



OIF OPTICAL
INTERNETWORKING
FORUM

Thermal Management at the Faceplate White Paper

ABSTRACT:

This whitepaper represents the work of the OIF to consider the system issues for thermal management at the faceplate of a line card. Issues associated with air cooling of pluggable modules in a line card are examined including a brief tutorial of the basic physics, examples of methods that can be used to improve cooling and discussion of design limitations and trade-offs. Guidance on the necessary communications between optics plug suppliers, the system architects and system thermal designers is provided.

Table of Contents

| | | |
|--|---|----|
| 1 | Executive Summary | 3 |
| 2 | Introduction | 4 |
| 3 | Project Overview | 6 |
| 4 | Detailed Thermal Modeling | 10 |
| 5 | Discussion of Contact Resistance Issues | 21 |
| 6 | Recommendations | 23 |
| 7 | Summary | 25 |
| 8 | References:..... | 26 |
| Appendix A: List of Contributors | | 27 |
| Appendix B: About the OIF..... | | 27 |

List of Figures

| | |
|---|----|
| Figure 1: Main Heat Removal Paths from Optical Plugs | 5 |
| Figure 2: Typical Blade Airflow Patterns | 7 |
| Figure 3: QSFP Air Temperatures for Cases 1 to 3 of Figure 2 – Air Inlet 40°C at 3000m. | 9 |
| Figure 4: Module Thermal Resistance Through Heatsink | 10 |
| Figure 5: The QSFP Flotherm® Model | 11 |
| Figure 6: The Simplified Switch Blade Model Showing Case 3 Baffles and Axial Location of QSFP's..... | 12 |
| Figure 7: Model Cross-section Through QSFP – Open Area Total =3179mm ² Above Board=2481mm ² Used in Cases 1 through 7..... | 13 |
| Figure 8: Detailed Thermal Modeling Cases | 15 |
| Figure 9: Air Temperatures With and Without Baffles at Plane at Mid-height of Heatsink Fins: Thermal Modeling Cases, 2, 3, 4 and 5. | 16 |
| Figure 10: Baffle Design for Cases 3, 6 and 7 | 17 |
| Figure 11: Baffles Used with Extended Heatsinks - Cases 4 and 5 | 18 |
| Figure 12: Extended Heatsink Temperatures for Downstream QSFP's – a) Aluminum Heatsink per Case 4 – b) Al Heatsink with Embedded Nanosreader. Sections are through second QSFP from the left..... | 18 |
| Figure 13: Air Flow for Cases 3, and 6 with Flow Under the Board a), and Case 7 with Flow Under the Board Blocked b). Section is at Mid-height of the Fins | 19 |
| Figure 14: Small Flow Channel for Cases 8 through 15, Section through QSFP and Plan View | 20 |
| Figure 15: Effects of Changing Airflow in the Small Channel Cases 8 through 11 | 20 |
| Figure 16: Effects of Changing the Applied Contact Resistance in the Small Channel Cases 12 through 15..... | 21 |

List of Tables

| | |
|---|----|
| Table 1: Optical Module Form Factors versus Power Density | 5 |
| Table 2: Summary of CFD Model Results | 14 |
| Table 3: MSA Plug Thermal Interface Specifications | 23 |

1 Executive Summary

The power and therefore heat dissipation of optical pluggable modules is expected to increase at the same time as plugs are reducing in size and increasing in number per blade. As a result their heat flux is expected to increase from 2 to 5 times over the next 2 to 4 years. Current thermal designs are approaching their functional limits. The objective of this White Paper is to provide recommendations and guidance on necessary technical exchange among optics plug suppliers, system architects and system thermal designers to provide improved thermal management of systems with optical ports.

With the aid of a detailed conjugate heat transfer model of a QSFP optical plug module, a series of analyses have been conducted on a simplified switch blade platform. On this basis, recommendations for cooling the high power modules expected in the near future were produced. The analyses identified that front to back and channeled side-to-side airflow can provide the lowest temperature air for cooling plug modules. Increasing air flow can deliver reduced air temperature for downstream components but has lesser benefit for upstream components. By utilizing PCB layouts that separate high dissipation and hotter components from the plugs, air temperatures arriving at plugs are kept low as are board temperatures in the region of the plug.

Air flow in the regions of plugs can be improved by including baffles. Baffles can be used to separate plug air from hot air over the ASIC's, increase the speed of air at the plugs, and improve flow mixing to reduce hot air pockets or air streams in the regions of the plugs. Baffles can also be used to increase flow locally in the faceplate region and thereby increase the performance of heatsinks.

Increasing the size of heatsinks on devices usually makes major improvements to heat removal capacity. Heatsinks for current plugs have planform areas similar to the plug, however the expected increase in heat flux means that thermal resistance must be reduced to achieve plug temperature limits. The convective component can be reduced by increasing the surface area and/or the air speed. For plugs in gang cases, increasing the heatsink area necessitates extensions rearward from the faceplate, and can bring large temperature gradients from front to back in the heatsink. It is likely that heatpipes or advanced highly conductive materials must be embedded in the base to reduce this temperature variation, i.e. the conductive component to an acceptable level.

The requirement to remove plugs means that only small forces are available to reduce interface resistance between heatsink and plug. However, contact resistance literature shows that it can be easily reduced by reducing surface roughness as well as increasing the contact pressure. Separate from contact areas, gaps between the mating surfaces form a parallel thermal resistance. Since many Multi-source Agreements (MSA's) do not even mention surface roughness of the plug or heatsink they require review to ensure sufficiently tight tolerances are included. In addition, flatness tolerances are larger than can be achieved through standard fabrication techniques and tighter tolerances would reduce gap heights between the mating faces.

The analyses showed that through the combined actions of separating air streams, increasing flow velocity at the plugs, keeping high dissipation components located away from faceplate area, increasing heatsink surface area, and decreasing the interface resistance between the heatsink and plug it is possible that case temperature limits for plugs can be met at the expected higher powers when ambient is 40°C. For telecommunications applications where 55°C is the maximum ambient this will be much more difficult unless contact resistances are significantly reduced. The analyses did not include heat dissipation in the board near the plugs so they are optimistic in that respect.

2 Introduction

The mission of the Optical Internetworking Forum (OIF) is to promote the development and deployment of interoperable networking solutions and services for optical networking products, network processing elements, and component technologies. The Physical and Link Layer (PLL) Working Group develops Implementation Agreements and White Papers related to physical and data link layer interfaces between Optical Internetworking elements and between their internal components, and reuses existing standards when applicable. The purpose of this white paper is to identify thermal issues specific to air-cooled pluggable optical modules and propose methods for surmounting these issues. In particular, it provides recommendations and guidance on necessary technical exchange among optics plug suppliers, system architects and system thermal designers to provide improved thermal management of systems with optical ports.

As described in References 1 and 2, the problem is that the accessible heat transfer area of the individual modules – generally restricted to the upper surface area of the module – is such that power surface densities are expected to increase from ~ 0.1 W per square centimeter (W/cm^2) to 0.3 or 0.5 W/cm^2 over the next 2 to 4 years. Currently, switch bandwidth has reached terabits per second (Tbps) and is defined by the number of switch blades and the bandwidth of each switch blade.

The smaller size of second generation 100GbE modules will enable higher port counts on switch blades so that individual switch blades can support a Tbps. With second generation modules 8, 16, 32 or even 44 module ports can be placed on a single blade to yield the Terabit throughput. A consequence of the reduced size of these smaller second-generation 100GbE modules is the high power density of the modules and blades. Table 1 lists the expected power densities for various module form factors and shows that most anticipate power densities of over $0.3\text{W}/\text{cm}^2$ (2 W/sq in). This will require improved thermal design, both at the system level and at the plug level.

Table 1: Optical Module Form Factors versus Power Density

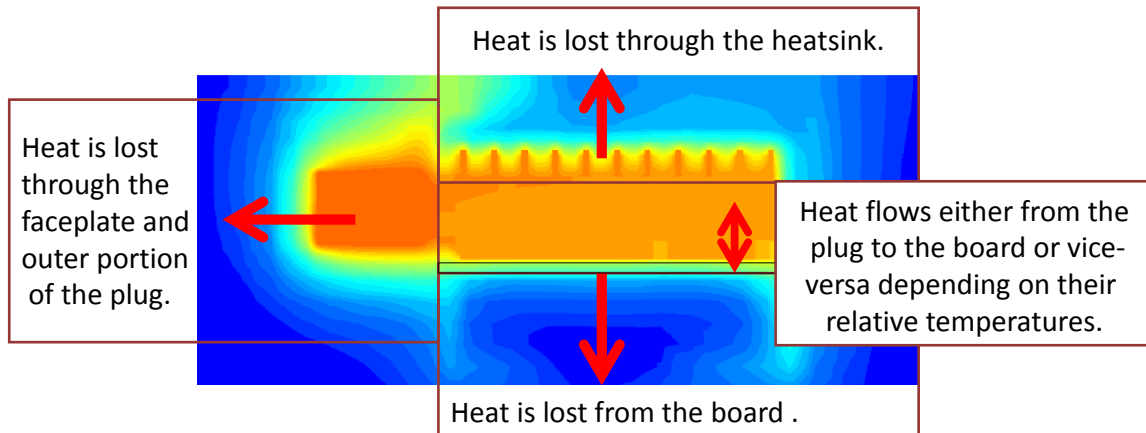
| Module Form Factor | Top Face Dimensions mm | Top Face Surface Area cm ² (sq inch) | Power W | Power Density Range W/cm ² (W/sq in) |
|--------------------|---------------------------|--|-----------------|--|
| SFP+ | 47.5 x 13.6 | 6.5 (1.0) | 1.0/1.5 | 0.15 to 0.23 (1.0 to 1.5) |
| QSFP | 52 x 18 | 9.4 (1.45) | 1.5/2.5/3.5/5.0 | 0.16 to 0.53 (1.0 to 3.5) |
| XFP | 70 x 18.4 | 12.9 (2.0) | 1.5/2.5/3.5 | 0.12 to 0.27 (0.75 to 1.75) |
| CFP4 † | 66 x 21.7 | 14.3 (2.2) | 1.5/3.0/4.5/6.0 | 0.10 to 0.42 (0.7 to 2.7) |
| CFP2 † | 84 x 41.5 | 34.9 (5.4) | 3.0/6.0/9.0/12 | 0.086 to 0.34 (0.6 to 2.2) |
| CFP | 114 x 76 | 86.6 (13.4) | 8/16/24/32 | 0.092 to 0.37 (0.6 to 2.4) |

†The power classes and form factors for CFP2 and CFP4 modules are not finalized at time of writing but are current as of March 16, 2012. Ref. 3

Fiber optic modules or transceivers have unique thermal constraints because the laser reliability is dependent on maintaining relatively low case temperatures of under 70°C which is at least 15°C lower than other blade components. Pluggable optical modules have additional challenges since they must fit through a standard opening in the faceplate. This restriction severely limits the fin surface area and the effectiveness of heat spreading in the base when integrated heatsinks are used. For higher power density devices, heat transfer is insufficient to maintain case temperature limits. Accordingly, riding heatsinks are deployed to increase heat transfer area on the top side. However, they introduce what can be a poorly controlled thermal contact resistance between the module and the heatsink.

