Enabling QSFP-DD1600 Ecosystem With Performance-Driven Thermal Innovations

Contributors:

Alvin Chen, Samuel Kocsis - Amphenol
Peter Ajersch, Bonnie Mack - Ciena
Giovanni Giobbio, Joe Jacques, Mark Nowell– Cisco
Terry Little, Woo Park, Paul Yang - Foxconn Interconnect Technology
Attila Aranyosi, Vu Le, Jeffery Maki – Juniper Networks
Hani Daou, Sany Farajalla - MultiLane
Hasan Ali, Scott Sommers – Molex
Remesh Kuzhikkali, Kyle Sammon- TE Connectivity

September 2023

Abstract

The enduring race to accelerate network performance is a tale older than Ethernet itself. With the advent of Artificial Intelligence and Machine Learning use cases, interconnect and system vendors must rise to the challenge of evolving their designs for 1.6 Tb/s ports.

As was the case for previous generations, the universally-adopted QSFP-DD family continues to extend its performance threshold, use cases, and backwards-compatibility. QSFP-DD1600 establishes a roadmap for systems and pluggables to improve cooling capacity thanks to a wide array of thermal innovations. Superior cooling capacity notably enables faster speeds at the port level and can reduce overall system power consumption.

This whitepaper offers comprehensive details of QSFP-DD1600 performance enhancements achieved by both module and system designers. Various analyses show the evolved cooling capabilities of 64-port 2RU system designs, including the capability to cool modules operating at 40W or even higher. This will further extend the utility and presence of the QSFP-DD form factor in network topologies of the near future.

Table of Contents

1 Introduction

The QSFP-DD family of pluggable modules are the extension of the QSFP pluggable modules that currently dominate the broad market for pluggable modules at 40G and 100G. The #1 priority and market driver for QSFP-DD was to enable the support of 400 Gb/s modules without giving up the backwards compatibility with QSFP which the market expected and required.

At 400 Gb/s, the QSFP-DD module met all these goals and was the most widely adopted solution for 400 Gb/s modules. The initial challenges of supporting 400 GbE (the higher power modules, the faster and wider 8x50Gb/s electrical interfaces) were solved.

As Ethernet switch ASIC capacities have increased, switch systems have required more and faster ports to keep up. QSFP-DD800 was developed with enhancements that supported both the thermal and electrical needs. At the same time, system designs were evolving to meet the challenges of these dense higher power modules.

One of the key advantages of the QSFP-DD family is the unchanging flat-top module mechanical form factor. This design provides the backwards compatibility that is critically beneficial to many network operators but also offers the system designers a design point that offers considerable flexibility to optimize their system architectures. The QSFP-DD800 specification provided mechanical enhancements that allowed module designers, connector designers, cage designers and system designers to continue to innovate. Many of these improvements were outlined in QSFP-DD800 Thermal Whitepaper.

Building on this innovation-friendly module form factor, QSFP-DD MSA has released its updated specs for QSFP-DD1600 which supports 1.6 Tb/s of data capacity. Again, enhancements have been incorporated for both electrical and thermal optimizations.

In this whitepaper, these thermal enhancements will be presented and their impacts quantified.

Once again, the inherent flexibility for system designers that the QSFP-DD module offers has provided new insights and improvements that a module designer or system designer can take advantage of.

While the previous conversations around QSFP-DD or QSFP-DD800 performance have centered around maximum power cooling abilities, with QSFP-DD1600 this has not been the sole focus; reducing overall system power due to increased thermal efficiencies has been a key aspect of QSFP-DD1600 development.

Network operators have been raising the challenges with overall power consumption and the ever-increasing need to reduce it. With each generation's speed increase, efficiencies improve from a W/Gb perspective, but overall capacities of new solutions outpace this improvement, resulting in an net increase in power consumption. Many ideas are being proposed to reduce this power increase trend, such as Co-Packaged optics, or Linear pluggable optics. Both approaches are under investigation but have impacts to the current mode of operation that the industry has relied on since pluggable optics were invented, and would therefore require a significant overhaul of network operations.

QSFP-DD1600's enhancements are a path to support the network operators in their constant thermal challenges around increasing power with minimal impact on current infrastructure.

The whitepaper will summarize these QSFP-DD1600 enhancements and how they affect module design, cage design and system design. One key insight the reader should remember is that the inherent solution flexibility enabled by the QSFP-DD design means that many of these ideas and concepts are independent of one another and can be implemented as needed to meet the cost, power, and performance targets for the application. Whether the goal is to design a module and system capable of supporting 50W pluggables, or to design a system with maximum overall power efficiency for a high-density pluggable solution exhibiting typical dissipated power, , QSFP-DD1600 offers a solution with minimal disruption to existing infrastructure.

2 Introduction to QSFP-DD1600 Modules

The QSFP and QSFP-DD module form factor family is the industry's broadest adopted pluggable module providing the highest port bandwidth density.

The QSFP-DD1600 specifications published by the QSFP-DD MSA include the expected design criteria:

- 1.6 Tb/s support (8x 200 Gb/s electrical signals)
- mechanical enhancements to improve both thermal and electrical performance
- backwards compatibility with previous QSFP-DD and QSFP form factors

The integration of these QSFP-DD modules into a system offers the highest density possible for these bandwidths. This level of density translates to optimized designs across both 32-36 port RU fixed switches as well as larger 2RU/4RU fixed boxes and modular line cards.

The QSFP-DD/QSFP-DD800/QSFP-DD1600 module form factors are the industry's smallest 400G/800G/1.6T modules. This is particularly important to system design due to the increasing ASIC power needed to support these higher capacities. Maximizing colder airflow to cool the ASICs is critical, and the smaller faceplate area in a QSFP-DD system allows for a greater volume of air to bypass the optical modules and thus more efficiently be utilized for ASIC cooling.

2.1

2.2 QSFP-DD MSA specification and QSFP-DD1600 enhancements

The Hardware specifications for QSFP-DD, QSFP-DD800 and QSFP-DD1600 are available at www.qsfp-dd.com.

The QSFP-DD1600 specifications cover the following items:

a) Updated electrical interfaces including pad assignments for data, control, status and power supplies and optimized host PCB layout requirements that include an enhanced grounding approach to improve signal integrity.

b) Enhanced optical interfaces specifications with improved breakout cable details to ensure broad compatibility with known industry patch cords.

c) Mechanical specifications enhancements to improve both electrical and thermal performance without compromising backwards compatibility to previous generations. The main enhancements include:

- New module angled latch design to tighten the position tolerances of the seated module in the connector and cage. This allows module pad dimensions to be optimized to reduce parasitics and improve signal integrity.
- Use of bottom side riding side heat sink under the module will be feasible due to the definition of a heat sink contact area with surface flatness specifications. Additionally, a cut-out of the PCB under that area and an opening of the faceplate to allow ingress airflow to pass through the heat sink. See Figure 2-1.
- Recessed channels in the module case bottom and sides to allow the bottom side airflow to egress into the system. These channels offer both enhanced cooling for the module but also increased airflow into the system to aid with ASIC cooling. See Figure 2-2.
- Increased height on module nose integrated heat sink. While the initial version of the QSFP-DD1600 specification only addresses the 1x1 connector cage designs, further updates will include specifications around stacked connector and cage designs. However, it has been determined that additional height can be made available in the nose heat sink without impacting system density. This will increase the upper to lower port spacing in a stacked configuration and will be identified as a Type 2C module type. See Figure 2-3.

d) Further enhancements are possible that exist outside of the QSFP-DD MSA specification but do not impact interoperability and intermateability which is the purpose of the MSA. Some of these techniques will be described more in this whitepaper for the consideration of module or system designers. These include:

- Riding heat sink designs with thermal interface material (TIM) coatings to improve thermal conductivity to the heat sink from the module. Solutions are available to minimize or eliminate wear and abrasion on the TIM. Additionally, work continues on coating materials which could improve tolerance to wear and tear while improving conductivity.
- Enhanced heat sink designs to maximize fin area through all available airflow
- The ability to create openings in the module casing to allow air ingress from the bottom side airflow into the module body to further improve module cooling.

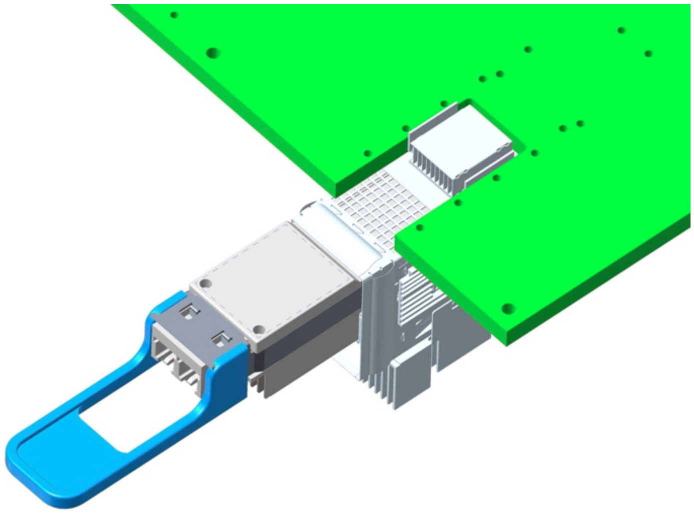


Figure 2-1: Bottom side riding heat sink on QSFP-DD1600

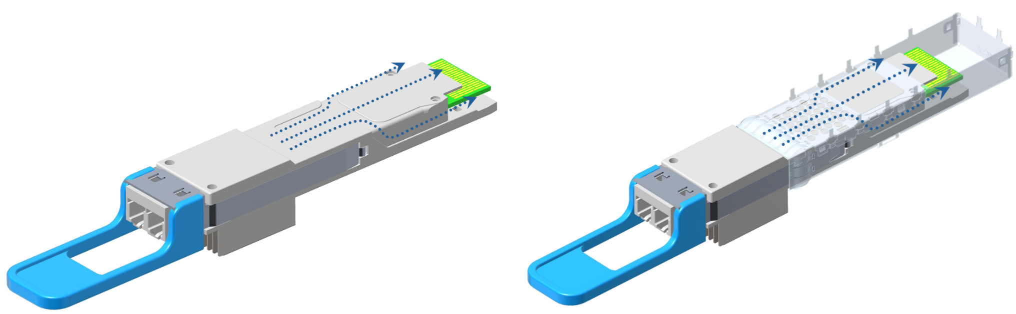


Figure 2-2: Bottom side airflow enhancement features of QSFP-DD1600 showing the module (left) and the module within the cage (right). The arrows indicate where airflow channels are increased due the enhancement features.

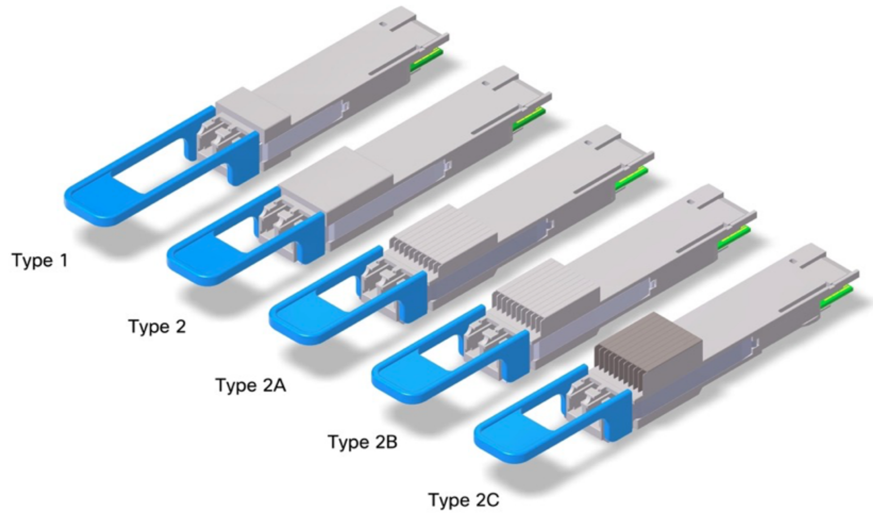


Figure 2-3: QSFP-DD module variants: QSFP-DD Type 1; Type 2; Type 2A; Type 2B and Type 2C

2.3 QSFP-DD module cage configurations

There is no change in the module cage configurations that are possible with QSFP-DD1600. System designs utilize both 1x1 connectors and 2x1 (stacked) connectors. These are often used in systems designs as:

- fixed 32 port 1RU where the 1x1 connectors are implemented belly-to-belly on either side of the system PCB
- 36 port modular line cards, where stacked connectors are used to support the PCB biased lower in the slot.
- Higher capacity systems where both stacked and belly to belly configurations are used.

These higher capacity systems are becoming increasingly important with a fixed 64 port system based being a good example. These system configurations with QSFP-DD1600 will be studied in this whitepaper.

3 Module Design Considerations

With increasing capacities from 400G to 1.6 Tb/s bringing with them a commensurate increase in module power, effective internal module thermal design becomes critical. The internal design of an optical module can have a large impact on the thermal performance of a QSFP-DD module in a system. The goal of a system thermal design is to remove the heat from the module case to ensure that the internal components within the module stay within certain temperature ranges to ensure optimal performance and reliability.

This was discussed in detail in the QSFP-DD800 Thermal Whitepaper and many of the findings are equally true for QSFP-DD1600 but perhaps more important to be implemented as average module power increases.

The study in the QSFP-DD800 paper explored the improvements that were possible with a consistent 25W module design by adjusting such aspects as: improved case material conductivity; improved surface flatness where the riding heat sink contacts; improving thermal conduction within modules with case design/thickness and improved thermal interface material (TIM); and increased nose heat sink heights. In that study, the effects of changing design aspects on the critical component temperatures and case temperatures were studied. It was shown that temperature reductions of 15-16°C were possible with relatively simple optimizations.

