface between the switch IC and the optical modules will be CEI-112G-XSR. This project is still active within the OIF. The expectation is that CEI-112G-XSR will support PCB losses of 10dB at 28GHz.

5.4. Reference Clock

Optional reference clock signals can be provided to the optical modules from the switch main board. Preliminary guidance is a maximum of one reference clock per 800Gb/s of capacity, or a maximum of 64 clock signals for a 51Tb/s CPO assembly. The reference clocks will be generated/broken out on the main switch board. The CPO connector and substrate must support delivery of the reference clock signals from the main switch board to the optical modules. It is also possible to implement fan-out buffering of the reference clocks on the CPO when a large number of reference clocks are required by the optics.

The switch IC will require reference clocks as well as listed in Table 3. The exact number of reference clock and clock specifications will vary by the specific switch IC used.

5.5. High-Speed Electrical Connector

In some cases, high-speed-data-path electrical connectors will be required. For example, on a modular switch with a back plane, it may be desired to have high-speed electrical connections from the switch IC to the back plane. In such a case, the CPO assembly will have high-speed optical connections to the front panel of the modular switch card, and high-speed electrical connections to the back plane.

Figure 2 below shows an embodiment of a CPO assembly with high-density, high-speed connectors on a portion of the switch IC's high-speed IO. It is also possible to drive the high-speed-data-path electrical signals through the SMB connector. For such designs, it may be required to re-time/re-gen the XSR signals on the SMB.