The main heat removal paths for a plug are depicted in Figure 1. In general, convective heat removal through the sides of plugs is restricted by the EMC cages, adjacent plugs and guide rails. Under the plug, the thin air gap between the plug and the board allows little air flow to remove heat. Conductive paths to the board through connectors, air and cages also allow transfer but the direction depends on the relative temperature of the plug and the board. Finally, some heat is also dissipated via natural convection from the parts of the plug that are outside of the faceplate.

Figure 1: Main Heat Removal Paths from Optical Plugs



The relative amount of heat removed from each of these paths is dependent on blade design but for a given plug and cage installation, the only ones that can be controlled are the loss through the heatsink and the loss through the board. Heat loss through the sides of plug in the center of gang cages is minimal due to the presence of heat dissipating plugs on either side. The natural convection cooling on the outer portion of the plug is also relatively small due to the low air speed.

3 Project Overview

To explore the consequences of these factors, this paper investigates a series of fifteen progressive Flotherm^{®1} models of a switch blade with eight 5-Watt Quad Small Form Factor Pluggable (QSFP) optical modules. Tight design constraints limit the cooling capabilities of an Ethernet switch blade and tradeoffs between airflow, heatsinks, baffles and contact resistance are discussed in this paper. Design issues and suggested solutions are generally applicable to all plug types.

An example layout of a switch blade with high power optical plugs is shown in Figure 2 where two 1x4 QSFP cages are mounted at the faceplate with an ASIC complex between the QSFP's and the back of the blade. Three cases for airflow for typical installations are illustrated and in practice when leakage is considered, flow through a blade is a combination of one or more of these. Note that the side-to-side flow is essentially the same as would occur in a switch where cards are mounted vertically with fans either top mounted to pull air through the blades from the bottom and exhausting out the top, or bottom mounted to push air in from the bottom.

The grouping of plugs used in the analysis is one of many possible and the drive to increase the bit rate per blade leads to a desire to locate plugs as closely as possible along the faceplate. For example, see Ref. 3 where it is proposed that up to 16 CFP4 or 18 QSFP28 modules be located side-by-side along the faceplate. When examined in light of the current results such layouts are seen to give a more severe thermal environment than that described here.

¹ Flotherm is a commercial software program from Mentor Graphics that is specialized for analyzing conjugate heat transfer in electronics cooling applications.

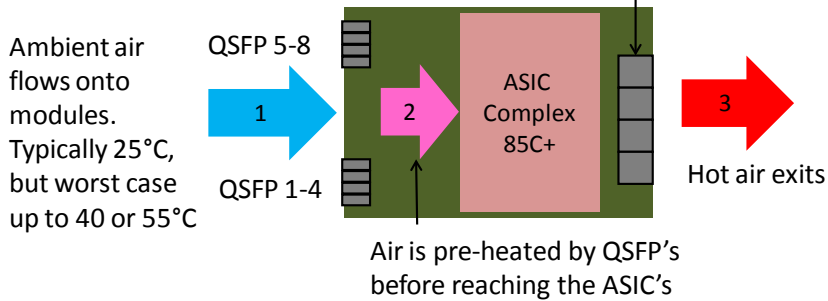
Figure 2: Typical Blade Airflow Patterns

Note: 40C becomes 55C for Telecommunications

Case 1: Front to Back Airflow

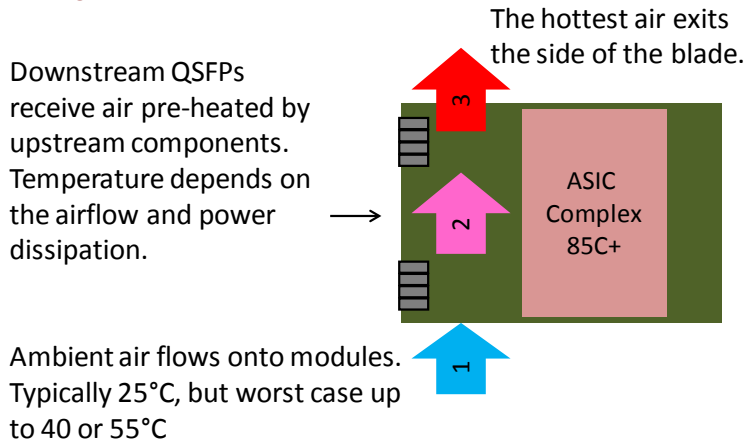
Best Air Temperature Scenario for 1U Switches

1U fan array can pull 10's of m³/h (CFM) of air through the Switch for nominal 2.5m/s (500LFM) airspeed.



Case 2: Side to Side Airflow

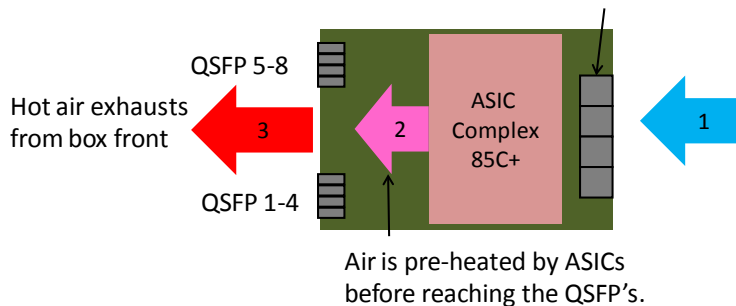
Average Case Scenario for Modular Switches



Case 3: Back to Front Airflow

Worst Case Scenario for 1U Switches

1U fan array can push 10's of m³/h (CFM) of air through the Switch for ~2.5m/s (500LFM) airspeed



For the optical plugs the local flow conditions are important. In cases 2 and 3, the air arriving at the plugs has been preheated by a combination of other board components or upstream QSFP's. The increase in the temperature of this air is governed by the equation:

$$\Delta T = P_{\text{air}} / (\rho \times C_p \times Q_{\text{air}}) \quad (1), \quad \text{i.e.}$$

$$\begin{aligned} \Delta T [^{\circ}\text{C}] &= T_{\text{air}} - T_{\text{amb}} \\ &= \text{Power [W]} / (\text{Density [kg/m}^3\text{]} \times \text{Specific Heat [J/kg-}^{\circ}\text{C]} \times \text{Airflow [m}^3\text{/s]}) \end{aligned}$$

Where the Power, P_{air} , is the portion of device power removed as heat from upstream components and transferred to the air stream, the Density of air is a function of temperature and altitude (i.e. atmospheric pressure), the Specific Heat is a constant for the fluid, and Airflow is the volume of air flowing over the modules. T_{air} is the local air temperature after upstream heat is absorbed, and T_{amb} is the temperature of air at the blade inlet.

The heat removed by convection from the module to the air is described by:

$$q = h \times A \times dT_{\text{conv}} \quad (2), \quad \text{i.e.}$$

$$\text{Heat Flow [W]} = \text{Heat Transfer Coefficient [W/}^{\circ}\text{C-m}^2\text{]} \times \text{Area [m}^2\text{]} \times (T_{\text{surf}} - T_{\text{air}}) [^{\circ}\text{C}]$$

The heat transfer coefficient is a function of the air properties, flow speed, and the geometry of the surface. Area is the surface area of the module or heatsink exposed to the air, and the driving temperature difference, $dT_{\text{conv}} = T_{\text{surf}} - T_{\text{air}}$, where T_{surf} is the temperature of the module or heatsink surfaces contacting the air. For a given geometry and fluid, the heat transfer coefficient increases with flow speed by

$$h \propto u_{\text{air}}^n, \quad (3)$$

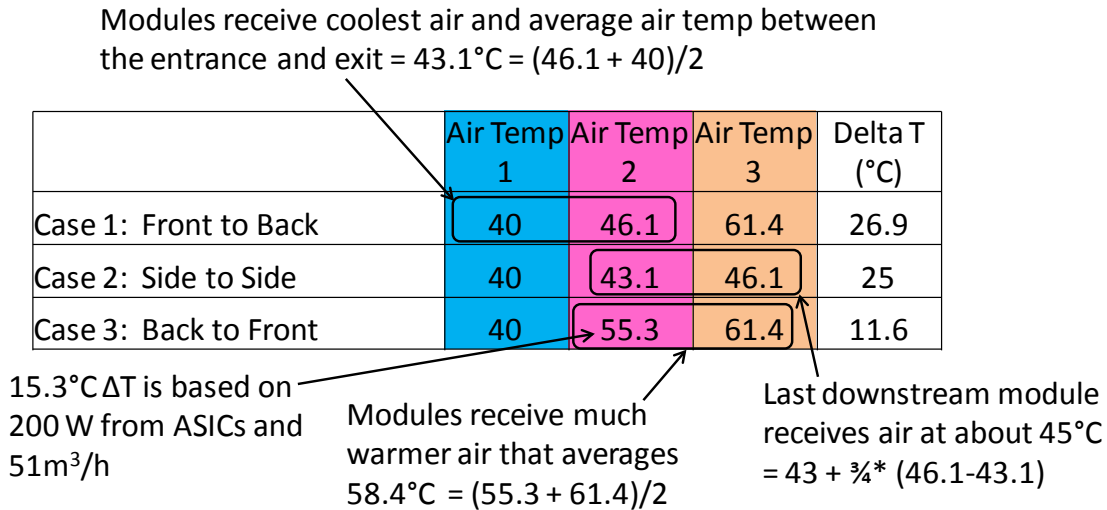
Heat Transfer Coefficient is proportional to (Air Speed)ⁿ, where for analyses used in heat exchanger design, n typically ranges from 0.33 for laminar flow to 0.8 for high speed turbulent flow in pipes, Ref. 4.