With QSFP-DD1600, where typical intensity modulated direct detection (IMDD) module powers may be around 25W, module designers are highly encouraged to incorporate these improvements. Technology advancements and miniaturization of advanced conductive elements are able to be integrated into many module case designs if desired. The increased Type 2C nose heat sink also offers considerable improved performance.

As mentioned earlier, the incorporation of recessed channels on the module bottom and PCB cutouts increases airflow around and even into the module body if case openings exist. These aid in overall system airflow design but also are very effective at dissipating module heat.

3.1 Enhanced module temperature monitoring

One important factor with QSFP-DD1600 is that the traditional approach of testing, qualifying and monitoring module temperatures leaves a lot of potential margin on the table. These modules are complex designs and a simple case temperature spec is inadequate for today's system design use.

The traditional approach to monitoring module temperature is to define a monitor point on the module, likely under the heat sink contact area and design the system cooling to not exceed the max Tcase specification (typically 75°C). This monitor point is typically inaccessible to verify without disturbing the heat sink and also lacks fidelity with the actual internal component temperatures – which are actual temperature limits which define whether the module will operate to specification. However, internal sensors report the Tcase using the Digital Optical Monitoring (DOM), which can be read by the software management interface (i.e., CMIS).

An example of the margin that is given up using this too-simple approach is shown in Table 3-1.

Table 3-1 shows the readings from a typical module located in the lower port of a stacked cage. The temperature limit specifications are shown for the module case and also for the actual critical components in the module which actually are the true temperature limits to ensure module operation and performance.

Module in Lower port	Limits	Actual	Margin
Tcase (above DSP)	75°C	72.6°C	2.4°C
Laser	85°C	76.4°C	8.6°C
TIA/Driver	105°C	81.4°C	23.6°C
DSP	105°C	93.5°C	11.5°C

Table 3-1: Comparing the differences in temperature margin between the module case and the critical internal components

Using the module case temperature as the target limit yields only a 2.4°C margin, but, as has been demonstrated, modules have a lot more margin before they reach maximum specification. In the above example, there is a further 6.2°C available before any performance impact from the laser would occur.

Therefore, it is proposed that module DOM reading for QSFP-DD1600 be modified in such a way to reduce this limitation, while maintaining compatibility with the currently adopted CMIS software.

To maintain backwards compatibility with the existing CMIS and system software, the value reported in the DOM register becomes:

$DOM = 75C - MIN(\Delta t_{laser}, \Delta t_{DSP}, \Delta t_{TIA}, etc.)$

where:
Δt_{laser} = is the temp margin between actual and max for the laser
Δt_{DSP} = is the temp margin between actual and max for the DSP
etc.

This simple change in reported DOM values in the module immediately offers system designers and network operators a much better fidelity metric to monitor and manage their systems. This change also reduces any "wasted" margin in the thermal architecture and provides much better module control and operation for system management.

3.2 Module power trends

It is informative to look at the trends of the pluggable modules as they progress from 400G to 800G and 1600G. One observation of particular note is that while for every generation the W/Gb decreases, overall system power increases. Interestingly, IMDD and Coherent technologies are not scaling at the same rate, due to the different combination in the number of lasers, amount of digital logic power in DSPs, and the amount of analog SERDES power. With each advanced CMOS node, digital power sees more power efficiency benefits than analog power.

Looking at the historical trends and projections for IMDD modules, typical power is increasing at a higher rate than coherent module power as shown in Fig 3-1. A key takeaway from this with regards to

QSFP-DD1600 is that it becomes increasingly important for pluggable thermal solutions to be both capable of handling the higher power coherent modules but also be highly efficient at cooling the IMDD modules whose power is no longer so differentiated from coherent modules.

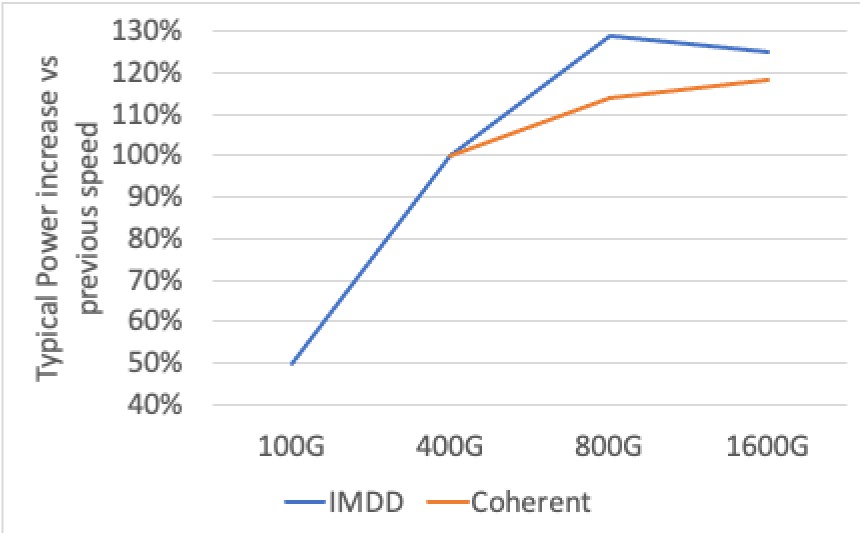


Fig 3-1: Optical module power increase trends

Finally on the topic of module power trends, a useful observation that will be discussed later in the whitepaper is that module power varies as a function of module temperature. While operating at its highest temperature, it is also operating at its highest power due to the increased leakage current in chips, increased laser bias and increase in any cooling components that might exist. Being able to operate the module at lower temperatures also reduces the power alongside improving component reliability.

In summary, 1.6 Tb/s-capable pluggable module design can optimize its thermal performance within a system with careful internal design choices. Numerous design options and trade-offs exist to support the wide breadth of module architectures and design variants that will be deployed. As we move towards QSFP-DD1600 designs this emphasis on power reduction becomes even more important.

4 System Design Considerations

An effective system thermal design encompasses many considerations from overall line card architecture, ASIC cooling, fan operation, design of heat sinks and airflow design and management. We will explore some of the options for cage design and line card design as well as some of the enhancements that are independent of line card design. Front-to-back airflow direction is the dominant system utilized by the industry, and on which this whitepaper will concentrate.However, the inherent design flexibility that QSFP-DD offers will be used to show how a side-to-side airflow based system can utilize QSFP-DD too.

4.1 General design considerations

The basic principles of system design considerations were covered well in the [QSFP-DD800 Thermal Whitepaper] and are still valid and true. The details will not be replicated in this paper but instead additional insights or capabilities will be discussed that further enhance the ability of the systems to cool QSFP-DD1600 modules.

As overall power consumption of the network switches has become a more critical consideration for data center operators, the focus of the thermal analysis around QSFP-DD1600 has also shifted.

Being able to cool higher power modules is an important capability of any solution, as it provides the user with assurance that they will be able to support any high-power variant of modules that they want to use. This is important for QSFP-DD1600 where Coherent modules (1600ZR) or breakout modules (2x800G-FR4) are required.

However, the more important aspect of this thermal capability is that these same system design optimizations mean that the system is easily capable of handling the more typical powered modules efficiently. This results in some useful advantages:

- lower overall system power
- lower module operating temperatures

4.1.1 Faceplate features

Faceplate features play an important role on the thermal performance of optics and other downstream components on the card. QSFP-DD is the smallest module used in the industry and this provides increased faceplate area for airflow to the rest of the system. This is a key advantage of any QSFP-DD based system design. With QSFP-DD1600 there are some specific enhancements:

- With the PCB cutout region below the cage (Figure 2-1), the faceplate should have openings/perforations aligned with the cutout to allow air ingress through the bottom heat sink (if present) or through the recessed bypass channels in the module body to increase bypass airflow to the rest of the system
- With the Type 2C modules with the increased fin height, faceplate should have openings/perforation aligned with the nose heat sink

Common to any system design is the concern about balancing airflow between upper port modules, lower port modules, and the rest of the system. Since all heat sinks and faceplate openings are completely under the control of the system designer, this allows this airflow balancing to be optimized without concerns for port to port variations as different vendor modules are used.

4.1.2 Riding heat sinks

Riding heat sinks on the pluggable module cage are the main heat rejection path to the forced airflow around them.

The flat module surfaces of QSFP-DD allow significant flexibility in the design and size of these heat sinks. Since overall fin area is the dominant performance criteria in an air-cooled system, this flexibility of design is foundational to QSFP-DD1600's thermal performance.

Suitable contact between the module surface and the heat sink surface is important and surface flatness specs and retention clips have been proven out to be sufficient for purpose.

In Section 5, a number of different system configurations will be presented including 1RU 32 port belly to belly cages and 2RU 64 port belly to belly stacked cages. heat sink configurations for all of these are possible including bottom riding heat sinks between belly to belly ports if desired.

Some examples of different riding heat sink configurations are shown in Figures 4-1, 4-2, and 4-3.

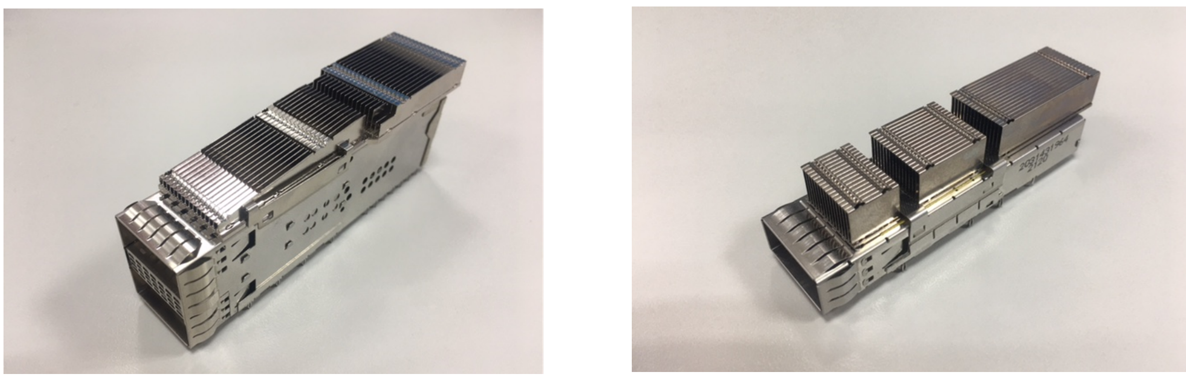


Fig 4-1: Examples of a 2x1 stacked cage (left) and a 1x1 cage (right) with customized riding heat sinks optimized for the line card design

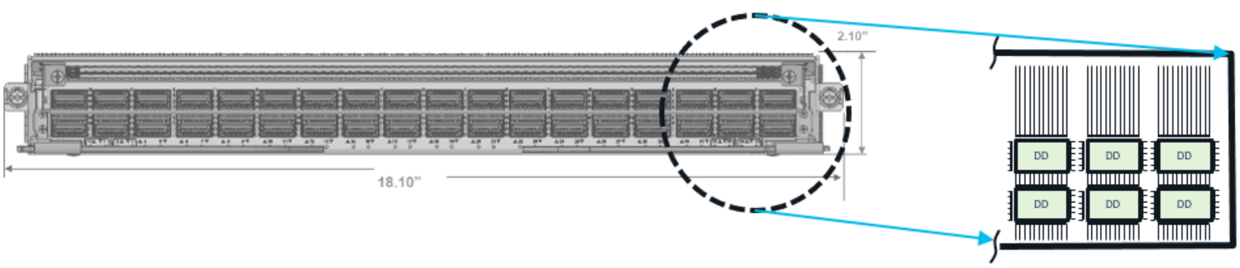


Fig 4-2: Example of 36 port modular line card using optimized heat sink configurations optimized to the products use case. In this case higher power modules were to be used in the upper ports

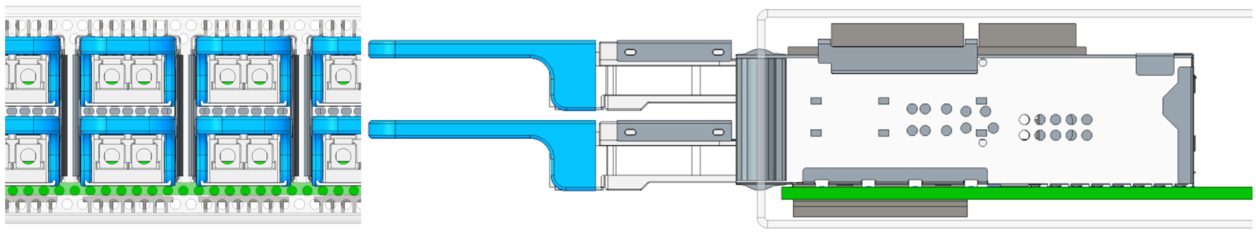


Figure 4-3: Example of a bottom heat sink attached to a stacked cage

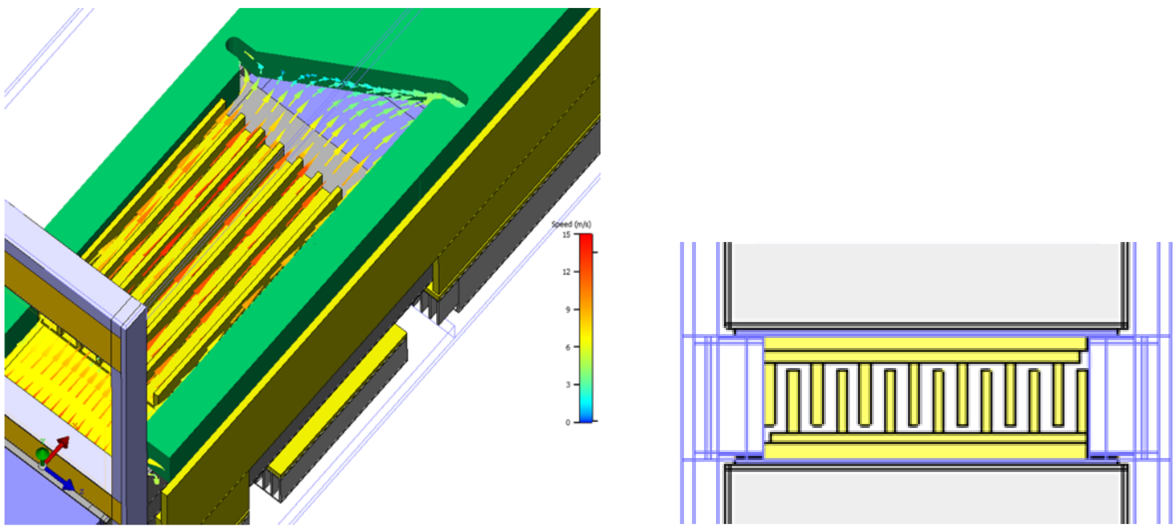


Fig 4-4: Examples of bottom side riding heat sink implementations. On the left showing the underneath view of a PCB with the cutout allowing airflow through a bottom riding heat sink. In a belly to belly configuration, the fins can be designed to interleave with each other as shown on the right.