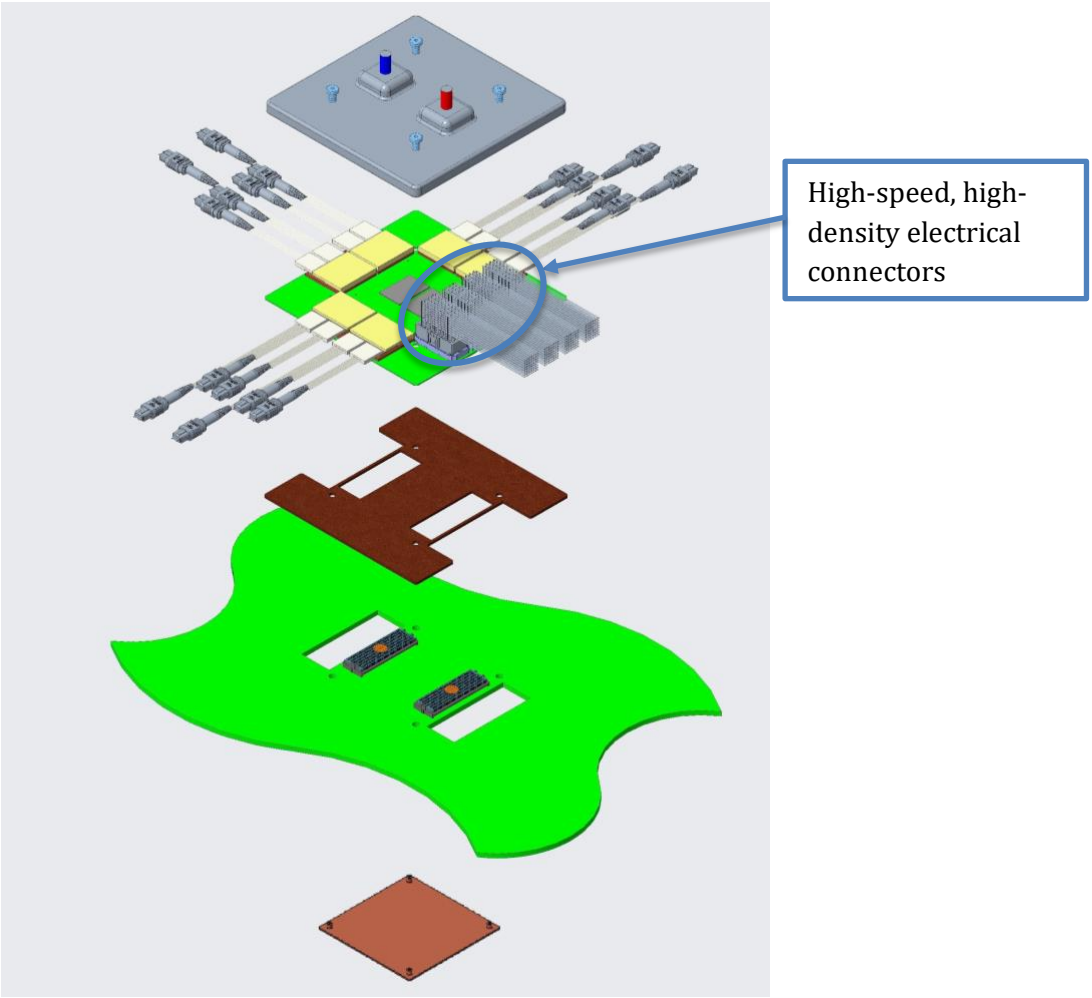

Figure 2: Exploded view of CPO assembly with optical and high-speed electrical connectors

6. Management Interface

CPO assembly will incorporate two separate management interfaces; one for the optical modules and for the switch IC. See below for details.

6.1. Optical Module Management Interface

The management interface proposed is the same specification used for 400G pluggable modules, QSFP-DD Common Management Interface Specification (CMIS). The current CMIS covers 16 lanes and there is a desire to extend this to 32 for applications like CPO. Optical modules with >32 lanes will require multiple instances of CMIS. For the electrical interface for CMIS on CPO, we are proposing one Serial Peripheral Interface (SPI) for each CMIS instance. Reference the CPO JDF's Optical Module Guidance Document for details.

6.2. Switch Management Interface

The switch IC management interface requirements will vary by the switch IC used. Some general guidance is captured below.

The switch IC is primarily managed by the CPU on the SMB through the PCIe interface. The switch's internal registers and tables are accessible through the PCIe interface, and it also supports packet transfer between the CPU and the switch. In addition, some switch ICs may have a management Ethernet interface for packet transfers to/from the CPU.

Table 3: Switch IC clocks and management interfaceTable 3 below lists the typical control pins and management interface for the switch silicon. Note that the quantity refers to the number of interfaces, not the pin count.

Signal/Interface Name	Qty.	Direction	Description
Reset	1	Input	Chip reset input
PCIe reset	1	Input	PCIe interface reset input from the root complex
PCIe data	x4	I/O	CPU management interface to the switch silicon. Can be Gen 3 or Gen 4. Preferred to be x4 width
PCIe reference clock	1	Input	100MHz differential clock
Core reference clock	1	Input	Differential clock
Serdes reference clock	4 or 8	Input	See section 5.4
Hardware timestamping reference clock	1	Input	Differential clock
Core voltage AVS interface	1	Output	The interface tells the external circuit the voltage level the core voltage should be at
SPI	1	I/O	The interface connects to an off-chip SPI flash that stores the initialization firmware of the switch silicon
Management ethernet	1	I/O	10GBASE-KR interface for switch management
On-die temperature sensor interface	1 or more	Output	The interface(s) provides the on-die temperature information to the external circuit for system thermal management

Table 3: Switch IC clocks and management interface

6.3. Other Management Interfaces

Other CPO assembly components or subsystems may require management interfaces. For example, the voltage regulators on the CPO assembly may need control/status pins and a management interface like I2C.

7. Mechanical Considerations

7.1. Mechanical Overview

The CPO assembly consists of a high-density organic substrate, switch IC and optical modules arrayed around the perimeter. An example of a CPO assembly is shown in Figure 3 for reference; it is not intended to represent the preferred embodiment. Note that the switch IC shown below is represented as two dies. This is just an example to show that future switch ICs may not be monolithic and be composed of several chips that together act as a single device.