These equations show that increasing air flow will decrease T_{air} at the module and decrease dT_{conv} to improve module (heatsink) surface temperatures for downstream objects while objects at the inlet only benefit from the generally smaller improvement in heat transfer coefficient. This is often the simplest way to improve performance, but increases cost, energy use and noise.

To illustrate the benefits of wise choice of flow pattern, Reference 5 provides an example where the blade of Figure 2 has 50.4m³/h (30cfm) total flow of which 25.2m³/h (15cfm) flows over the QSFP modules. Inlet air is at 40°C with ambient pressure corresponding to 3000m (10,000ft). The QSFP's are assumed to dissipate 5W each, and the ASIC complex is assumed to dissipate 200W. The results, summarized in Figure 3, show that the average air temperature over the QSFP's is the same for Cases 1 and 2. However, the side-to-side flow of case 2 results in ~2.4°C higher average air temperature for the downstream module,

than for any in case 1. The back-to-front airflow of case 3 is markedly poorer with average air temperatures into the QSFP's that are 15.3°C higher than the front-to-back flow of case 1. The Delta T numbers in the figure represent the difference between the average air temperature over the worst-case module and the maximum case temperature of 70°C.

Figure 3: QSFP Air Temperatures for Cases 1 to 3 of Figure 2 – Air Inlet 40°C at 3000m.



Although the front-to-back flow of case 1 provides the best average air temperature at the modules, side-to-side flow is often used because many applications require filtering, in particular, telecommunications equipment per Section 4.5.4 R4-87 of Reference 6. The provision of filters at the faceplate poses design and service issues particularly when the optical modules are fully connected with optical fibre.

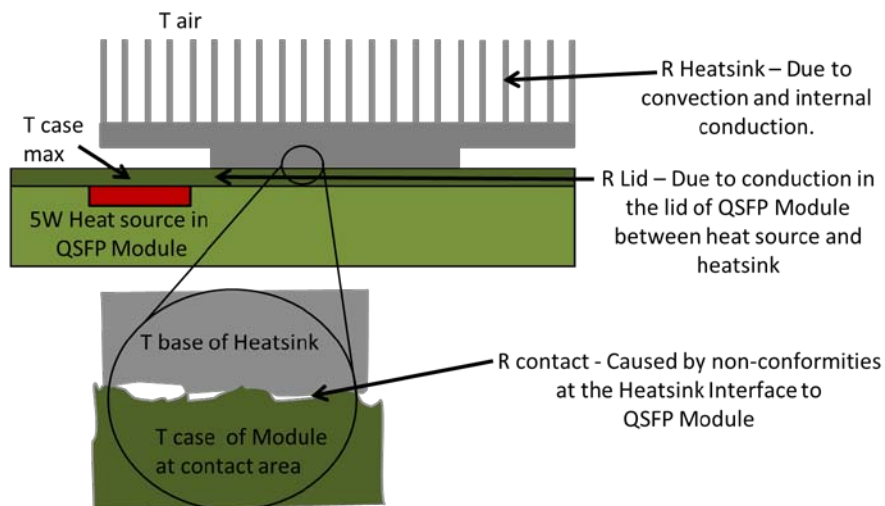
The blade flow considerations become more critical when telecommunications environments are considered. There, design inlet conditions increase to 50°C for large switches and 55°C for smaller ones, Reference 6 section 5.1. With the above design, case 3 flow would have 68.2°C air at location 2 with 55°C inlet at 1800m elevation conditions vs. 55.3°C above. In comparison, case 1 flow has average QSFP air temperature at 57.6°C vs the 43.1°C above. At a 70°C maximum case temperature, case 1 provides almost 2.2 times the ΔT to drive heat removal through convection. At that, only 12°C remains of the original 26.9°C to drive heat from the module cover to the air. Note that the corresponding European telecommunications standard, Reference 7, is less restrictive with a maximum design inlet of 45°C but this still reduces the available temperature difference by ~19% to 21.9°C.

Case 2 flow pattern was selected for detailed modeling because of its wide applicability and good air temperature performance. The ΔT data in Figure 3 gives the overall driving temperature, available for 40°C inlet air 3000m elevation conditions. With 25°C available for the highest temperature plug, the design goal for the modeling was to use 20°C to cover the resistance between the module and air. This includes convective resistances from module and heatsink surfaces and internal conduction resistances.

When a heatsink is used, heat is conducted from the source through the module case, across the contact resistance of the interface between the module and heatsink base and then to the outer surface of the heatsink. Heat is then convected to the air. Common practice is to combine the internal conduction resistance of the heatsink with the convective resistance into a single heatsink resistance, R_{heatsink} , that includes the resistance due to spreading heat from a smaller source throughout the base and fins and the airspeed dependent convective resistance. The contact resistance, R_{contact} , of the interface between the module case and base of the heatsink is difficult to estimate as will be discussed later and is therefore treated separately. In addition, the heat must conduct through the lid of the module to a location under the heatsink contact area; this is included as R_{lid} . Using this, the case temperature is calculated from

$$T_{\text{case max}} = T_{\text{air}} + R_{\text{heatsink}} \times q + R_{\text{contact}} \times q + R_{\text{lid}} \times q \quad (4)$$

Figure 4: Module Thermal Resistance Through Heatsink



For the design example, $T_{\text{case max}} - T_{\text{air}} = 20^{\circ}\text{C}$, so that at 5W per QSFP, the allowable total thermal resistance, $R_{\text{lid}} + R_{\text{contact}} + R_{\text{heatsink}} = R_{\text{total}}$, is 4°C/W . For design example purposes, it was assumed that half the resistance would be allocated to the interface resistance, and the remainder to R_{heatsink} .

4 Detailed Thermal Modeling

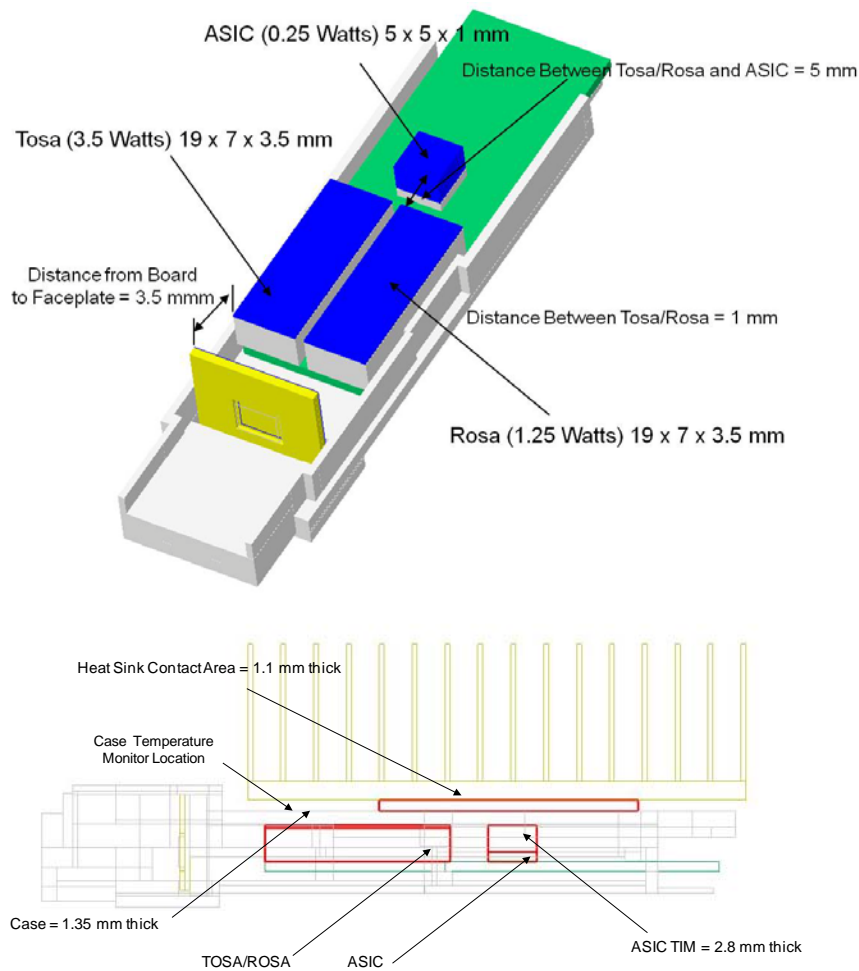
The thermal modeling work is described in Google Thermal Group presentations, References 8-10, and an extended abstract Reference 11. The results and model descriptions are summarized here.

The models of the switch blade consist of three components:

1. The QSFP module
2. The QSFP cage and heatsink
3. The switch blade

The Flotherm[®] model of the QSFP is shown in Figure 5 and shows the three active components inside the module. The model was originally donated by Avago Technologies but was modified by Brocade and Electronic Cooling Solutions. The most critical component of the QSFP module is the TOSA or Transmitter Optical Sub-Assembly that houses the 4 lasers. The 4 lasers are the critical components of the TOSA because they need to operate at a low temperature for high reliability. Since the internal design of the QSFP varies from vendor to vendor, the main specification for optical modules is a 70°C case temperature on the power dissipation surface in contact with the heatsink.

Figure 5: The QSFP Flotherm[®] Model



The Flotherm[®] models for the cages were donated by Molex Incorporated. The cages are made of tin and form the port or enclosure where the QSFP modules are inserted. Each 1X4 cage holds four QSFP modules and holds an independent heatsinks for each QSFP. The riding heatsinks are attached to the cage with a heatsink clip. According to the QSFP standard, Ref. 12, the heatsink clip exerts a 5 newton force on the module. When the module is inserted, the heatsink slides over the top of the module with a contact interface

317mm² (0.49 sq in) and 12 mm by 26.45 mm or about one third of the top surface area of the QSFP inside the cage. Increasing the heatsink interface area may be a way to reduce the contact resistance between the module and the heatsink.