4.1.3 Improving module to heat sink interface

With QSFP-DD modules, the riding heat sinks are the dominant heat rejection path for cooling pluggable modules. The thermal performance of the module to heat sink interface is a critical design aspect of any solution. Well known techniques have been utilized and specified in the QSFP-DD specifications, as needed, to ensure thermal performance. These include surface flatness and roughness specifications or clip retention design to maximize conductivity across the interface. Further improving the performance of this interface would have benefits.

One approach is to use a thermal interface material (TIM) on the pedestal of the riding heat sink. This would increase the thermal conductivity by reducing the impacts of any mechanical inconsistencies on that riding interface. One challenge to this approach is that, typically, TIM are not resilient to the wear and tear associated with module insertion and removal. Modules have some sharp edges on them and may be inserted at slight angles as they are pushed in. Figure 4-5 highlights some of these issues and shows the wear and tear to the TIM that may occur.

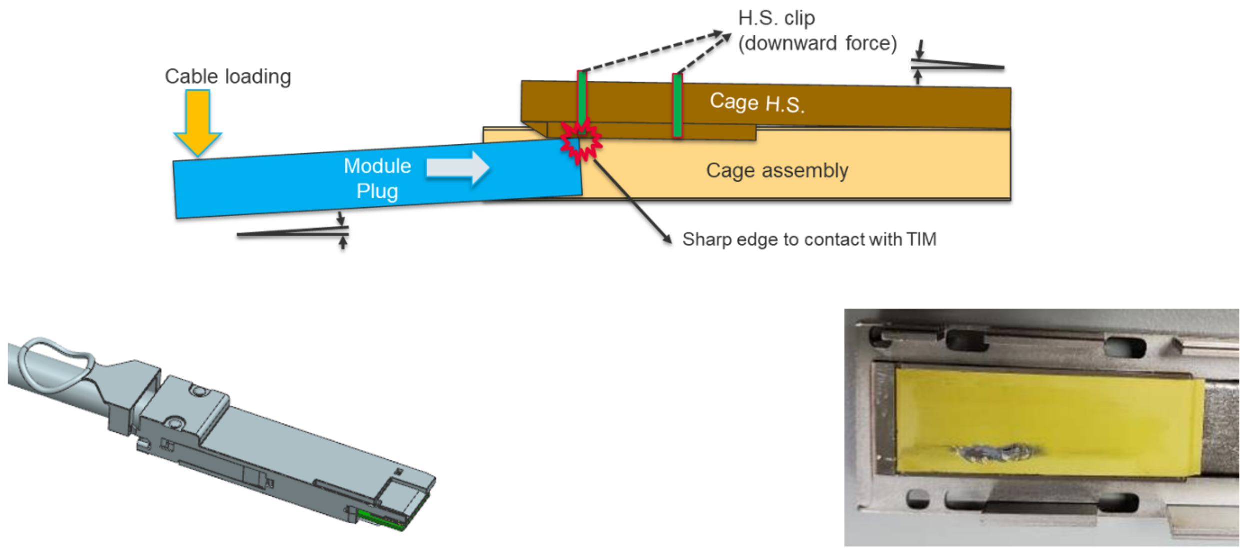


Figure 4-5: Insertion and removal of a pluggable module can cause damage to a thermal interface material present on the riding interface of a heat sink

However, QSFP-DD1600's innovation friendly aspects enable a solution to this opportunity to improve the thermal conductivity of this heat sink to module sliding interface. Cage-based solutions are already available that enable a TIM usage on this interface while ensuring the design complies with durability and insertion cycle requirements of a pluggable module.

An example of this is the Drop Down heat sink (DDHS) concept, which has been validated to preserve the TIM and resulting performance improvement up to 200 insertion cycles. The DDHS concept is based on a mechanism built into the cage that protects the TIM from any physical contact (and resulting damage) while the module is being inserted or removed and the heat sink (with TIM) lowers on to the flat module surface only when the module is almost fully inserted. This prevents the TIM from being damaged by any sharp edges or angled module impact during insertion.

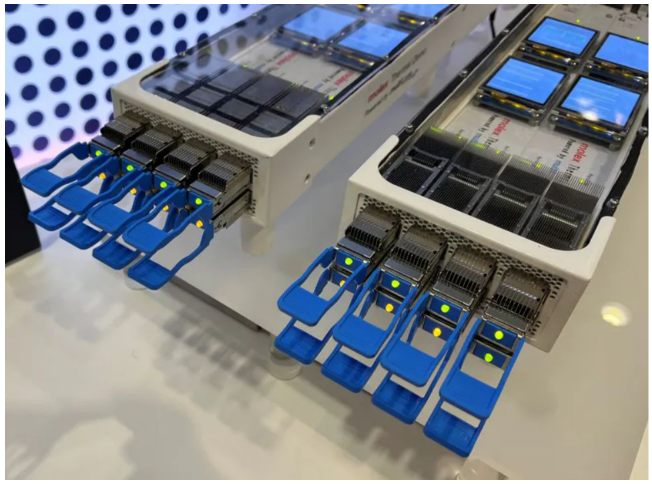


Figure 4-6: Live demonstration of Drop Down heat sink (on right) compared to traditional heat sink (on left)

The DDHS is shown in Figure 4-6 besides some traditional cages and riding heat sinks. It should be noted that there is no impact to the system's module density.

These solutions have been built and tested and improved performance is verified. To try to quantify the impact of these cages, simulations were carried out using a thermal module consistent with a MultiLane thermal module operating at 40W. The details of this model are explained later in the paper. Figure 4-7 shows the temperature measured at the Test Sensor (TS) #4 (consistent with where a DSP would be located in a real module) for various levels of improved thermal conductivity that usage of a TIM would offer.

Added to the curve are the data points of TIM from two suppliers that are currently being tested. Just looking at the result for "TIM supplier 1" that has completed lab testing and validation the results show that a reduction of temperature from 104°C (dry contact) to 98.5°C occurs and is consistent with the experimental validation. As shown in the graph, further improved TIM yielded improved results.

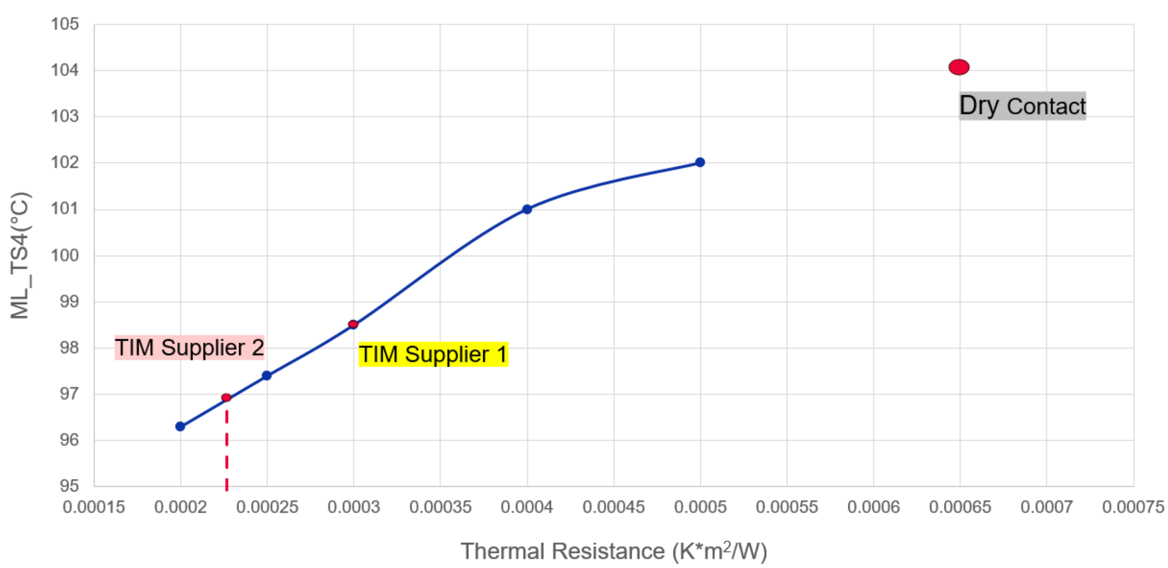


Figure 4-7: Reduction in internal temp sensor in module for different TIM that could be used in a Drop Down heat sink implementation. The dry contact data point is shown for comparison.

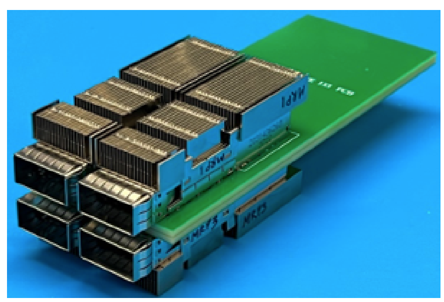
An experimental study was conducted to measure the relative impact of these various techniques to improve the conductivity of the module to heat sink interface. A wind-tunnel apparatus was built incorporating four QSFP-DD1600 ports configured in a belly-to-belly format. The initial study compared three design variants:

1. Riding heat sink (RHS) with typical dry contact (2 mil flatness)
2. RHS with a TIM-A (TIM with durable coating)
3. DDHS with TIM-B

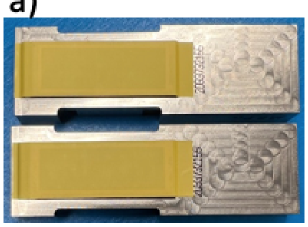
Some details of the experimental setup are shown in Fig 4-8. The experiment was conducted using a MultiLane 30W module. The measured module temperatures (i.e. at TS#4, the hottest location on the inner surface of the module's top housing) are shown in Figure 4-9. It is noted that the RHS with TIM-A case is included to verify the impact of improved thermal conductivity, but as previously mentioned, addressing the wear and tear due to insertion cycles remains a consideration.

At the applied module power of 30W, TIM-A provides ~3°C lower temperatures than the baseline option, RHS with dry metal-to-metal contact. The DDHS, with a better performing TIM-B, reduces the module's case temperature by a further 4°C, or 7°C in total when compared to the RHS with dry contact case which is in line with the simulation results shown in Figure 4-7.

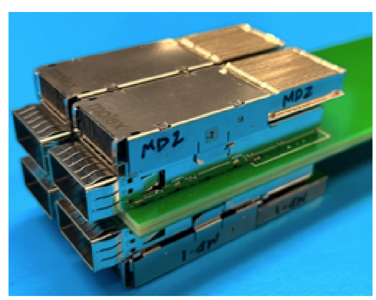
a)



c)



b)



d)

Figure 4-8: Details of experimental setup. a) wind tunnel setup with four QSFP-DD modules. b) RHS showing TIM-A on pedestal. c) QSFP-DD1600 cage with RHS. d) QSFP-DD1600 DDHS cage

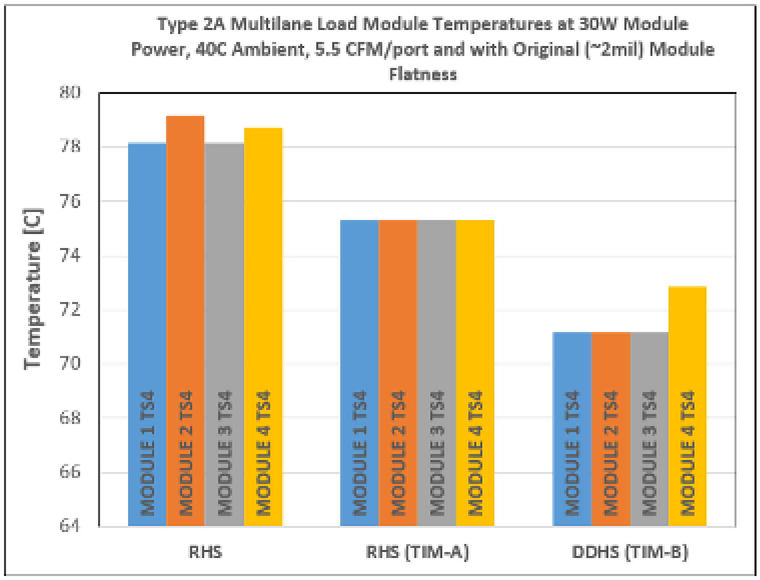


Figure 4-9: Experimental results comparing the differing RHS interface options.

To further understand the importance of the modular surface flatness, the experimental cases were run again but this time including a module with improved surface flatness of 1 mil. Module flatness was measured with a 3D optical profilometer. The results are shown in Figure 4-10. As the results show, with improved module flatness, the performance of the RHS with metal-to-metal contact improves significantly (by ~7°C). For the other configurations the improved surface flatness had a consistent improvement as expected with even a 2-3°C improvement with the DDHS. This highlights the importance of the module surface flatness to improve thermal performance. Considering the dry-contact case, the improved surface flatness case out-performed the result of the 2mil flatness case when the RHS with TIM-A was used. It should be noted that at higher module powers the improvements would be higher.