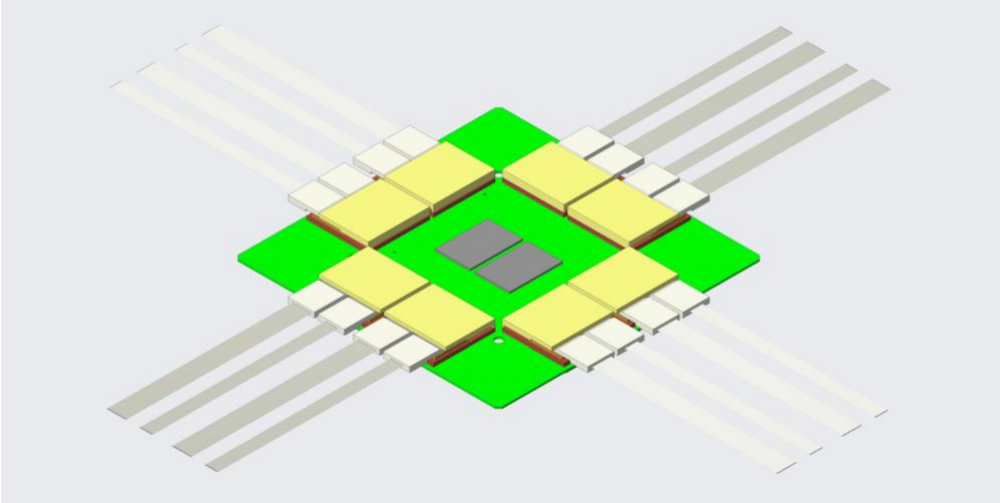

Figure 3: Example embodiment of a 51Tb/s CPO assembly

The maximum size of the organic substrates available will evolve over time; we anticipate a range of 100 to 150mm per side for first-generation CPOs. A major benefit of CPO is lower power. As switch SERDES speed and radix increase, minimizing the distance between the switch and optical module electrical interfaces will be critical for reducing power consumption.

The optical modules may be directly soldered to the CPO substrate or attached using a high-speed, LGA connector. The optical modules must have a heat spreader on the top surface for mating to a heat sink or cold plate. More details can be found in the CPO Optical Module Guidance Document.

7.2. Optical Connectors

The optical connectors provide a means to connect the optical module to the front plate of the host switch. The optical module may be pigtailed or have an integrated optical connector.

The pig-tailed version would have SMF ribbons for the Tx and Rx terminated in an optical connector. For example, a 16x 400GBASE-FR4 (6.4Tb/s) module would have 16x fibers for the Tx and 16x fibers for the Rx functions. These fibers could be terminated in any connector arrangement desired on the front panel. The most common arrangement today is duplex LC connectors.

For a module with an integrated optical connector, an optical patch cord would be used to connect the optical module to the faceplate of the switch. High-density optical connectors available today (e.g. MPO) are not ideal for CPO applications. It is likely that new optical connector scheme will need to be developed to make this approach practical and the JDF is interested in reviewing design proposals.

For optical modules with high lane counts using a single-fiber/lane optical standard, there will be a large number of fibers attached to the optical module. For example, a 16x 400GBASE-DR4 (6.4Tb/s) module would have 128 fibers attached for Rx and Tx functions. For high-fiber count applications, dense fiber connection schemes (e.g. MCF or 3D waveguides) are desired and the CPO JDF is interested in proposals on how to address such cases.

7.2.1. Fiber Connections for External Laser Source

Some CPO designs will include an External Laser Source (ELS) to provide a CW light source to the optical modules. In such cases, each optical module will require a minimum of one fiber per wavelength connected to the ELS module. The number of ELS fibers required will depend on the optical module requirements and the power/fiber launched from the ELS. For most SiPho-based optical modules, PM fiber will be required between the ELS and the optical module.

For example, a CPO design with eight 6.4Tb/s optical modules (using 400GBASE-FR4) would require ELS's with 32 fibers, with four fibers routed to each optical module. In the example shown in Figure 2 , the Tx side of the optical module has a 24-fiber array, in which eight fibers (two per wavelength) are routed to the ELS module and the remaining 16 are used for the Tx function.

7.3. Thermal Considerations

Components to consider when designing a cooling solution for the CPO assembly are (but not limited to) – switch IC, optical modules and main digital voltage rail regulators. The following paragraphs outline the thermal considerations for each.