The switch blade model is essentially the same as Case 2 from Figure 1. From the preliminary analysis, it was determined that any successful design would prevent mixing of hot air from over the ASIC complex from mixing with the air cooling the plugs. Accordingly, other components were removed from the blade so that the only heat sources in the blade were the QSFP modules. To simplify the Flotherm[®] model, the channel was reduced to the dimensions shown in Figures 6 and 7. The airflow through the channel was usually 17m³/h (10 cubic feet per minute, cfm) but the airflow was varied in some cases. Outside of the QSFP cages, average airspeed at 17m³/h would be 1.3m/s (254LFM) at cage location with no heatsinks this increases to 1.5m/s (297LFM). The natural convection regime existing outside the faceplate was included in the analysis as were the details of the air flow between the module and cage, and the cage and circuit board.

Figure 6: The Simplified Switch Blade Model Showing Case 3 Baffles and Axial Location of QSFP's.

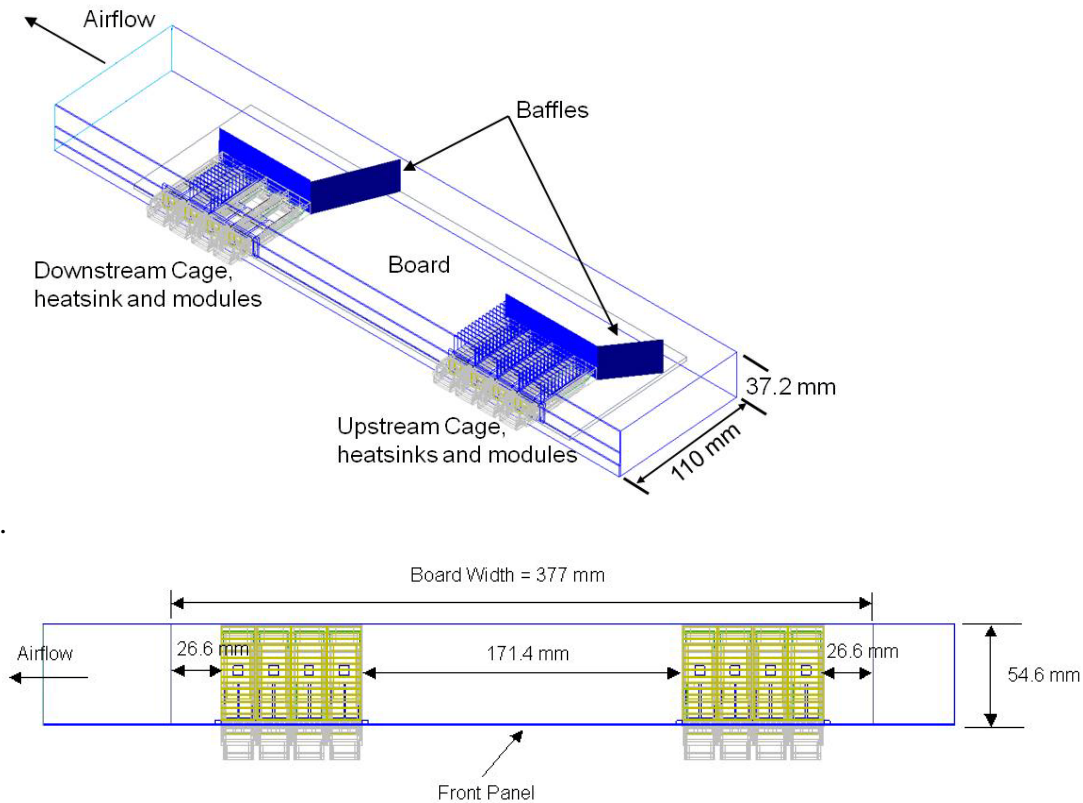
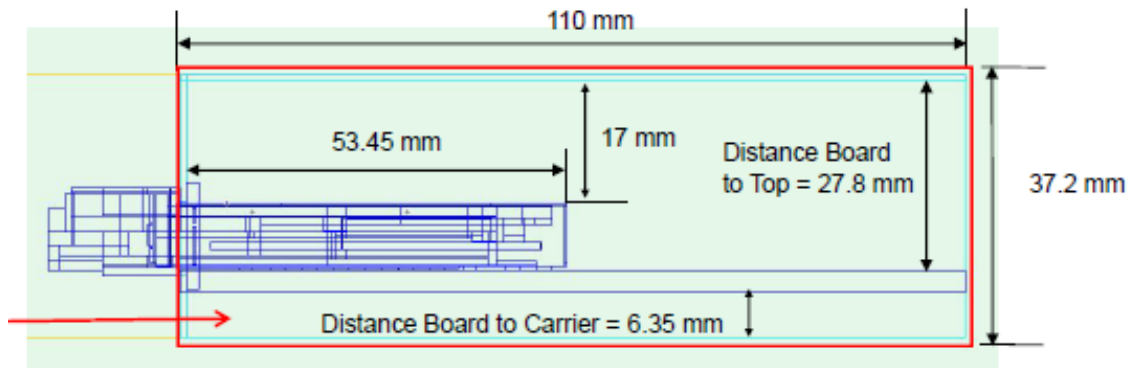


Figure 7: Model Cross-section Through QSFP – Open Area Total =3179mm² Above Board=2481mm² Used in Cases 1 through 7.



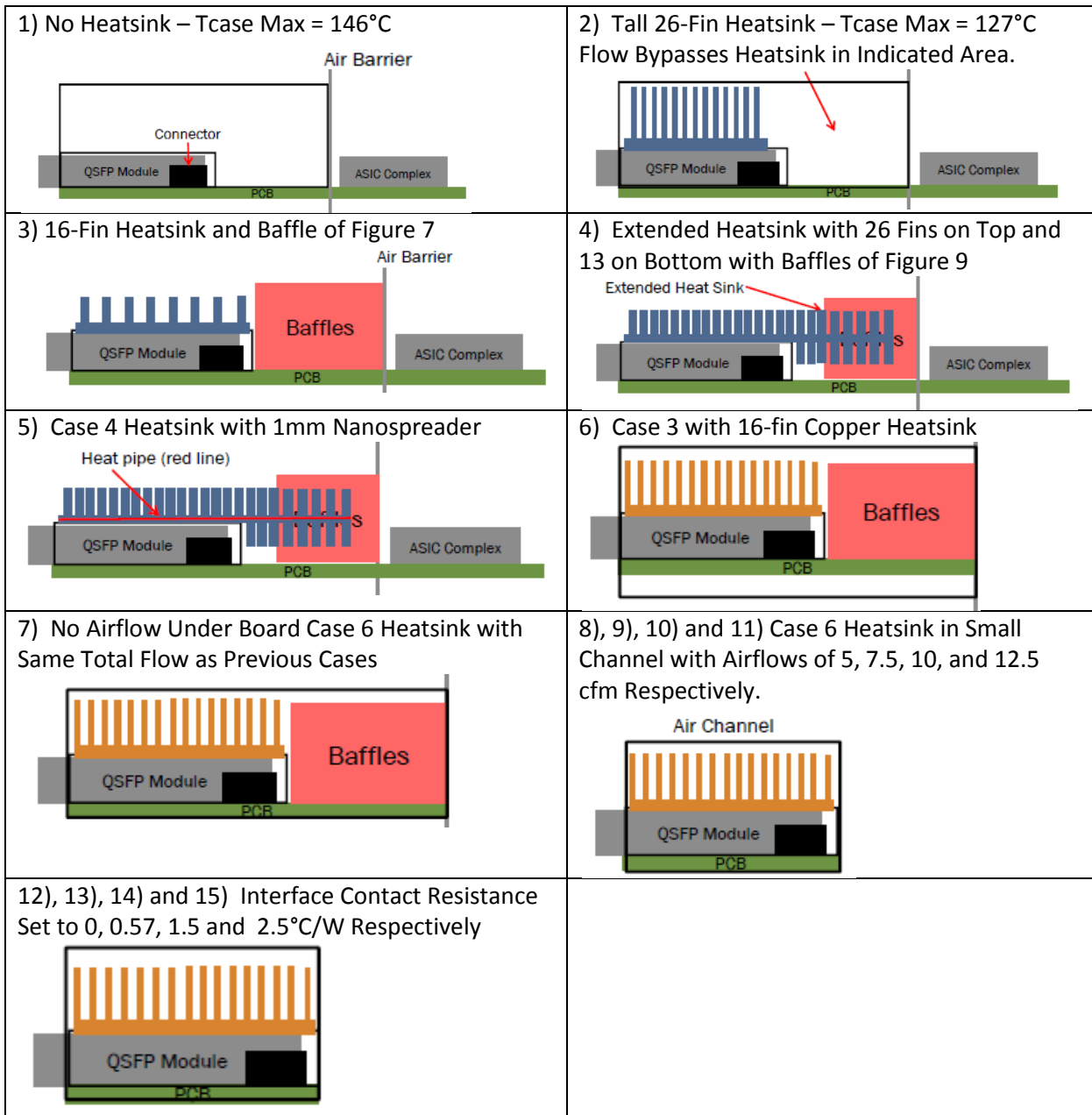
Since the switch and optics are cooled by air, the initial conditions of the air are an important parameter of the simulation. The inlet air temperature was set to 40°C and the elevation was 3000m or about 10,000ft. The temperature difference (ΔT) between the air and maximum allowable case temperature was 30°C for the first module. A baseline contact resistance equivalent to a 0.02mm air gap giving ~2.1 C/W or half the total design resistance was applied for the analysis and will cause ~10°C temperature delta between the module case and the heatsink if all 5W is dissipated through the top of the module.