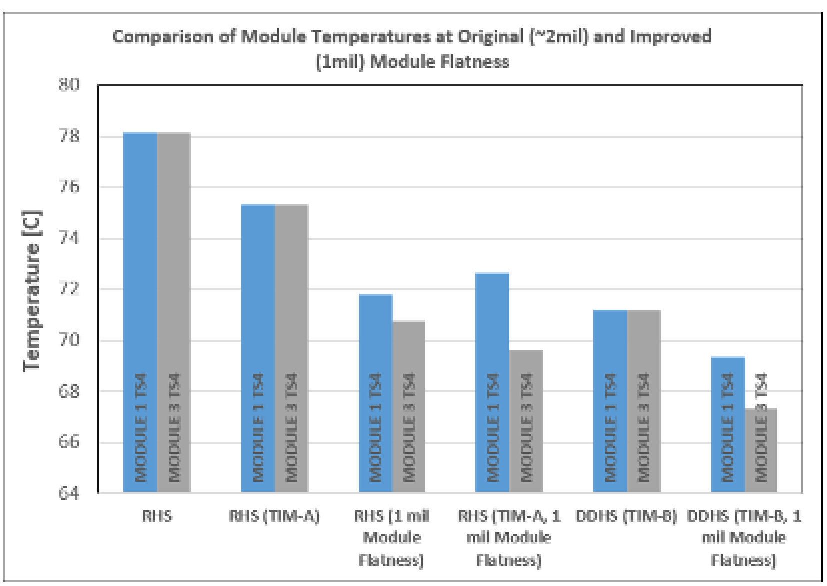


Figure 4-10: Experimental results comparing the differing RHS interface options including the effects of an improved (to 1 mil) module surface flatness.

Beyond the DDHS concept which can utilize a variety of TIM, further work is ongoing on more hardened coatings for the heat sink pedestal that would be better tolerant of the insertion and removal cycles. These coatings would improve the conductivity of the module to heat sink interface without requiring techniques to minimize wear and tear.

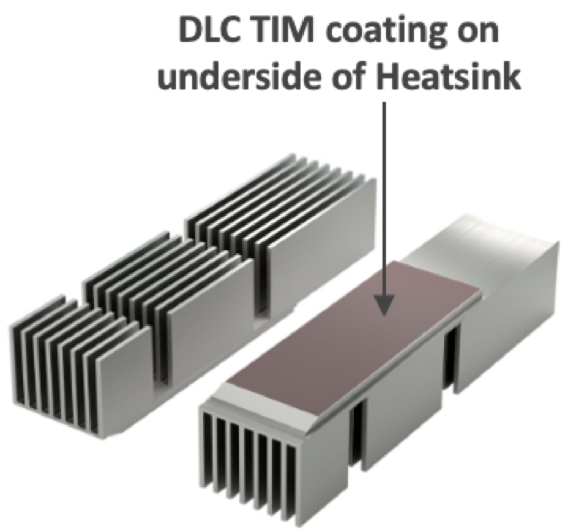


Figure 4-11: Example of a hardened heat sink coating to improve the thermal conductivity of the riding heat sink interface

This technology is still being developed by several different companies. For the application, the coating would be applied to the underside of the heat sink. This method of contact resistance improvement is intriguing as it has the goal of reducing the complexity of the solution while improving the conductivity of that interface. Figure 4-11 shows where a coating, such as the Diamond-like coating (DLC) technology would be applied to the heat sink pedestal.

Given the thermal improvements that result from improving this riding heat sink to module interface, a number of options are available to improve performance. A simple improvement to the module's case flatness spec shows marked improvement. At the system side, a pragmatic solution to incorporate a TIM within a DDHS-based cage is possible. And as more data becomes available, investigations into validating hardened coatings will be an interesting opportunity for system designers to consider.

As will be consistently pointed out in this whitepaper, all techniques are additive to any of the other techniques shown in this paper and within the control of the system or module designer to implement.

4.1.4 Side-to-side airflow support

While the majority of data center networking products use a front-to-back airflow architecture, line card architectures utilizing the previously common side-to-side airflow architecture need a solution. The flexibility of the riding heat sink design of QSFP-DD makes this a fairly simple solution by rotating the fin direction on the riding heat sink to match the airflow, as shown in Figure 4-12.

More details of these system and design considerations will be explored in Section 6 & 8.

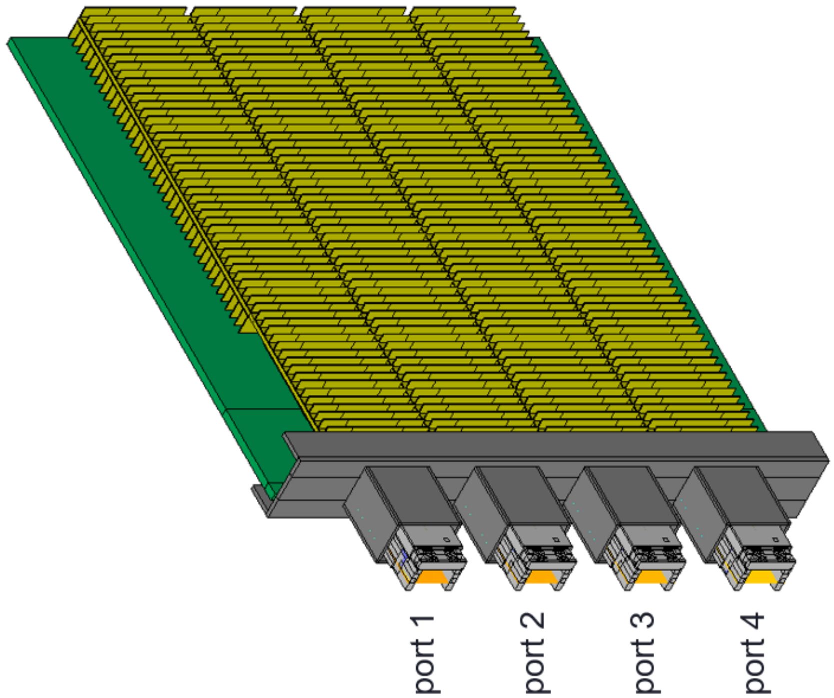


Fig 4-12: Riding heat sink design can be designed to support a side-to-side airflow architecture; sample configuration shown

Numerous systems are designed utilizing a side-to-side airflow approach for various design, environmental or legacy reasons. Being able to design a solution for these higher port speed interfaces is necessary. The intrinsic flexibility of system design around QSFP-DD ports allows system designers to build and cool these high-speed modules.

5 Module Thermal Analysis

In order to assess the relative merits of different system design options a series of analyses are executed using different configurations of cages and modules by the authors.

5.1 Modules under test

Unless otherwise indicated, the analyses are based on a QSFP-DD thermal test module provided by MultiLane, with variable power dissipation and a representative thermal simulation model. Modules have multiple thermal loads throughout the internal cavity of the module to allow for simulations of different power spot placement and overall power dissipation of the internal components.

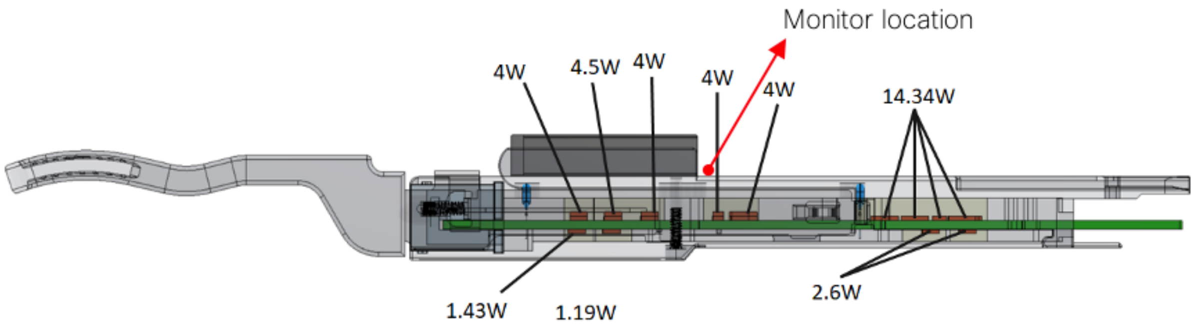


Figure 5-1. Details of the thermal module used

Figure 5-1 shows some details of the module with a representative power profile of a 40W module, with increased thermal dissipation capabilities. The load sources have been scaled proportionally from a 25W module and thermal simulation enhanced to be consistent with higher power modules, including higher performance TIM. Common boundary conditions and fan curves were used consistent with a typical system design. The common monitor test point used in experimental verification is shown.

5.2 Increased cage port separation

As indicated in Section 2, one of the cage changes recommended for QSFP-DD1600 is to further increase the separation between the upper and lower ports within a stacked cage. The V7.0 release of the QSFP-DD MSA spec which includes the QSFP-DD1600 specifications only specifies the 1x1 cage and connector and the full stacked cage definition will follow in later versions of the spec. However, based on system mechanical analysis, it was determined that the spacing between the upper port and lower port in a stacked cage could increase by ~ 1.8mm without impacting faceplate density.

The impact of this further separation is two-fold. It allows a taller internal riding heat sink on the lower port module which increases the fin area but the larger opening also allows the airflow to straighten in that middle area effectively increasing the air velocity. Both of these effects lead to an improvement in the cooling of the lower port module which has always been the limiting position in the stacked cage solutions. The increased height does effectively reduce the heat sink fin height possible on the upper port, but the overall outcome is that it brings the upper and lower ports into more balance from a thermal perspective.

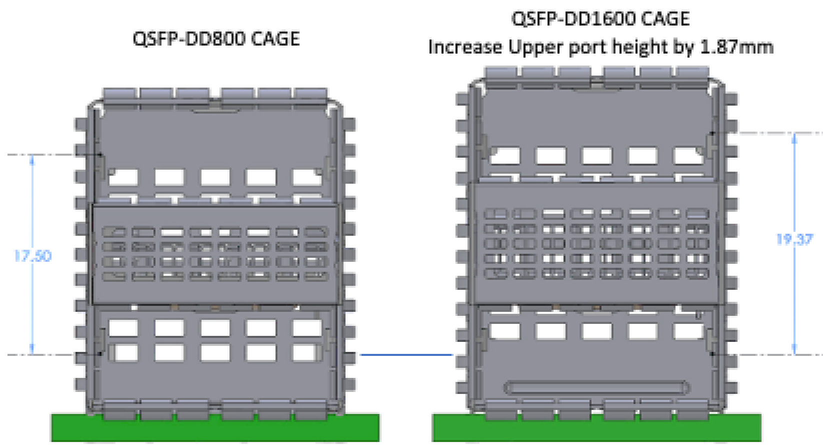


Figure 5.2: The proposed increased port separation height between the upper and lower ports in a QSFP-DD1600 stacked cage

The simulation analysis utilized a 40W module and assumed a 30°C ambient condition. A consistent 3-6°C reduction in temperature readings was found to be the outcome of the increased port separation.

5.3 Bottom side cage opening

The QSFP-DD1600 cage and module specifications include vents in the bottom of the cage to allow bypass air under and around the module as well as the possibility to include a bottom riding heat sink. These enhancements were shown in Figures 2.1 and 2.2.

To study the effects of these modifications, a belly to belly mounted 2x1 stacked cage configuration was used. This configuration would be consistent with a 64 port 2RU system design. Focusing the simulations to just the single slice allows quicker analysis than a full system design. As above, realistic 2RU boundary conditions were utilized in this analysis. Figure 5.3 shows the configuration under analysis.

The cage has the opening of 14x29mm on the bottom side of the cage on bottom port for improved bottom port cooling, heat sinks are presented on both top and bottom port and AL material was used in the simulation. A few height options of the cage are included in the simulation for comparison. Full details of this study are included in Section 8.

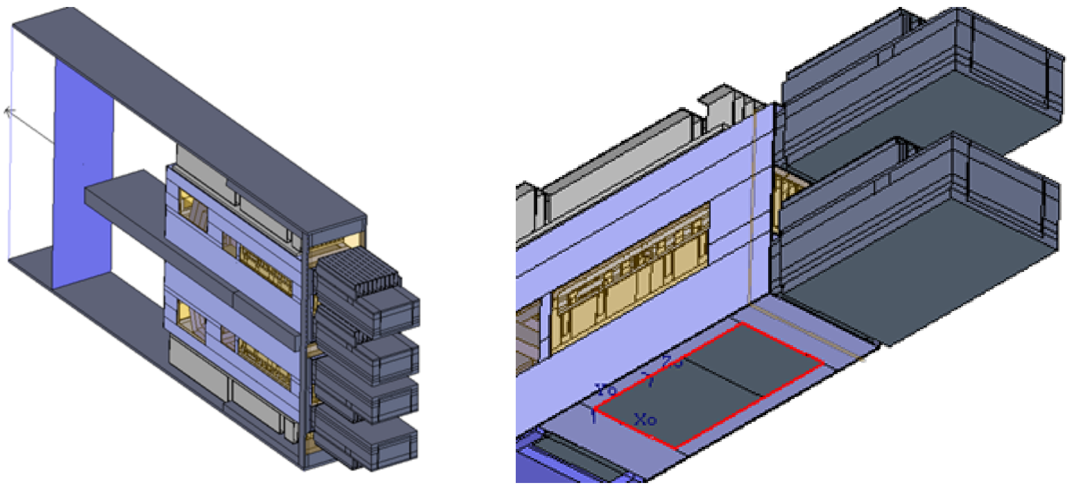


Figure 5.3: (left) Belly-to-belly stacked QSFP-DD1600 cages configured consistent with a 64 port 2RU system design. (right) Bottom side of cage showing opening that allows increased airflow bypass and enables possible inclusion of bottom riding heat sink.