Depending on how the switch IC is packaged as part of the CPO assembly (bare-die vs. lidded), the cooling solution requirements may vary (for example – acceptable loading force). In either case, TIM (thermal interface material) with a low thermal resistance is desired to counteract the relatively higher power density of the switch IC compared to the optical modules (based on the example embodiment). The entire thermal stack (packaged switch IC, TIM and cooling solution) will need to be optimized to ensure adequate thermal margins under different operating conditions. The CPO JDF is looking for vendor input on switch IC packaging to evaluate options (bare-die vs. lidded) with regards to IC protection (physical damage) and minimizing the thermal stack.

Optical modules present a different problem to the switch ASIC – the mechanical tolerance across multiple components will not allow for a thin TIMs (less than 100um). The aforementioned assumes a consolidated cooling solution for a given grouping of optical modules. To compensate for this, a gap filler TIM may be needed, and the right product must be chosen to ensure adequate thermal margins under different operating conditions.

Available real estate on the top surface of the CPO substrate is limited. Locating the main digital voltage regulator/s at the bottom side of the CPO substrate is a feasible option. The regulators will require heatsinks to dissipate power. In the Figure 4 below, an example using a heat sink mounted to the underside of the main switch PCB with clearance for tall pedestals to contact voltage regulators (through a gap filler TIM) is shown. Depending on the operating conditions, TIM selection and heat sink design, cooling 200W of heat loss from the voltage rail modules should be possible.

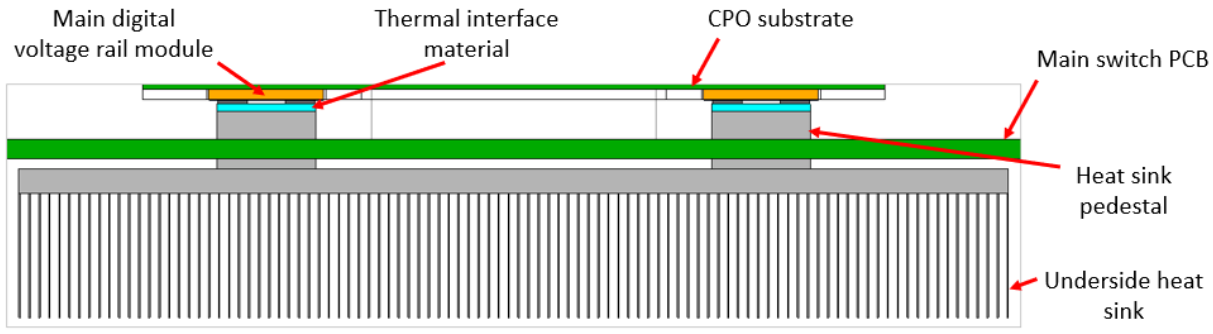

Figure 4: Cross section of a CPO assembly with back-side heatsink for voltage regulators

Each solution described above will depend on the implemented mechanical features (securing space for mounting features) and loading mechanisms to achieve the required thermal performance.

Figure **5** shows the example embodiment with corresponding mechanical and cooling components that form the assembly against the SMB. A bolster plate, with openings for DC connectors and voltage regulators provide mechanical rigidity to the CPO substrate. Cutouts in the switch PCB allow pedestals (part of the underside heat sink) to contact the voltage regulators through a TIM for cooling. Components at the top surface of the CPO substrate are cooled via a cold plate. Depending on component power and operating temperatures, an air-cooled solution (for top surface components) may also be feasible. All components in Figure **5** mechanically interface with each other under load to form a rigid assembly.