Fifteen simulation cases were defined to explore changes in flow boundaries, heatsinking and contact resistance. The results are summarized in Table 2 and the cases are outlined in Figure 8. In the analyses, heat transfer and flow were only solved for the region left of the air barrier. All heatsinks modeled have 13mm tall x 0.55mm thick fins on a 2mm thick base on the top side, and the extended base is 4mm thick with bottom fins 9mm tall. Case 1 results are included to show that heatsinks are necessary on the QSFP's. Case 2 results show that without baffling, heatsinks that don't extend past the end of the QSFP cage do not provide sufficient improvement because much of the air bypasses both the modules and heatsinks without removing heat as shown in Figure 9 a) by the cool air extending to the downstream end of the blade.

Table 2: Summary of CFD Model Results

| Case No. | Flow [m ³ /h] (<i>cfm</i>) | Contact Resistance R_{contact} [°C/W] | Channel Flow Area A_{channel} [mm ²] | Overall Pressure Loss [Pa] (<i>in H₂O</i>) | T _{case} Max Upstream [°C] | T _{case} Max Downstream [°C] |
|----------|---|--|---|--|--|--|
| 1 | 17 (10) | 2.09 | 3767 | | | 146 |
| 2 | 17 (10) | 2.09 | 3767 | | | 129 |
| 3 | 17 (10) | 2.09 | 3767 | 16 (0.066) | | 78.3 |
| 4 | 17 (10) | 2.09 | 3767 | 8.7 (0.035) | 68.4 | 74.8 |
| 5 | 17 (10) | 2.09 | 3767 | 8.7 (0.035) | 64.5 | 70.9 |
| 6 | 17 (10) | 2.09 | 3767 | 16 (0.066) | | 77.0 |
| 7 | 17 (10) | 2.09 | 3058 | 42 (0.17) | | 70.4 |
| 8 | 8.5 (5) | 2.09 | 1518 | 12 (0.094) | | 81.6 |
| 9 | 12.7 (7.5) | 2.09 | 1518 | 40 (0.16) | | 74.2 |
| 10 | 17 (10) | 2.09 | 1518 | 62 (0.25) | | 70.6 |
| 11 | 21 (12.5) | 2.09 | 1518 | 85 (0.34) | | 68.2 |
| 12 | 21 (12.5) | 0.0 | 1518 | | | 60.5 |
| 13 | 21 (12.5) | 0.57 | 1518 | | | 64.8 |
| 14 | 21 (12.5) | 1.5 | 1518 | | | 68.0 |
| 15 | 21 (12.5) | 2.5 | 1518 | | | 70.0 |

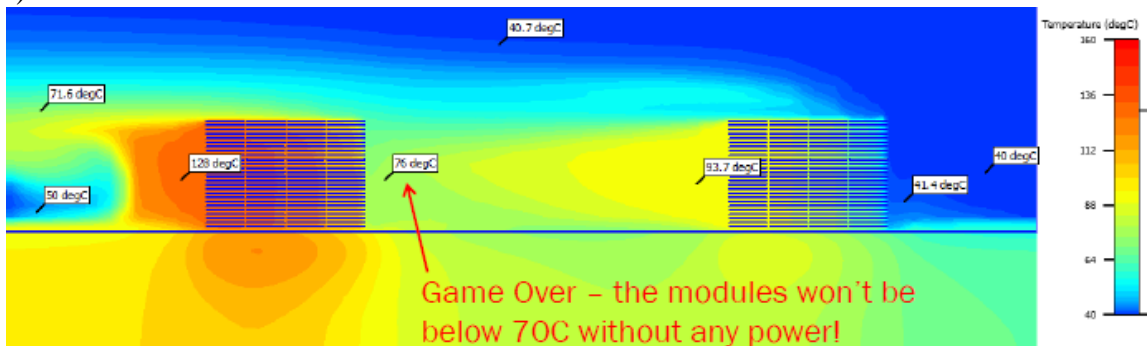
Figure 8: Detailed Thermal Modeling Cases



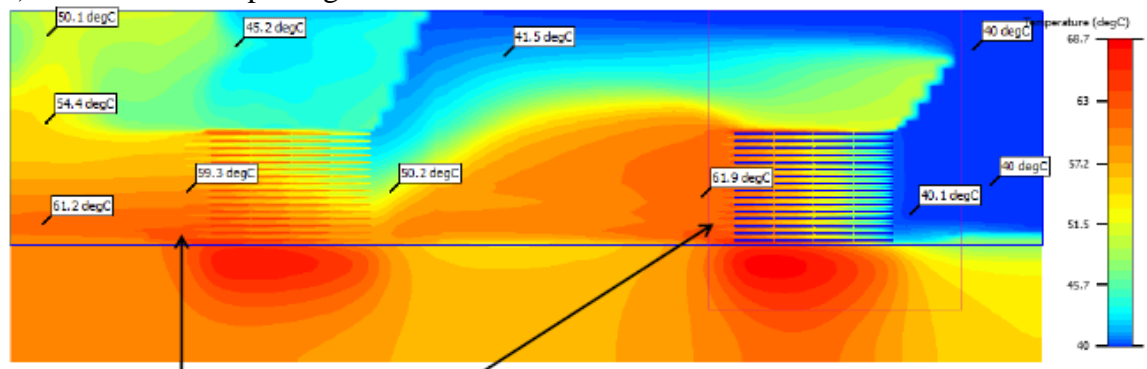
The airflow control provided by the baffles of Figure 10 was added for case 3. This resulted in a major reduction in maximum case temperature to 78°C as indicated in Table 2. This improvement is primarily due to the >20C reduction in air arriving at the downstream modules in Figure 9b relative to 9a. For the more dense layouts of Ref. 3, baffled or ducted designs that introduce adequate cool air to the downstream modules could also be effective. However, providing this cool air gets progressively more difficult since the channels required for cool air will likely be gained at the expense of fin area.

Figure 9: Air Temperatures With and Without Baffles at Plane at Mid-height of Heatsink Fins: Thermal Modeling Cases, 2, 3, 4 and 5.

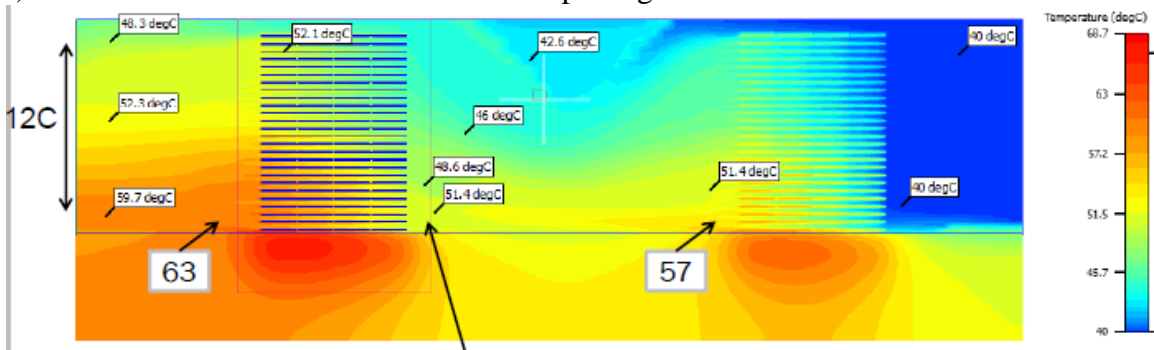
a) Case 2: No Baffles



b) Case 3: Baffles per Figure 10



c) Case 4: Extended Heatsink with Baffles per Figure 11



d) Case 5: Extended Heatsink with Nanospreader and Baffles per Figure 11

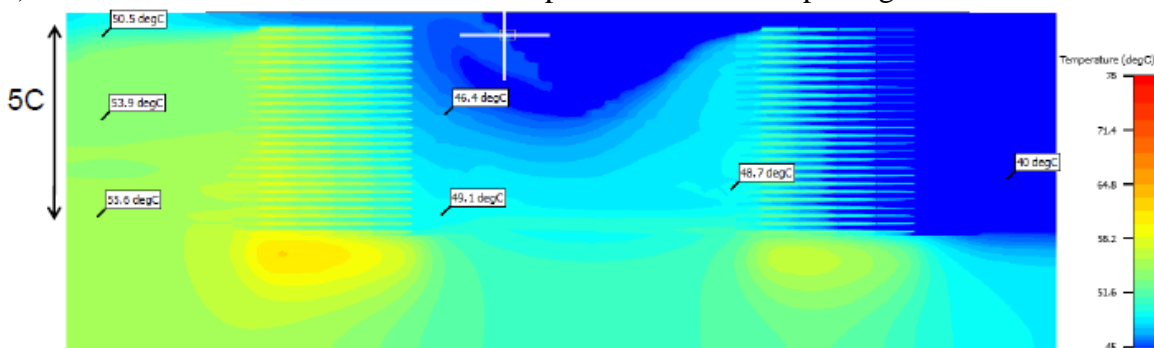
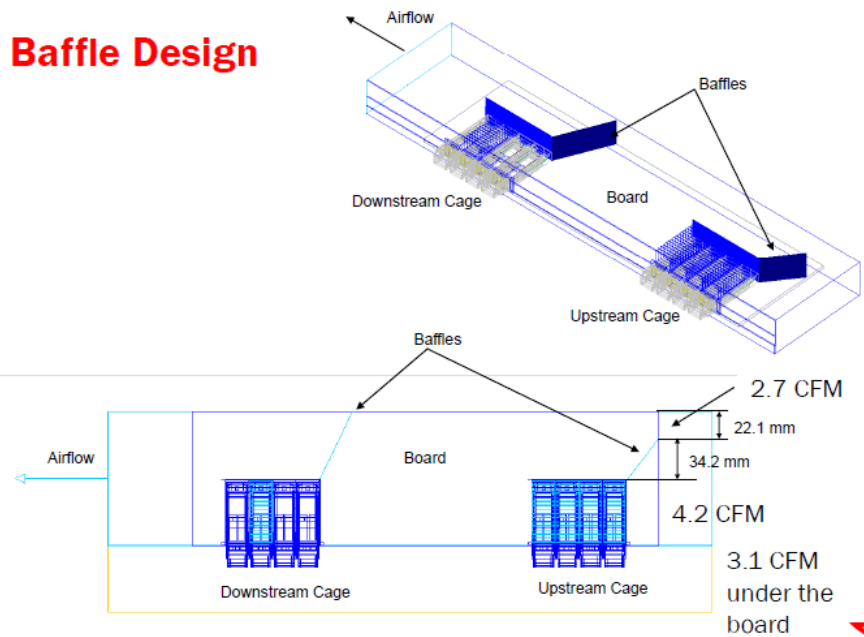


Figure 10: Baffle Design for Cases 3, 6 and 7



Cases 4 and 5 explore the idea of extending heatsinks toward the back of the channel. The heatsink surface area is increased by a factor of ~ 2.2 . In addition, the V shaped baffle and 2 angled baffles shown in Figure 11 were introduced to promote mixing of the flow between the upstream QSFP's and the downstream QSFP's. Case 4 results reduced the maximum case temperature to 75°C but it was apparent from inspection of the results that further improvements could be obtained by reducing the $\sim 10^{\circ}\text{C}$ temperature variation across the base of the heatsink, see Figure 12a. Case 5 is the same as case 4 but has a 1mm thick vapor chamber or “nanospreader”, modeled as a 5000W/m-K conductive solid, embedded in the base of the heatsink. Temperature variation across the heatsink base of Figure 12b was then $< 2^{\circ}\text{C}$ and maximum case temperature was reduced to 70.9°C . This improvement is also shown in relative uniformity of air temperatures in Figure 9d relative to 9c.