Thermal performance comparison_45˚C Ambient, 2U B2B (result from bottom side)							
Configuration	2x1 SMT cage						
Connector	2x1 SMT connector						
TIM option	N/A (dry contact only: contact r: 1.0C.in^2/W)						
Module power-30W	Y	Y	N/A	N/A	N/A	N/A	N/A
Module power-40W	N/A	N/A	Y	Y	Y	Y	Y
Module IHS type 2A	N/A	N/A	N/A	N/A	Y	N/A	N/A
Module IHS type 2B	Y	Y	Y	Y	N/A	N/A	Y
Module IHS type 2B+2.0mm	N/A	N/A	N/A	N/A	N/A	Y	N/A
Cage height per 800G	Y	N/A	Y	N/A	N/A	N/A	N/A
Cage height + 2.0mm	N/A	Y	N/A	Y	Y	Y	Y
Cage bottom open + faceplate vent	Y	Y	Y	Y	Y	Y	N/A
Air flow one side (CFM)	9.00	9.33	9.00	9.33	9.36	9.24	10.00
Static Pressure (In.H2O)	1.66	1.60	1.66	1.60	1.59	1.61	1.69
Top port T of ML module #4 (˚C)	51.2	52.4	60.4	63.4	66.9	60.1	62.4
Bottom port T of ML module #4 (˚C)	54.2	52.5	66.5	64.2	67.2	61.2	69.0

Table 5.1: Summary of results from belly-to-belly stacked configuration analysis

A number of variations of module power and cooling configurations were analyzed and are outlined in Table 5.1. The main variables under test were module power, module nose integrated heat sink (IHS) height, cage bottom opening (including faceplate opening and PCB cutout) and stacked cage port separation. Based on the simulation result, here are some of the key observations:

- As power increases from 30W to 40W, the delta T on the hottest spot of the top side of module increases between 9.2-12.3°C, bottom port increases 3.1°C more than the top port on baseline

QSFP-DD800 cage height and only 0.7°C with a increased 2.0mm higher port separation configuration.

- To improve the bottom port cooling, increasing the distance by 2.0mm between top and bottom port compared to the QSFP-DD800 cage height allows the lower port heat sink height to grow by the same amount and this will reduce bottom port temperature by 2.3°C at 40W power.
- The taller module nose IHS on the module shows further improvement. With a 40W module, increasing module IHS from type 2A to 2B, the delta T improves 3.0-3.5°C, from 2B to an additional 2mm, the delta T improves a further 3.0-3.3°C.
- The effect of the bottom side cage opening was included in most of these configurations in this analysis. When it was removed it would make the 40W bottom port worse by 4.8°C .

This analysis yielded three important observations. First, that a 64 port, 2RU system design with 40W modules was feasible and showed good operating conditions. Second, a bottom cage opening yielded a marked improvement in the lower port module cooling which is typically the limiting port in a thermal design. It should be noted that no bottom riding heat sink was included in this analysis. And finally, the third observation was that an increased nose IHS height, above that of QSFP-DD800 also had a marked improvement.

5.4 Bottom side cooling enhancements

As shown in Section 5.3, the opening of the cage on the bottom to allow bypass air into the cage has marked improvement on the thermal performance of the QSFP-DD1600. While not analyzed in the module level simulations, this bypass air also is important in improving the overall system cooling design as it aids the cooling of the large ASIC behind the modules. However, some extra advantage can be taken of this bottom side airflow path and this is analyzed in this section.

In this section, four different configurations are analyzed to understand the possibilities and impact of including bottom side cooling in a system and cage design.

- Baseline: no bottom cage vent opening
- Bottom cage vent opening
- Bottom vent opening + bottom heat sink
- Bottom vent opening + bottom heat sink + module case venting

The setup for this comparison was a simple 2x1 stacked cage configuration consistent with a 32 port 1RU design using 40W modules with an ambient temp of 45°C with an elevation of 900m. As shown in Section 5.3, this can be extended to a 64 port 2RU design.

Examples of the bottom side designs to utilize the bottom cage opening have been shown in Figures 2-2, 2-3 and 4-4. The last option listed above of having module case openings is shown in Figure 5-4 where it is possible to consider small openings in the module case bottom and sides to allow that bypass airflow to ingress and egress the internal module area. This option would be available to a QSFP-DD1600 module designer to implement if they had high power designs. It would not interfere or impact and MSA interoperability specs and preliminary analysis around EMI, and reliability – such as tested with mixed-flow gas tests – suggest that there are no concerns.

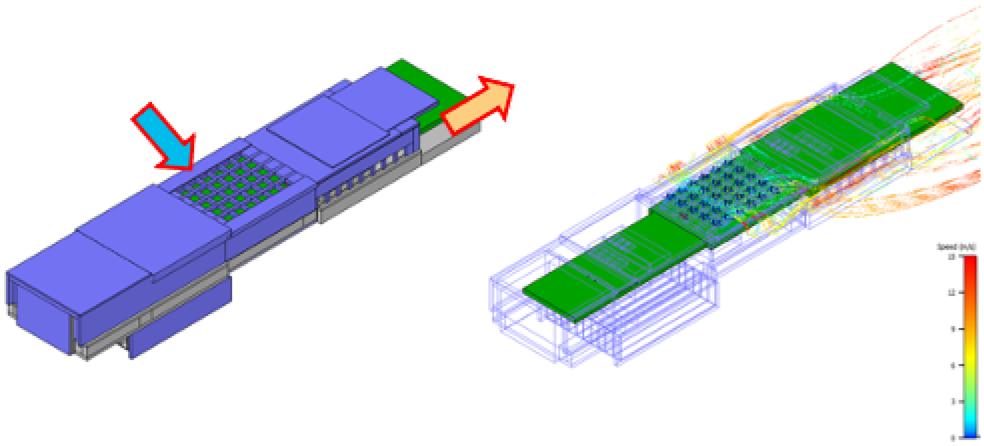


Figure 5-4: Example of module case venting to allow bottom side bypass air to enhance module cooling

The results are shown in Table 5-2. The considerable improvements due to enabling some enhanced bottom side bypass airflow are clear. Both the system designer and module designer have options available to them to take advantage of these QSFP-DD1600 enhancements.

When just considering the cage vent only, these results are consistent with the previous analysis in Section 5.3 contributed by a different co-author. However, it is shown that having that bottom side bypass airflow pass through a bottom heat sink or through the module case itself, additionally increased improvements. As shown in Table 5-2, an ~ 6°C improvement in a lower port module temperature is possible.

40W module 2x1 Stacked cage	ΔT improvement °C	
	Upper port	Lower port
Baseline	-	-
Cage bottom vent only	4.7	3.5
Cage vent + bottom heat sink	4.6	4.9
Cage vent + module case vent	5.6	5.5
Cage vent + bottom heat sink + module case vent	5.8	6.6

Table 5-2: Results from analysis of impact of options to utilize bottom side airflow in QSFP-DD1600

Interesting in Table 5-2, the upper port benefitted from this enhancement too. This was due to two reasons, firstly the lower port module was cooler and therefore less adjacent heating was occurring and secondly, the cage under study included an opening in the bottom of the upper port's cage surface to allow the benefits of the bypass air to be available to the upper module too.

6 System Thermal Analysis

A series on full system power analysis was performed in order to study the potential for system power efficiencies that arise when using QSFP-DD1600. These simulations are much more complex than module-centric simulations, accounting for all aspects of system design. This includes fan design and performance, ASIC thermal dissipation and airflow, module power temperature variability, and other measures to optimize overall system power.

As mentioned in Section 3.2, module power varies with module temperature and this characteristic was included in the optimizations. More details are available in Section 8.

To study the improvements possible with QSFP-DD1600 modules, two system comparisons were done.

6.1 System improvements with QSFP-DD1600 in a 32 port 1RU

Firstly, a 32p 1RU design fully populated with either QSFP-DD800 based modules operating a nominal 30W worst case or QSFP-DD1600 modules also operating at a nominal worst case 30W was studied. The goal of the study was to quantify the impact of the improvements possible with QSFP-DD1600 (vs QSFP-DD800) with everything else being kept constant.

The QSFP-DD800 and the QSFP-DD1600 modules used in the analysis are shown in Fig 6-1.

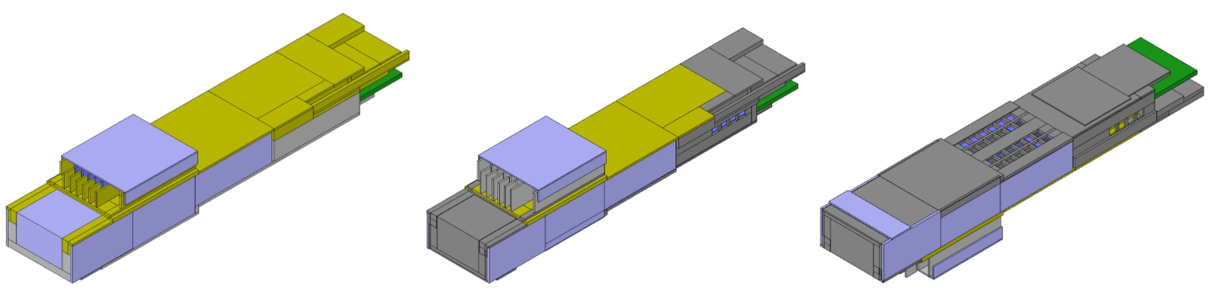


Figure 6-1: QSFP-DD800 module (left); QSFP-DD1600 module (middle); QSFP-DD1600 module underside (right)

The QSFP-DD1600 and QSFP-DD800 modules were consistent with MSA specifications with more details available in Section 8.

The full 1RU 32 port system simulation is shown in Figure 6-2. The modules were representative of a 30W coherent module. Also shown in Figure 6-2 are the surface temperature plots on the QSFP-DD800 and QSFP-DD1600 modules showing that the QSFP-DD1600 module is noticeably cooler. The riding heat sinks used the worst case dry metal contact of 6.45 °C-cm^2/W. Fan throttle was at 55%.

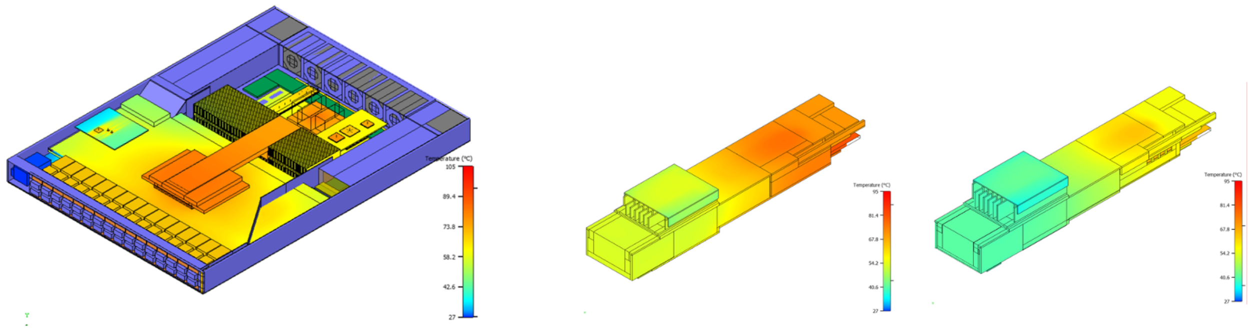


Fig. 6-4: Full 1RU 32 port system simulation (left) with either QSFP-DD800 modules (middle) or QSFP-DD1600 modules (right)

Module	Air Flow	NPU Margin	DSP margin	Module Power	System Power
QSFP-DD800	102.0 CFM	+7.1 °C	+22.7 °C	27.7W	2095W
QSFP-DD1600	103.7 CFM	+9.4 °C	+36.2 °C	25.5W	1995W

Table 6-1: Results from 1RU 32 port system analysis

In Table 6-1, the results of the comparisons can be seen. Due to the enhancements of the QSFP-DD1600 module, the system was able to more efficiently cool the nominally equivalent modules. This increased cooling resulted in modules operating at a lower temperature (as shown by the DSP margin). As a result, the module power itself was reduced which provided a significant 5% overall system power saving.

6.2 System improvements with QSFP-DD1600 in a 64 port 2RU

A second, more challenging system design was analyzed next using a 64 port, 2RU system design. The same modules described in Section 6.1 were used.

Both wind tunnel configurations and boundary conditions were matched with equal wind tunnel flow of 6.4-6.5 CFM and equal wind tunnel pressure drop of 1.08” H20.

The belly to belly stacked connector configuration was consistent with the one previously studied and shown in Figure 5-3.

The full 2RU 64 port system simulation is shown in Figure 6-5. The modules were again representative of a 30W coherent module. Also shown in Figure 6-5 are the surface temperature plots on the QSFP-DD800 and QSFP-DD1600 modules showing that the QSFP-DD1600 module is noticeably cooler. The test conditions for the simulations were consistent with the 1RU system described in Section 6.1 with the exception that the fan throttle was at 45% and a dust filter was included with a 65% arrestance.