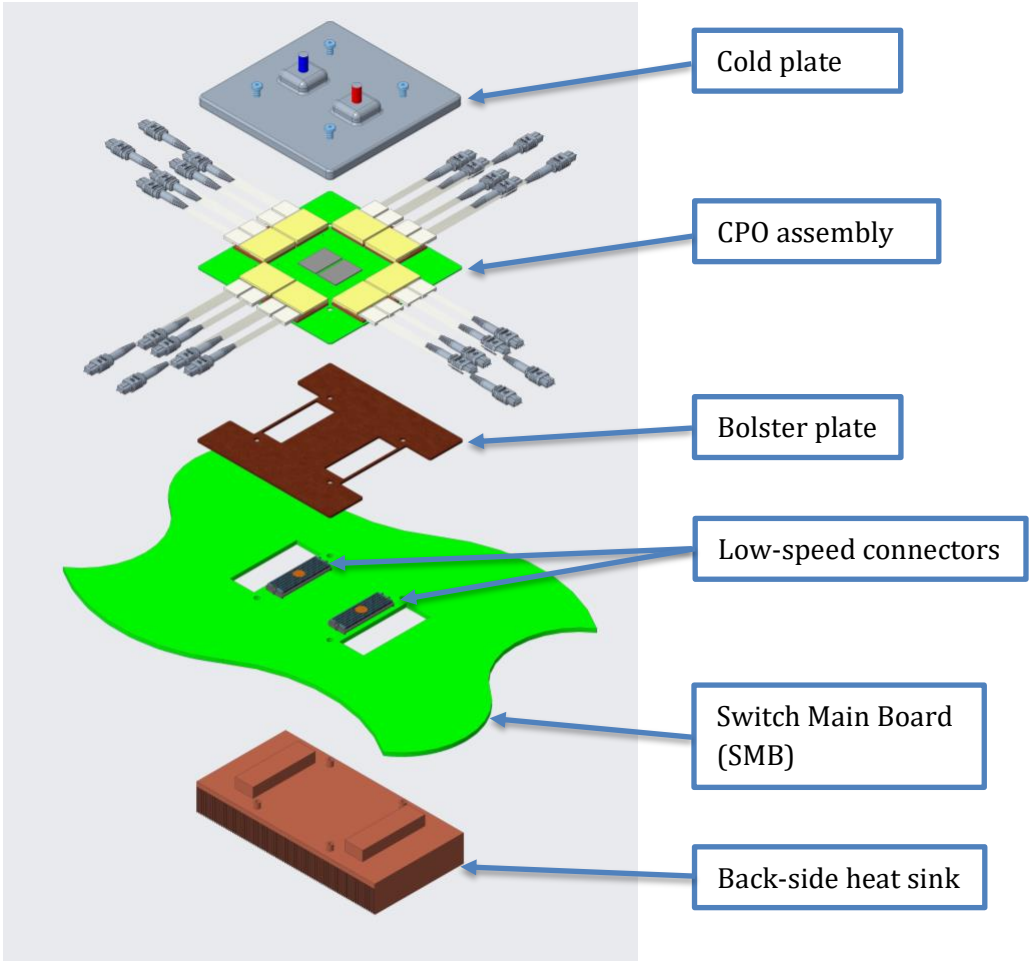

Figure 5: Exploded view of CPO assembly with back-side heatsink

8. Environmental Conditions

The values in Table 4 below are for general guidance and discussion and subject to change.

Parameter	Symbol	Min	Typical	Max	Unit	Note
Altitude, during operation, to sea level		-50		1800	m	
Acoustic, during operation		55			dBA	
Vibration, during operation				TBD		Note1
Altitude, storage, to sea level				5	km	
Altitude, transportation, to sea level				12	km	
Gaseous contamination						Note2
Note: 1. Refer to IEC 68-2-36, IEC 68-2-6 2. Conform to Severity Level G1 per ANSI/ISA 71.04-1985.						

Table 4: Environmental conditions

9. Quality and Reliability

CPO-based network switches must be resilient to failure of a single optical lane. The CPO JDF is seeking input on how to achieve this goal. For example, at the switch level, unused ports could be re-provisioned to take the place of a failing lane, or at the laser level optical modules could be designed with back-up laser sources in case of a laser failure. Please share your thoughts for future revisions of this document

Detailed quality and reliability requirements will be tailored by customer and program. High level guidance for optical modules used in CPO applications is provided below.

9.1. Environmental Stress Testing

Environmental testing for the optical module and ELS will be as described in Telcordia GR-468-CORE section 4.2 for Integrated Modules.

Environmental testing for the switch IC on the CPO substrate will be as described in JESD22.

9.2. Service Life

Maximum service life will vary between five and seven years.

9.3. Failure Rate

CPO assembly FIT rates but must be low enough to ensure the end user's Annual Failure Rate (AFR) targets for the finished network switch.

10. Revision History

The revision history of this document is captured in Table 5 below.

Revision	Date	Description of Changes
1.0	May 1, 2020	Initial release

Table 5: Revision history