In Case 6, the material of the heatsink is changed from the aluminum of Case 3 to copper. This resulted in slightly more than 1°C reduction in maximum case temperature. Examination of the results of Cases 3 through 6 showed that of the $17\text{m}^3/\text{h}$ forced through the channel, $\sim 5.3\text{m}^3/\text{h}$ flowed underneath the board and was ineffective in removing heat from the QSFP's. For Case 7 the channel under the circuit board (PCB) was blocked for airflow. Raising the total flow on the top for the PCB to $17\text{m}^3/\text{h}$ increased the flow through the heatsinks of the upstream and downstream banks of QSFP's as shown in Figure 13b relative to 13a. This also reduced the maximum QSFP case temperature to 70.4°C .

Figure 11: Baffles Used with Extended Heatsinks - Cases 4 and 5

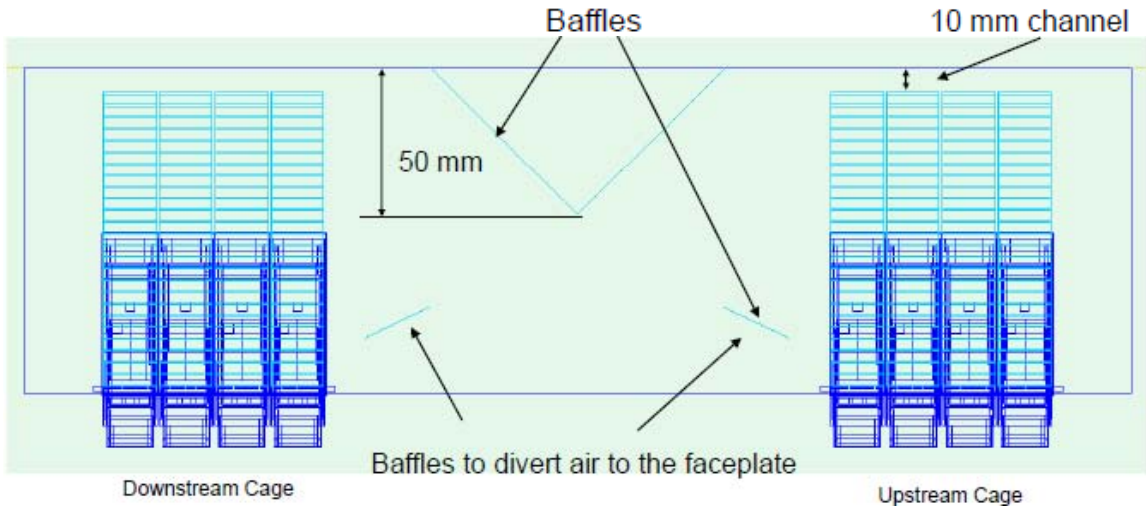


Figure 12: Extended Heatsink Temperatures for Downstream QSFP's – a) Aluminum Heatsink per Case 4 – b) Al Heatsink with Embedded Nanosreader. Sections are through second QSFP from the left.

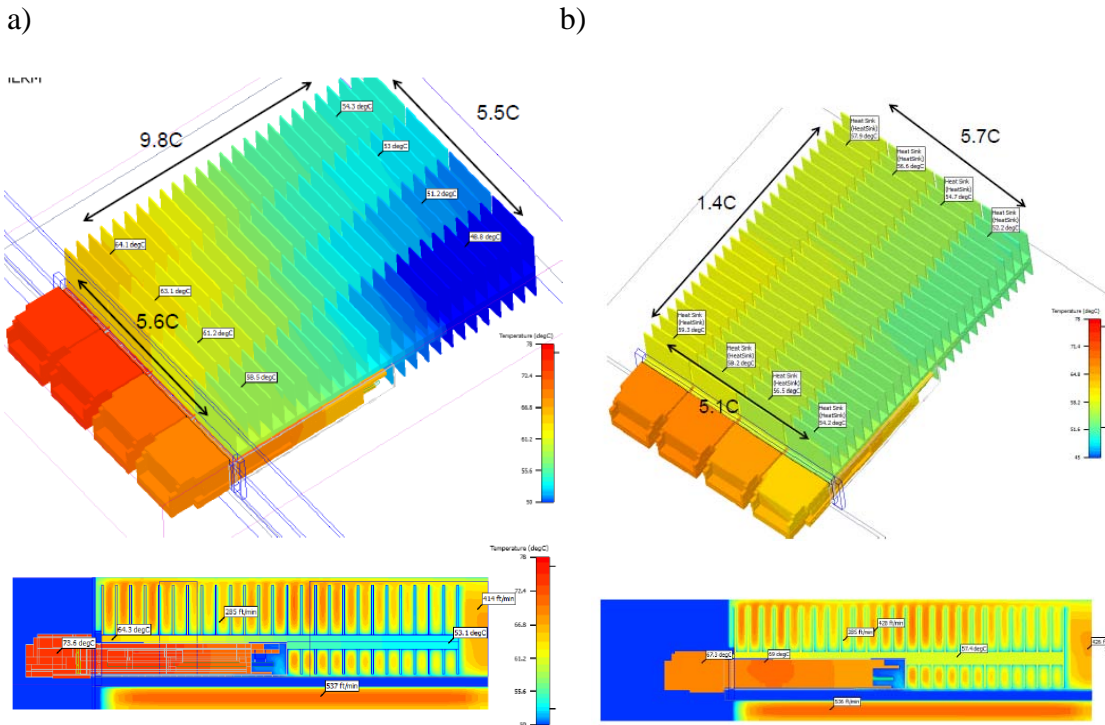
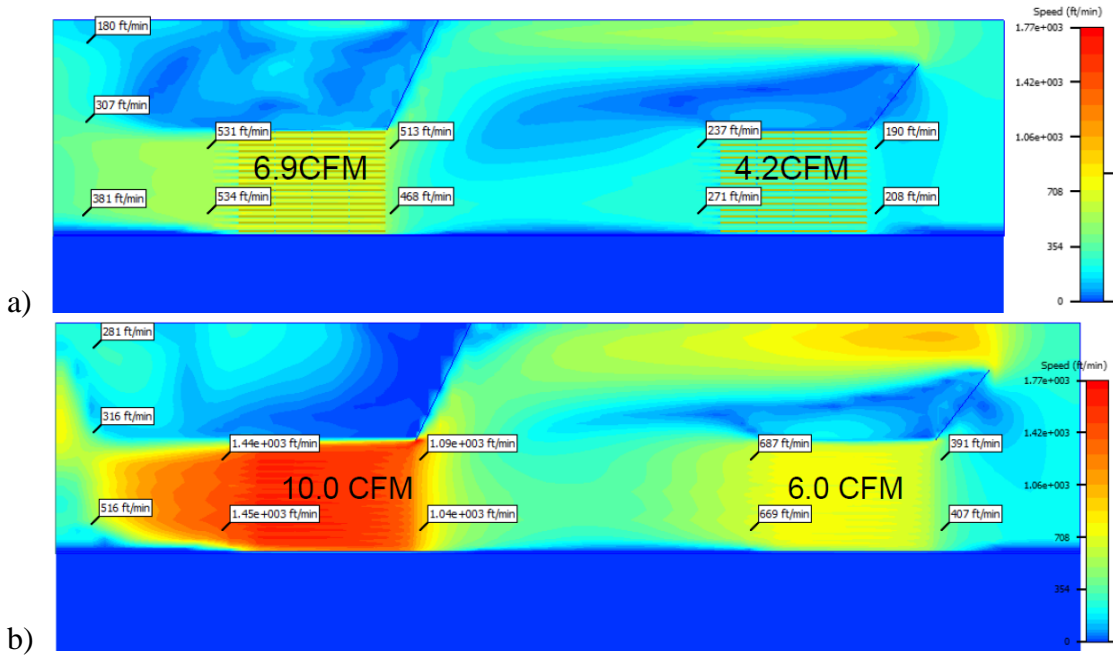


Figure 13: Air Flow for Cases 3, and 6 with Flow Under the Board a), and Case 7 with Flow Under the Board Blocked b). Section is at Mid-height of the Fins



For the remaining cases, 8 through 15, the flow channel was reduced to cover the area from the faceplate to just past the end of the un-extended heatsink as shown in Figure 14. In cases 8 through 11 air flow is varied from 8.5 to 21m³/h in 4.2m³/h steps. By examining the data from Table 2 shown in Figure 15, we find that increasing flow provides diminishing returns e.g. from 8.5 to 12.7m³/h reduces the max case temperature by 7.4°C while the increase from 17 to 21 m³/h only reduces case temperature by 2.4°C. Conversely, the cost in terms of pressure drop or fan power, increases by 16.9Pa from 8.5 to 12.7m³/h but by 22.4Pa from 17 to 21m³/h.