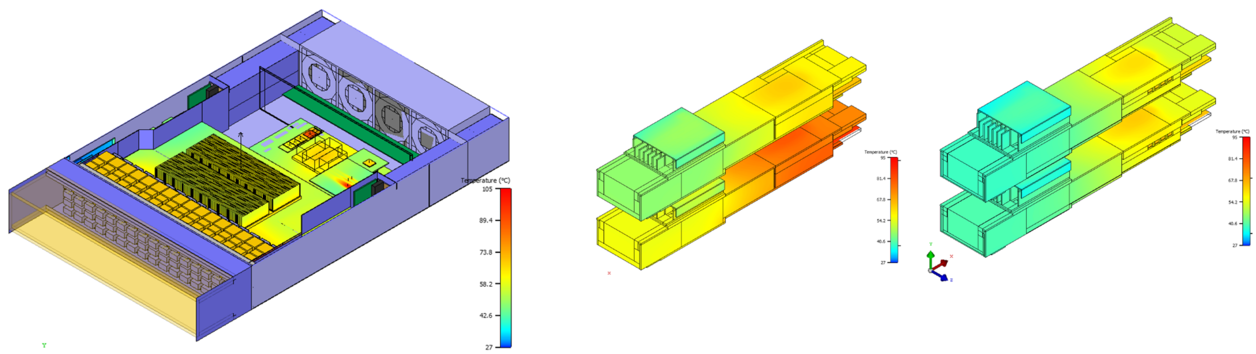


Fig. 6-5: Full 2RU 64 port system simulation (left) with either QSFP-DD800 modules (middle) or QSFP-DD1600 modules (right)

Module	Air Flow	NPU Margin	DSP margin (upper/lower)	Module Pwr (upper/lower)	System Power
QSFP-DD800	197.8 CFM	+10.1 °C	+36.6 °C/+18.6 °C	28.5W/25.5W	2095W
QSFP-DD1600	204.7 CFM	+12.0 °C	+39.5/36.9 °C	25.4W/25.0W	1995W

Table 6-2: Results from 1RU 32 port system analysis

In Table 6-2, the results of the comparisons can be seen. Again, due to the enhancements of the QSFP-DD1600 module, the system was able to more efficiently cool the nominally equivalent modules. This resulted in the module operating at a lower temperature (as shown by the DSP margin). As a consequence, the module power itself was reduced which provided a 5% overall power savings in the system which is significant considering everything else was equal. Notice the distinct reduction in differences between upper and lower port which is a typical challenge of stacked cage designs.

Sections 6.1 & 6.2 highlight that the enhancements defined for QSFP-DD1600 improve the ability to cool the pluggable modules to such an extent that overall the power of the system can be reduced compared to what was capable with QSFP-DD800 modules.

6.3 System cooling feasibility study with 50W QSFP-DD1600 in a 64 port 2RU

With all the thermal advancements and improvements identified for QSFP-DD1600, the whitepaper has focused on studying what would be considered typical configurations and with module powers that were high but not extreme. This study investigated whether a 50W module could be feasibly cooled in a system. Many of the enhancements identified for QSFP-DD1600 are additive in their impact and this study investigated what could be possible if such a design point existed. While this exceeds any known module power roadmap, it is important to understand the limits of this technology. As described in Sections 6.1 and 6.2, the enhanced efficiencies that are possible with QSFP-DD1600 are much more likely of interest to most network operators.

In the first series of simulations, a 2RU 64 port system was configured with 50W modules with air flow from the 2RU system with 90% fan throttle at 45°C+1829m. The wind tunnel airflow was 27.0 CFM with pressure of 3.13" H20.

Two options for module design were studied using different techniques described earlier. The prime difference is the use of a TIM on the riding heat sink to improve the thermal contact in Option 2. and with that advantage, other aspects of the module design were relaxed to reduce cost. This demonstrates the versatility and flexibility that QSFP-DD1600 based systems can provide:

Option 1: Dry metal-metal heat sink contact (4.5 C-cm^2/W); All copper heat sinks for cage; copper nose heat sink; embedded vapor chamber or heat pipe in lid; aluminum alloy housing.

Option 2: Pluggable-compatible TIM heat sink contact (1.8 C-cm^2/W); All aluminum heat sinks for cage; aluminum nose heat sink; embedded copper spreader in lid, zinc alloy housing.

Module	DSP Margin (upper/lower)	Optics (laser) margin (upper/lower)
Option 1	+4.7 °C/+0 °C	+3.0 °C/+0.5 °C
Option 2	+4.9 °C/+0.9 °C	+13.1 °C/+11.3 °C

Table 6-3: QSFP-DD1600 50W Module performance in 2RU system

Table 6-3 shows that it is feasible to cool a 50W module with reasonable airflow conditions using a QSFP-DD1600 module design. Unfortunately, building and populating a 2RU system with 64 ports of 50W isn't practical with today's power supplies nor is it consistent with today's typical port configurations, where a switch would not be fully populated with coherent modules.

To complete the study, a more realistic port fill configuration of 32 high power (50W) modules and 32 low power Active Copper Cables was simulated in the 2RU 64 port configuration. The 50W optical modules were placed in the lower ports of the stacked cage as a worst case environment. The only difference in the system configuration was the fan throttle was reduced to 85%.

The results are shown in Figure 6-8 and Table 6-4.

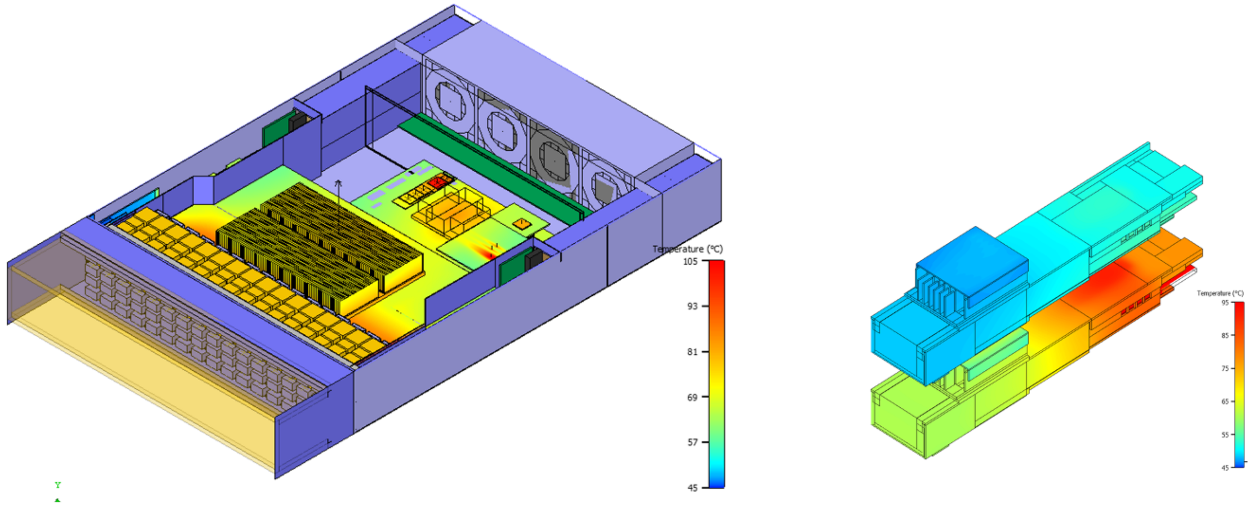


Figure 6-8: Full 2RU 64 port system simulation (left) with AEC modules (7W) in upper port and Coherent Optical modules (50W) in lower port (right)

Air Flow	NPU Margin	DSP margin	Optics margin
408.2 CFM	+0.0 °C	+3.2 °C	+11.8 °C

Table 6-4: QSFP-DD1600 50W Module performance in 2RU system populated with 50% optics and 50% AEC.

6.4 System cooling feasibility study with a 30W QSFP-DD1600 module in side-to-side airflow

For the final system thermal analysis incorporating QSFP-DD1600, we return to the side-to-side airflow configuration. Simulations were conducted to study various design options that can be applied to the system solution. These design options are mostly consistent with the design options studied previously with the QSFP-DD1600 in a front-to-back airflow approach. This further reinforces the system design flexibility of QSFP-DD1600.

In this study, a 30W QSFP-DD1600 was studied consistent with a coherent module design which includes the dominant DSP and optics (ITLA) heat sources. As shown in Figure 4-9, the riding heat sink design allows cooling with the side-to-side airflow.

A number of design variants were considered which are consistent with the earlier studies (increased airflow, improved case to heat sink thermal interface, and bottom side cooling) but also include some designs options unique to the side-to-side airflow configuration:

- integrated vapor chamber in riding heat sink along length

- cowling mechanism around module nose which serves to control the leakage airflow through faceplate to more effectively cool the module nose
- Increased heat sink dimension in length and width

The options summarized in Table 6-5 were considered as variables.

Option	Units	Value (baseline cooling)	Value (improved cooling)
bulk air speed	m/s	3.0	6.0
case to heat sink thermal interface	°C·cm²/W	6.5	3.2
secondary side cooling		no	yes
vapor chamber extends to front		no	yes
nose cowling		no	yes
heat sink width (streamwise direction)	mm	19.3	38.5

Table 6-5: Study inputs (design options) in side-to-side flow architecture

Figure 6-9 shows the module surface temperatures of the variants noted above. The application of multiple cooling improvements results in a 9°C benefit over the DSP, and 18°C at the nose, adjacent to the ITLA.

A detailed quantitative summary is provided in Section 8.

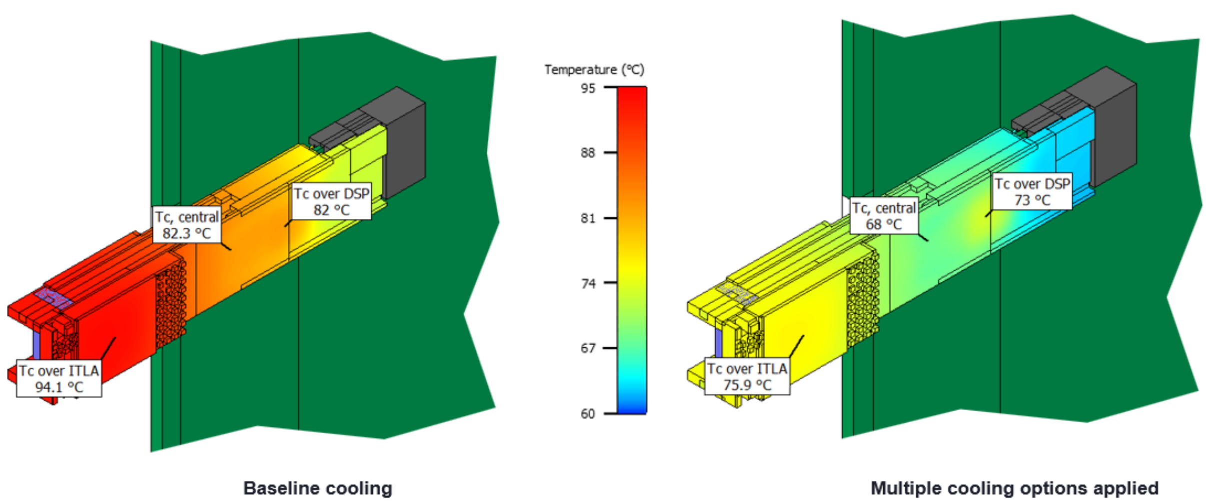


Figure 6-9: Surface temperatures of two different cooling implementations in side-to-side airflow. Module power = 30 W.

With these design options applied, thermal margin improvements result in an effective cooling solution for a 30W module. This initial analysis was considering cooling a single module. At the higher-power end of the scale, these design options are capable of cooling a 40W module but with near-zero margin (See Section 8 for details).

A distinct difference from the front-to-back airflow architecture is that the internal device with least margin in side-to-side airflow is the laser package (ITLA), not the module's DSP device. This has ramifications regarding the usefulness of the case top temperature monitor and it is recommended that the improved DOM reporting approach identified in Sec 3.1 be used.

The outcome of this analysis confirms that cooling solutions, even for module powers at 30W or higher are feasible.

6.5 System cooling feasibility study of multiple 30W QSFP-DD1600 modules in side-to-side airflow

When multiple QSFP-DD modules are used on a line card, the results of the previous Section 6.4 are roughly applicable to only the most upstream module. A system analysis was conducted with 4 modules in the airflow as shown earlier in Figure 4-12 and the various design options were studied.

The design challenge with multiple modules in a side-to-side airflow is that any downstream module encounters pre heated air from the upstream modules. A conclusion from the analysis is that any downstream module experiences a thermal penalty of 1-2°C for every module upstream. For example, if 4 modules are used on a card, the last module in the air stream is 3-6°C hotter than the first. As such, a system design with 4 x 30W modules might be under-cooled, and it is incumbent upon the system designer to determine the set of parameters that is viable; e.g. perhaps 4 x 28W modules at 40°C and 900m. More details are provided in Section 8.7.

7 Conclusion

As pluggable modules approach 1.6 Tb/s and systems that utilize those ports speeds are designed, it is clear that a comprehensive thermal strategy is required to successfully build reliable products. ASIC capacities and powers are increasing, port counts are increasing, and space constraints still exist. An overall successful solution requires every module or system designer to be able to optimize their specific designs and have those designs to be compatible and interoperable with each other.

This is where QSFP-DD's performance strength and leadership lies.

QSFP-DD's initial design has always been based on a flat-topped module without an integrated heat sink. While initially considered a weakness, once understood properly, this is a strength of the design. It frees both the module designer and the system designer to innovate in their own designs with a known point of interconnection between the designs.

This has been shown and demonstrated across QSFP-DD and QSFP-DD800 where every type of module and every type of system configuration that has been conceived can be built and deployed.

With QSFP-DD1600, this flexibility has further enhanced the ability to build solutions that meet or exceed the market's needs and goals.

Through 400G and 800G, much of the focus was on the ability to cool the ever increasing module powers that were needed for the technology. While for 1.6T modules this is again true, the overall reduction in power of a fully loaded system is top of mind for all network operators.

As described in this whitepaper, there are a wealth of improvements around QSFP-DD1600, either specifically in the MSA specifications or at the system level with how the devices can be used and cooled.

As the typical powers of the high-volume modules continues to rise into what was previously considered the "high-power" ranges, it is imperative that solutions can cool these modules with greater efficiency than previously. And at the same time, a solution needs to be ready to handle both the high-volume modules and the equally important coherent modules which may have higher powers than the shorter reach modules.

These improvements are a spectrum of options and possibilities enabling the module designer or the system designer to implement for purpose optimizing for their designs on such metrics as performance or cost. This whitepaper covers many of these options and builds upon the previous [QSFP-DD800 thermal whitepaper] which started exploring these design optimizations. QSFP-DD's inherent flexible design platform gives us these additive improvements.

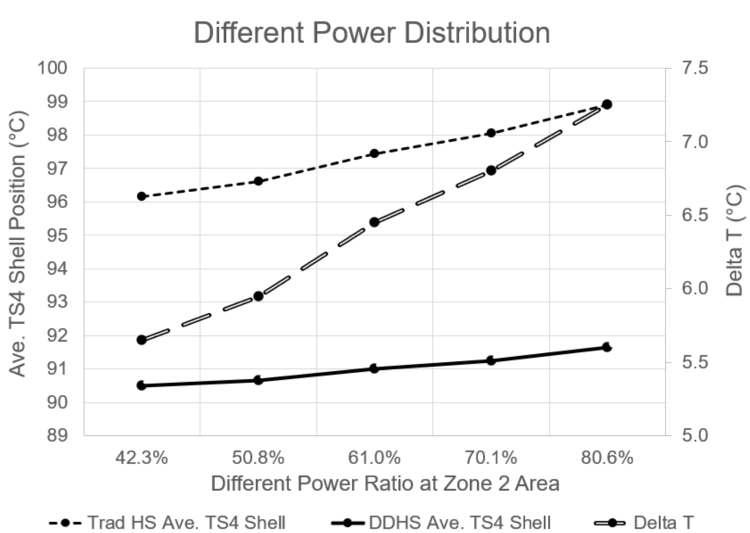
Most strikingly, it has been shown in this whitepaper that the enhancements available in QSFP-DD1600 can reduce the overall system power (compared to an identical system using QSFP-DD800). This improvement comes from the QSFP-DD1600 thermal enhancements that improve the efficiency when cooling these modules. If it takes less effort to cool a module, you can reduce the system fan powers accordingly or instead maintain fan powers resulting in cooler optical modules which consume less power (and likely have improved reliability).

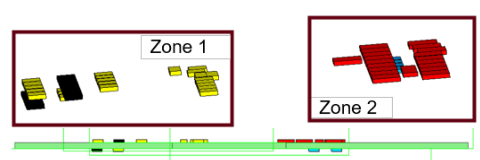
8 Additional details

In this section, further details are available for the motivated reader to acquire more information on some of the previously discussed results and analysis.

8.1 Sensitivity analysis of module heat source locations

One of the challenges (and advantages) of designing these solutions for thermal operation is the high degree of variability that is possible by module and system designers. As described previously, the simulations in this whitepaper were mostly conducted using a 40W MultiLane thermal model, but the specific power profile in the MultiLane module may not be consistent with all the variations of module design that a system would deal with. To study the impact of any variation, a simulation was conducted where the percentage of power in the module was varied from zone 1 (towards front of module) to zone 2 (towards back of module). With many optical module designs it is often more common to see more power towards zone 2 consistent with the placement of the DSP device. Figure 8-1 shows the relative improvement in performance with a DDHS as the power profile shifts towards zone 2, for example if ratio of power in zone 2 increases from 42.3% to 80.6% we see an appreciable increase in the delta improvement that DDHS can offer (from 5.3°C to 7.3°C) This isn't surprising as the riding heat sink makes contact more towards the rear of the module and this is an good example of how reducing the thermal impedance of a module heat path from their heats source to the heat sink contact is always a good idea as explored in the QSFP-DD800 thermal whitepaper.



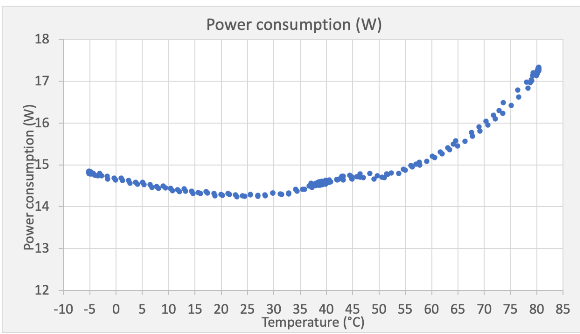


Zone 1 Power (W)	Zone 2 Power (W)	Zone 2 Power % of Total Power	Total Power (W)
23.12	16.96	42.3	40.08
19.69	20.35	50.8	40.04
15.62	24.42	61.0	40.04
11.95	28.09	70.1	40.04
7.77	32.30	80.6	40.07

Figure 8-1: Sensitivity analysis of Drop Down heat sink on module power profile variations

8.2 Module power temperature dependency

A key observation about optical module power dissipation is that it is a function of the module temperature. Keeping the module cooler results in a lower power dissipation of the module. By improving the cooling efficiency of the QSFP-DD1600 it is possible to operate the modules at a lower temperature resulting in an overall system power savings. Figure 8-1 shows two representative curves of both an IMDD and coherent module.



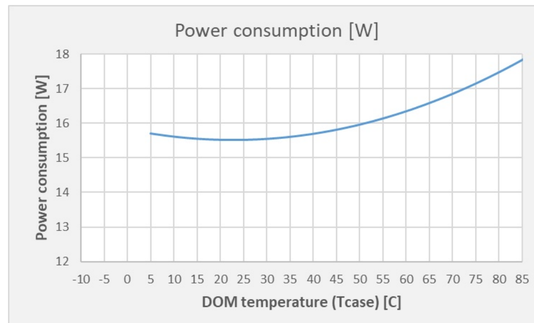


Figure 8-2: Example power versus module temperature for 800G IMDD (left) and 400ZR coherent (right) modules showing a typical relationship of power vs temperature.

8.3 System improvements with QSFP-DD1600

In Section 6.1, an system analysis of a 32p 1RU design fully populated with either QSFP-DD800 based modules operating a nominal 30W worst case or QSFP-DD1600 modules also operating at a nominal worst case 30W was studied.

The QSFP-DD800 and the QSFP-DD1600 modules used in the analysis are shown in Fig 6-1 and shown again here in Figure 8-3. More details on the modules used in this analysis are shared in this Section.

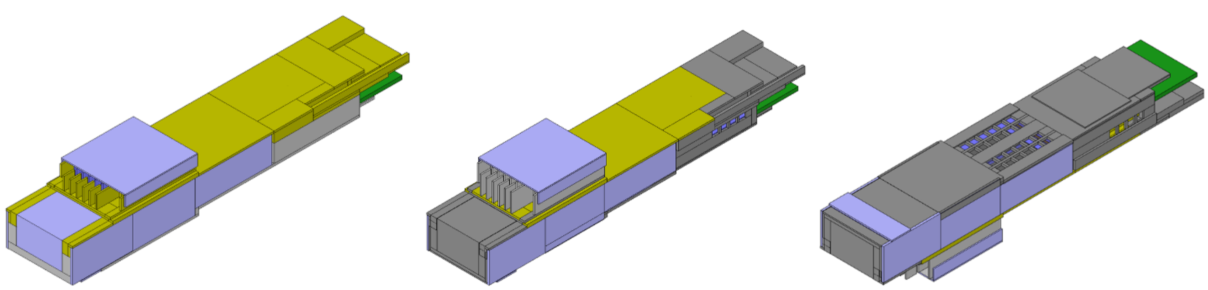


Figure 8-3: QSFP-DD800 module (left); QSFP-DD1600 module (middle); QSFP-DD1600 module underside (right)

The QSFP-DD800 module was consistent with the MSA specification and had the following features: Standard cage, PCB and module housing, Type 2B nose heat sink, Copper nose heat sink fins and housing lid; and Aluminum alloy housing base.

The QSFP-DD1600 module was consistent with the enhancements proposed in the MSA specification and in this whitepaper and had the following features: Cage with added vents at bottom, sides and back; Cage with bottom heat sink; PCB with slot and side notches; Type 2C nose heat sink (+1.87mm taller); Aluminum nose heat sink fins; Zinc alloy housing lid and base with minimal Cu spreader; Recess surface ducts on bottom and rear sides; Vent holes in housing at recesses on bottom and sides.

Both wind tunnel configurations and boundary conditions were matched with equal wind tunnel flow of 6.4-6.5 CFM and equal wind tunnel pressure drop of 1.08” H20.

The two belly to belly configs are shown in Figure 8-4.

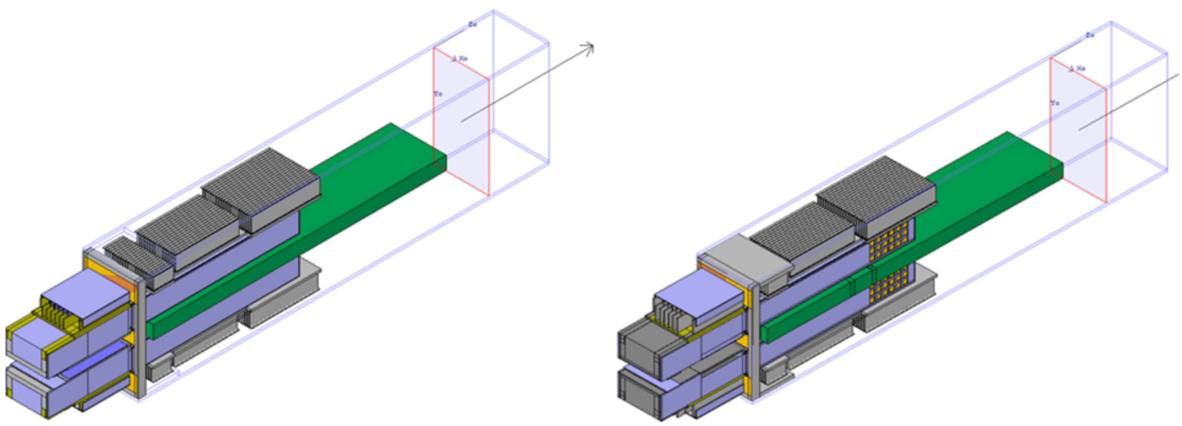


Fig: 8-4: QSFP-DD800 (left) and QSFP-DD1600 (right) configurations used in full 1RU 32 port simulations

8.4 System improvements with QSFP-DD1600 in a 64 port 2RU

As described in Section 6.2 a system analysis of a 64 port 2RU design fully populated with either QSFP-DD800 based modules operating a nominal 30W worst case or QSFP-DD1600 modules also operating at a nominal worst case 30W was studied.

Module details are consistent with what was shared above and the port configurations are shown in Figure 8-5.

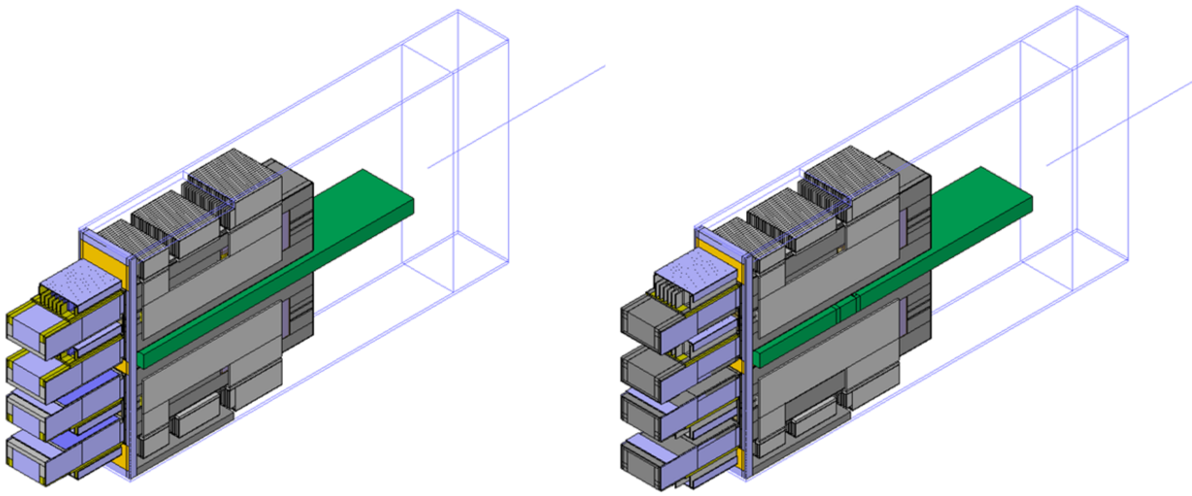


Fig: 8-5: QSFP-DD800 (left) and QSFP-DD1600 (right) configurations used in full 2RU 364 port simulations

8.5 System cooling feasibility study with 50W QSFP-DD1600 in a 64 port 2RU

As described in Section 6.3, an analysis was done to investigate whether it was feasible to consider being able to cool a 50W QSFP-DD1600 in a dense system configuration. Some additional details of the modules used in this analysis are shown below in Figure 8.6.

Option 1: Dry metal-metal heat sink contact (4.5 C-cm^/W); All copper heat sinks for cage; copper nose heat sink; embedded vapor chamber or heat pipe in lid; aluminum alloy housing.

Option 2: Pluggable-compatible TIM heat sink contact (1.8 C-cm^/W); All aluminum heat sinks for cage; aluminum nose heat sink; embedded copper spreader in lid, zinc alloy housing.