It is also worth comparing the pressure losses at 17m³/h; the small channel of Case 10 has 62Pa drop while the large channel with baffles of Case 7 has 42Pa at 17m³/h. The maximum case temperature is essentially the same in both cases. Here extra board area is gained for the ASIC complex at the cost of extra pressure drop that must be balanced against the pressure drop over the ASIC complex or the air will simply bypass the QSFP's as it did in Case 2. The increase in pressure drop may be significant in some systems, but lies within the capacity of high performance fans.

Figure 14: Small Flow Channel for Cases 8 through 15, Section through QSFP and Plan View

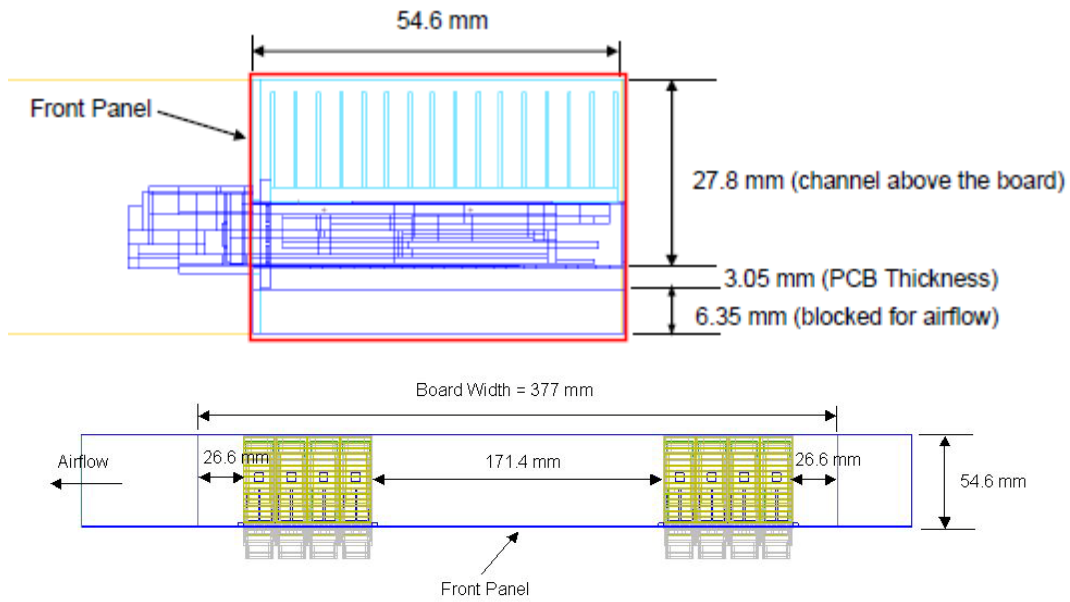
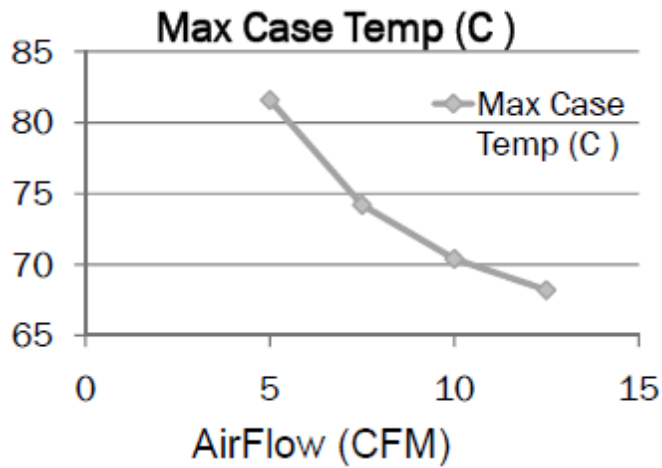


Figure 15: Effects of Changing Airflow in the Small Channel Cases 8 through 11



The remaining analyses for Cases 12 through 15 explore the effect of reducing the applied interface contact resistance with channel airflow set to 21m³/h with the results shown in Figure 16 as well as Table 2. If the interface resistance is set to 0°C/W per Case 12, the maximum module case temperature is reduced by ~7.7°C relative to Case 11 which had the 2.1°C/W that was applied for Cases 1 through 11.

Figure 16: Effects of Changing the Applied Contact Resistance in the Small Channel Cases 12 through 15

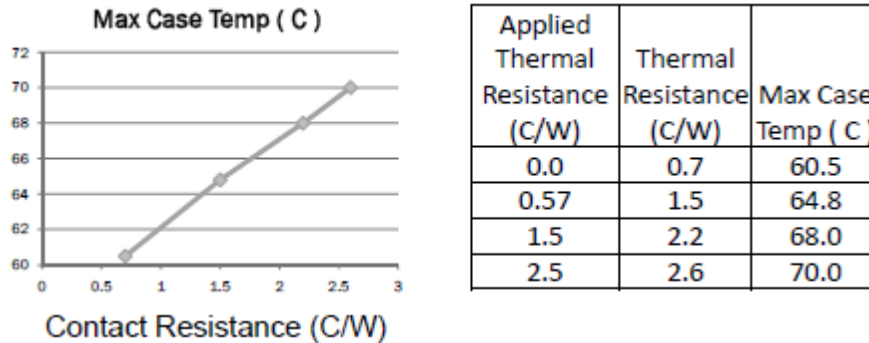


Figure 16 also gives an “effective thermal resistance” calculated by taking the temperature difference between the maximum case temperature on the module and the temperature at the centre of the heatsink baseplate and dividing by the 5W heat dissipation of the module. This effective resistance is larger than the applied contact resistance due to R_{lid} and the need to conduct heat from the maximum case temperature location to the heatsink.

The design goal for the heatsinking was to remove ~90% of the heat through the heatsink. Typically the results showed it to be more on the order of 80% indicating the heat removal outside the faceplate and through board to the secondary side was significant.

5 Discussion of Contact Resistance Issues

A detailed discussion of thermal contact resistance is beyond the scope of this paper; however, the physics and modeling of thermal resistance at various types of joints that occur in microelectronics problems are summarized in Reference 13. Here it is seen that the thermal resistance of contact regions depends on the force holding the 2 faces together and both the solid, usually metals, and fluid materials present at the interface. In addition, it is a complex function of the material and the dimensional properties of the solids comprising both faces, in particular, their thermal conductivity, micro-hardness, surface roughness (including geometric characterization of the asperities), and the size and locations of area(s) where contact occurs. Rearranging and dimensionalizing the correlation equation for dimensionless contact conductance given in Reference 13 gives, for a conforming contact area.

$$R_{\text{contact}} = (0.69 / (k_s \times A_{\text{contact}}) \times (\sigma / m) \times (H_c / P_i)^{0.95} \text{ [}^\circ\text{C / W]} \quad 5)$$

$$= 0.69 \times (1/(\text{harmonic mean of the surface conductivities} \times \text{contact area})) \times (\text{effective RMS roughness of the two surfaces}/\text{effective mean asperity slope of the interface}) \times (\text{contact microhardness}/\text{nominal joint contact pressure})^{0.95}$$

This shows that resistance varies in direct proportion to the surface roughness, and in inverse proportion to the surface area of contact. In addition the resistance varies with the 0.95 power of surface hardness and with the -0.95 power of contact pressure.

Owing to the need to keep insertion and removal forces low for pluggable modules, forces holding surfaces together are generally insufficient to cause macro-level deformation of the metals that would ensure that the entire nominal contact area has conforming contact. As a result, two regions need to be considered, 1) regions of surface-to-surface contact where air fills the surface asperities in the contact region, i.e. the micro-contact regions and 2) macro-scale air-filled regions where the nominal gaps between the surfaces are greater than the heights of the surface asperities. In the first region, conduction through the air parallels the surface-to-surface contact heat paths from the source to the heatsink. The macro regions are essentially air gaps that form parallel heat paths to the micro-contact regions.

In practice, the interface geometry is controlled by the fabrication process including casting, machining and/or other fabrication operations. The appropriate geometric parameters controlling planes that slide over one another are surface roughness, and flatness. A review of several MSA's, References 12, 14-16, shows that although flatness is generally specified, surface roughness is not; see Table 3. Where it is specified for CFP's, it is ~25x the surface roughness. If the surface roughness is the same as for CFP's, the specified module flatness of 0.075mm for XFP's, QSFP's or SFP's is ~12x the surface roughness. In either case, this is large enough that without macro-scale deformation of the interface surfaces, significant areas without contact can be expected mixed with small contact areas. Note also that for a QSFP, the contact resistance of 2.1°C/W applied in the baseline analyses is equivalent to that of a uniform air gap between the module and heatsink of 0.02mm (0.0007in) or ~1/4 of the specified flatness. Clearly the MSA specifications are inadequate for thermal designers to fully evaluate contact resistance.

Table 3: MSA Plug Thermal Interface Specifications

| Plug Type | Temperature Monitor Point Location Specified | Heatsink Contact Area Size (l x w) [mm] | Heatsink Interface Load [N] | Interface Surface Roughness [$\mu\text{m Ra}$] | Interface Surface Flatness ¹ [mm] | Other Interface Specification |
|-----------|--|---|-----------------------------|--|--|--|
| SFP+ | no | 26.5 x 9.0 | | | 0.075 | |
| QSFP | no | 26.45 x 12 | 5 | | 0.075 0.025 hs | |
| XFP | no | 44 x 12.8 | >5 | | 0.075 0.025 hs | |
| CFP | no | 110 x 70.0 | 10-15 | 6.3 | 0.15 0.15 hs | Shows an interface material on the heatsink. |

Notes: 1) Where possible, values for both the device and heatsink are given with the heatsink value followed by the letters “hs”.

Depending on how the out-of flatness is realized, other thermal losses can arise. Consider a typical geometry for heatsink loading where springs are located on either side of a high point in the contact area. Further assume that the module surface is convex with its high point at the centre when viewed from the side and that the heatsink interface surface is perfectly flat. If the maximum height point of the module lies between the springs, a moment balance on the heatsink shows that it will tilt and that contact can generally be expected at one end and at (or near) the top of the convex area on the module. As a result, the other end has a larger gap than the amount of out-of-flatness. If instead, the module surface was concave, an air gap would occur in the middle of the heatsink. Depending on the location of the heat sources, temperature loss due to conduction in the module case and heatsink base will be more or less significant in the overall temperature loss between case hot spot.