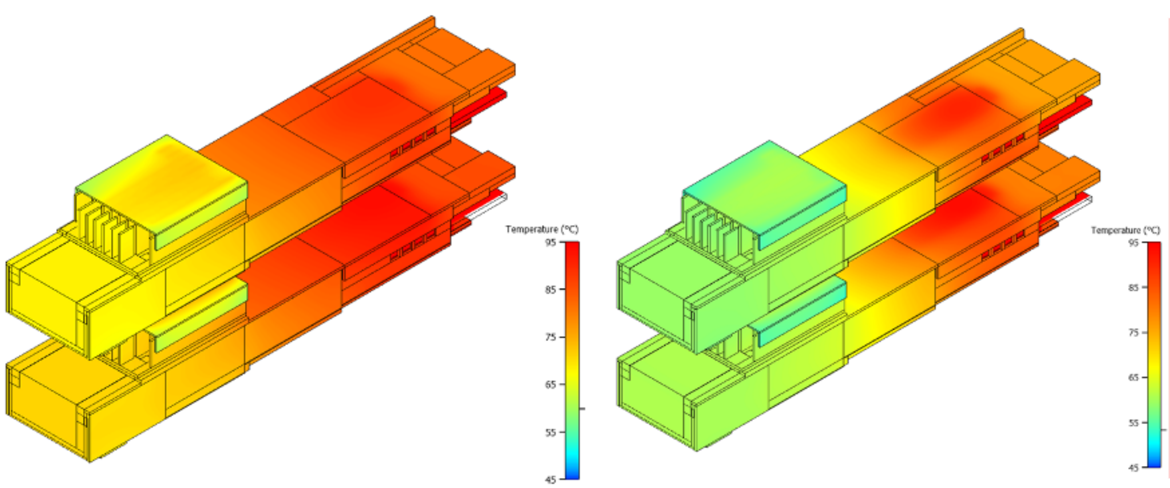


Figure 8-6: 50W QSFP-DD1600 modules: Option 1 (left) and Option 2 (right)

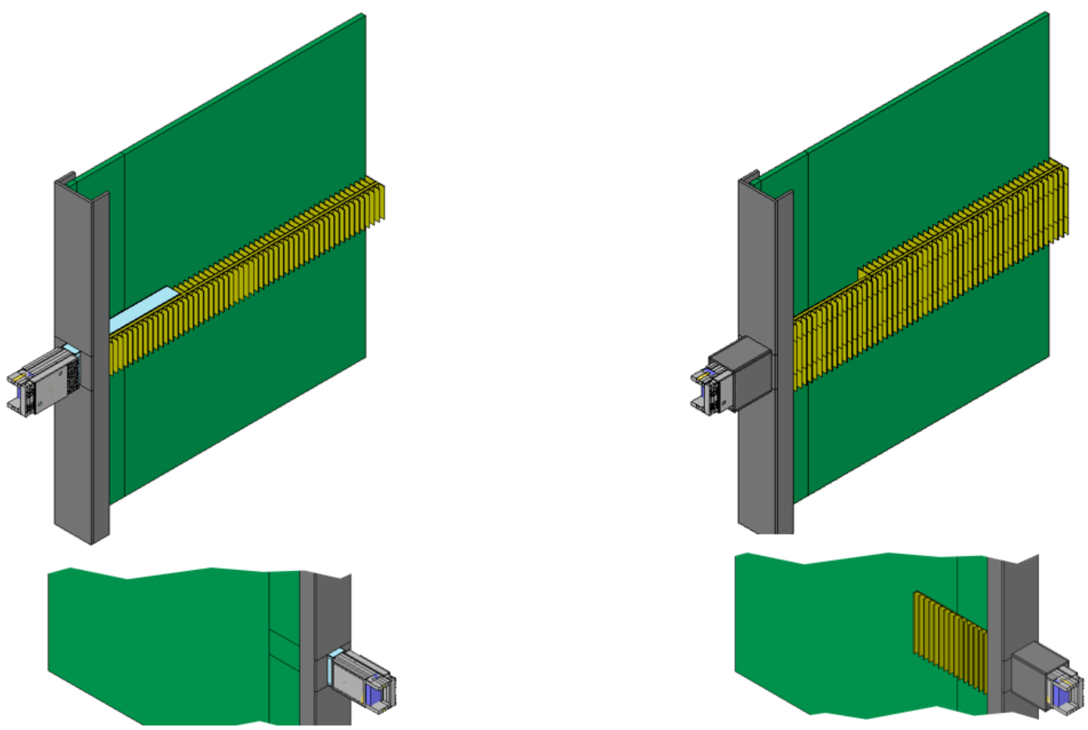
8.6 System cooling feasibility study with a 30W QSFP-DD1600 module in side-to-side airflow

As described in Section 6.4, analysis was conducted on the design trade-offs that would enable support of QSFP-DD1600 modules in a system utilizing side-to-side airflow.

The simulations were conducted to study the relative benefits of various design options available in side-to-side airflow architecture. The following conditions were common among all simulations:

- Pull-flow system
- Ambient temperature = 45°C
- Altitude = 1829 m
- PCB size = 200x200 mm
- Line card to line card pitch = 25 mm
- QSFP-DD thermal model includes DSP and ITLA
- QSFP-DD total power = 30 W
- DSP power = 17W
- Case material conductivity = 180 W/mK
- heat sink is a vapor chamber design

Some extra diagrams showing the configurations considered are shown in Figure 8-7.



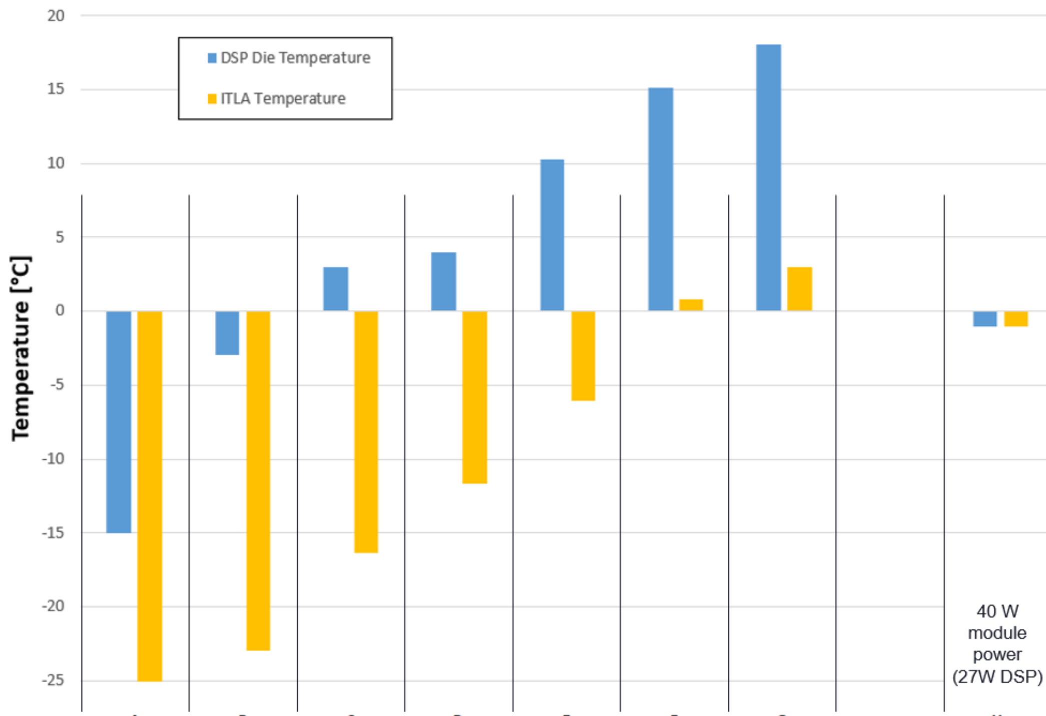
Baseline cooling

- bulk air speed = 3 m/s
- case-to-heatsink thermal resistance = 6.4 °C·cm²/W
- no secondary side cooling
- vapor chamber truncated at mounting clip
- no nose cowling
- heat sink width = 19.3 mm

Multiple cooling options applied

- bulk air speed = 6 m/s
- case-to-heatsink thermal resistance = 3.2 °C·cm²/W
- secondary side cooling
- vapor chamber extends all the way forward
- nose cowling
- heat sink width = 38.5 mm

Figure 8.7: Illustration of two different cooling implementations in side-to-side airflow



Cooling Option	A	B	C	D	E	F	G	H
Bulk air speed [m/s]	3.0	3.0	3.0	3.0	3.0	6.0	6.0	6.0
Case to HS resistance [°C·cm²/W]	6.5	3.2	3.2	3.2	3.2	3.2	3.2	3.2
Secondary side cooling	n	n	y	y	y	y	y	y
Vapor chamber extends to front	n	n	n	n	y	y	y	y
Nose cowling	n	n	n	y	y	y	y	y
Heat sink width [mm]	19.3	19.3	19.3	19.3	19.3	19.3	38.5	38.5

Figure 8-8: Device margins when using various combinations of cooling options; Module power = 30 W except where noted

The detailed results are shown in Figure 8-6 showing the temperature margin that resulted in each configuration for the module's DSP device or laser package (ITLA). It is readily apparent that the baseline cooling variant is under-cooled, and that at least some of the suggested options need to be applied. With more cooling options applied, thermal margin improves leading to a cooling solution for a 30W module. At the higher-power end of the scale, we see that a 40W module with a 27W DSP can be cooled with near-zero margins which suggests that a cooling solution is within target.

8.7 System cooling feasibility study of multiple 30W QSFP-DD1600 modules in side-to-side airflow

As described in Section 6.5, a further analysis was performed to study the impact of using multiple QSFP-DD modules on a line card. The configuration is shown again in Figure 8.9 and includes four widely spaced modules. In this analysis, all cooling improvements identified in Table 6.5 are applied.

The results of the analysis are shown in Figure 8-10. As noted earlier, the upstream module receives the airflow with the lowest pre-heating whereas the downstream ports all have increasing pre-heated airflow to deal with. In this analysis the upstream module is port 4 and is cooled with adequate thermal margin, while the most downstream module (port 1) slightly exceeds its rated limit, with the ITLA being the device with least-margin. Note that the circuit card in this study omits other circuitry; e.g. power supplies, CPU, FPGA, etc., and thus the margins would need to be studied with a full system analysis. As such, a high port count of high power QSFP-DD modules is challenging to cool in side-to-side airflow. Nevertheless, this analysis shows that it is feasible to consider support of multiple high-power QSFP-DD modules in a side-to-side airflow and that the system designer can use the inherent design flexibility that QSFP-DD provides to come up with a solution. With the assumptions included in this particular configuration, it would appear that a gang of three 27W modules at 45°C at sea-level .

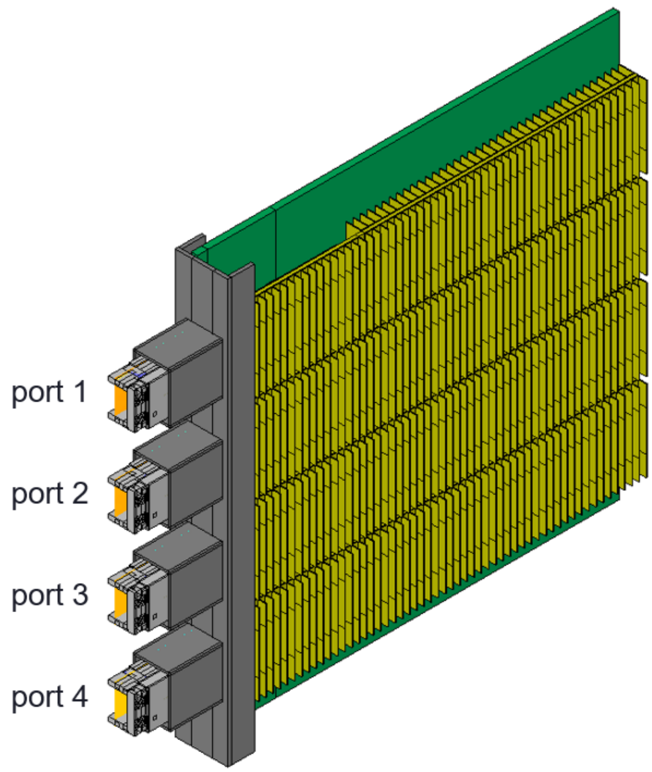


Figure 8.9: A configuration with 4 modules, generously spaced to allow for wide heat sinks (and all available cooling options)

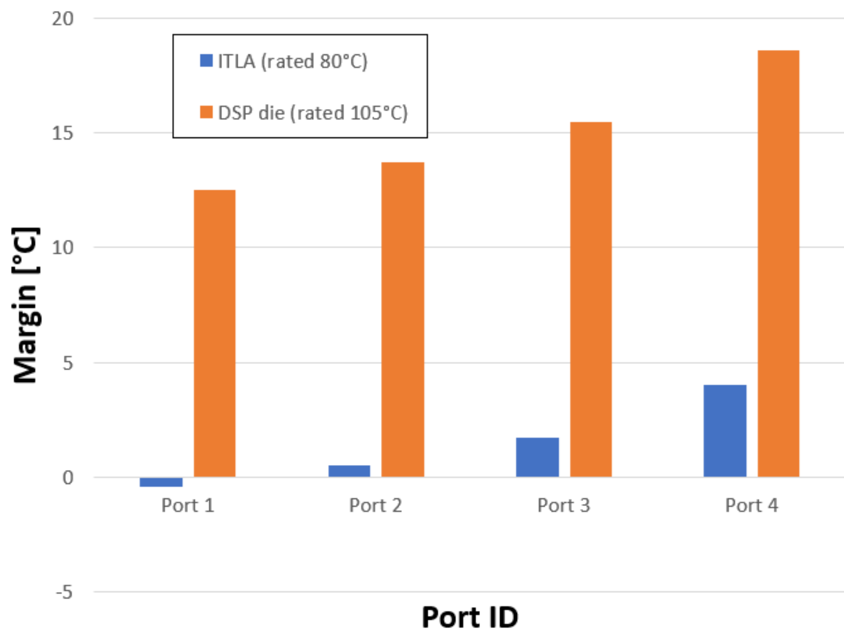


Figure 8-10: Device margins of a multi-module (4) configuration; various cooling options applied; module power = 30