6 Recommendations

The results and discussion above have addressed blade layout, airflow, heatsink design and contact resistance issues that impact the success of a thermal design with optical modules at the faceplate. Recommendations divided into several areas: for system designers, layout considerations, thermal design, and module design and common specifications.

Back-to-front airflow is ineffective for cooling optical plugs and is not recommended for cooling blades with ASIC complexes or other hot devices on the board. With overall increasing board powers, the difference between air and allowable maximum case temperatures with back-to-front flow is not sufficient to cool the plugs.

The preliminary analyses showed that the system providing the lowest temperature air for cooling devices at the faceplate has airflow through the faceplate from front to back. This

design is not suitable for telecommunications applications due to filtering considerations and a side-to-side (or equivalent bottom-to-top) airflow geometries are recommended in this case.

Physical layout of the blade should separate the high-power high-temperature devices from the optical plugs with the high power devices to the rear. This permits cool air to be channeled separately over the optical plugs.

Any power dissipated in the plug channel will increase the incoming air temperature for downstream devices so this should be minimized as much as possible.

Separate groups of faceplate plugs to encourage mixing of air for downstream plugs thereby reducing locally hot air from negatively impacting performance. Any increase in the number of plugs located adjacent to each other increases the difficulty of side-to-side cooling as the possibility of introducing cool air for mixing is reduced.

Use baffles when required to separate hot air from the ASIC complex from the generally cooler required for the plugs. Baffles may also be necessary for mixing in cool air between groups of faceplate optical modules to reduce incoming air temperatures for downstream modules.

Heatsinks are required for higher power optical modules. Extending the heatsinks rearwards allows more air to be heated and can make downstream air temperatures more uniform. Thermal designers must ensure that the heatsink base has sufficient conductivity for this. This also reduces losses arising from the module heat source not coinciding with the centre of the heatsink.

Increasing the local airflow improves performance by reducing the temperature rise of the air and increasing the heat transfer coefficient at module and heatsink surfaces. If this is to be achieved by channeling flow and blocking flow to other areas, the trade-off is that higher pressure losses may cause air to bypass the faceplate modules. Designers must exercise caution that this doesn't occur and that expected pressure losses are consistent with pressure provided by the fans. An upgrade of platform fans may be required and increased fan performance usually means more noise.

Although the analyses above showed that systems with 40°C incoming air could be made to work with a relatively large contact resistance, the same 2.1°C/W resistance on a 5W QSFP would leave only 5°C between incoming air at 55°C telecom shelf level ambient requirement and the 70°C case temperature maximum. For cost-effective air-cooling solutions, contact resistance between modules and heatsinks needs to be minimized. In addition, conduction losses in module lids should be minimized.

Module suppliers need to supply thermal designers with sufficient information to allow them to estimate contact resistance. This includes:

- Location of the monitor point for case temperature which may or may not coincide with the hot-spot.

- Maximum short and long term operating temperatures at that location.
- Sufficient detail of the module to allow calculation of the module surface temperatures in detail analyses. This includes location, size and wattage of heat sources connected to the top surface, the top surface material, the thickness distribution of the top surface in sufficient detail to allow calculation of heat spreading and, details of heat sources not connected to the top surface.
- Ideally this information would allow calculation of heat transferred out or in through other faces of the module to the PCB or air.

The following enabling improvements should be discussed to improve thermal design of the MSA and/or individual suppliers:

- Reduce contact resistance by increasing the force holding the heatsink against the module. Trade-offs with connectors and insertion/removal forces should be evaluated.
- Improve surface finish specification from 6.3 μ m Ra to 0.8 μ m Ra or better. The quoted value should be obtainable through standard machining operations or high quality casting.
- Reduce flatness tolerance to 0.001mm/mm (0.001in/in) to a maximum of 0.075mm or even less.

7 Summary

The power and therefore heat dissipation of optical pluggable modules is expected to increase at the same time as plugs are reducing in size and increasing in number per blade. As a result their heat flux is expected to increase from 2 to 5 times over the next 2 to 4 years. Current thermal designs are approaching their functional limits.

With the aid of a detailed conjugate heat transfer model of a QSFP optical plug module, a series of analyses have been conducted on a simplified switch blade platform. On this basis, recommendations for cooling the high power modules expected in the near future were produced. The analyses identified that front to back and channeled side-to-side airflow can provide the lowest temperature air for cooling plug modules. Increasing air flow can deliver reduced air temperature for downstream components but has lesser benefit for upstream components. By utilizing PCB layouts that separate high dissipation and hotter components from the plugs, air temperatures arriving at plugs are kept low as are board temperatures in the region of the plug.

Air flow in the regions of plugs can be improved by including baffles. Baffles can be used to separate plug air from hot air over the ASIC's, increase the speed of air at the plugs, and improve flow mixing to reduce hot air pockets or air streams in the regions of the plugs. Baffles can also be used to increase flow locally in the faceplate region and thereby increase the performance of heatsinks. Layouts with increasing numbers of optical plugs adjacent to each other are progressively more difficult to cool since the heat added to the air

increases its temperature at downstream plugs and the room to introduce cool air via baffles or ducting is reduced.

Increasing the size of heatsinks on devices usually makes major improvements to heat removal capacity. Heatsinks for current plugs have planform areas similar to the plug, however the expected increase in heat flux means that thermal resistance must be reduced to achieve plug temperature limits. The convective component can be reduced by increasing the surface area and/or the air speed. For plugs in gang cases increasing the heatsink area necessitates extensions rearward from the faceplate, and can bring large temperature gradients from front to back in the heatsink. It is likely that heatpipes or advanced highly conductive materials must be embedded in the base to reduce this temperature variation, i.e the conductive component to an acceptable level.

The requirement to remove plugs means that only small forces are available to reduce interface resistance between heatsink and plug. However, contact resistance literature shows that it can be easily reduced by reducing surface roughness as well as increasing the contact pressure. Separate from contact areas, gaps between the mating surfaces form a parallel thermal resistance. Since many MSA's do not even mention surface roughness of the plug or heatsink they require review to ensure sufficiently tight tolerances are included. In addition, flatness tolerances are larger than can be achieved through standard fabrication techniques and tighter tolerances would reduce gap heights between the mating faces.

The analyses showed that through the combined actions of separating air streams, increasing flow velocity at the plugs, keeping high dissipation components located away from faceplate area, increasing heatsink surface area, and decreasing the interface resistance between the heatsink and plug it is possible that case temperature limits for plugs can be met at the expected higher powers when ambient is 40°C. For telecommunications applications where 55°C is the maximum ambient this will be much more difficult unless contact resistances are significantly reduced. The analyses did not include heat dissipation in the board near the plugs so they are optimistic in that respect.

8 References

- 1) Joiner, Steve, Finisar: Thermal Management at the Face Plate, Optical Internetworking Forum Document No. OIF.2010.183.01, May 12, 2010.
- 2) Kipp, Scott, Brocade: Module Power Density, Optical Internetworking Forum Document No. OIF.2010.111.00, February 9, 2010.
- 3) CFP MSA Companies: Next Gen PMD CFP MSA Baseline Specifications, www.cfp-msa.org, March 16, 2012.
- 4) Serth, Robert: Process Heat Transfer: Principles and Applications, Academic Press, 2007.
- 5) Kipp, Scott, Brocade: Modeling 800 GBPS Blades – Part 1, Google Thermal Modeling Group, (<http://groups.google.com/group/thermal-modeling>), June 30, 2011.
- 6) Telcordia, NEBS™ Requirements: Physical Protection, Telcordia Technologies Generic Requirements, GR-63-CORE, Issue 3, March 2006.

- 7) ETSI European Standard (Telecommunications Series), EN 300 019-1-3 V2.3.2 Environmental Engineering (EE); Environmental conditions and environmental tests for telecommunications equipment; Part 1-3: Classification of environmental conditions; Stationary use at weather protected locations, (2009-11)
- 8) Kipp, Scott, Brocade: QSFP Modeling Results – 3, Google Thermal Modeling Group, (<http://groups.google.com/group/thermal-modeling>), July 25, 2011
- 9) Kipp, Scott, Brocade: QSFP Modeling Results – 4, Google Thermal Modeling Group, (<http://groups.google.com/group/thermal-modeling>), August 25, 2011
- 10) Kipp, Scott, Brocade: QSFP Modeling Results – 5, Google Thermal Modeling Group, (<http://groups.google.com/group/thermal-modeling>), September 21, 2011
- 11) Kipp, Scott and Romero, Adriana, Cooling 8x100GbE Switch Blades with High Power Optical Modules, Extended Abstract Submitted to Semitherm Conference
- 12) SFF Committee, SFF-8436 Specification for QSFP+ 10 Gbs 4x Pluggable Transceiver Rev. 4.1, SFF Committee, August 24, 2011, <ftp://ftp.seagate/sff/>.
- 13) Jovanovitch, M. Michael: Four Decades of Research on Thermal Contact, Gap and Joint Resistance in Microelectronics, IEEE Transactions on Components and Packaging Technologies, Vol. 28, No. 2, June 2005
- 14) SFF Committee, SFF-8432 Specification for Improved Pluggable Formfactor, Rev. 5.0, July 16, 2007, <ftp://ftp.seagate/sff/>.
- 15) Ghiasi, Ali, Technical Editor, SFF Committee, INF-8077i, 10 Gigabit Small Form Pluggable Module, XFP Rev 4.5, Copyright XFP Promoters August 31, 2005, <ftp://ftp.seagate/sff/>.
- 16) Traverso, Matt Editor, CFP MSA Hardware Specification, Rev. 1.4, June 7, 2010

Appendix A: List of Contributors

| | | |
|----------------|------------------------------|--|
| Terence Graham | Ciena | tgraham@ciena.com |
| Scott Kipp | Brocade | skipp@brocade.com |
| Adriana Romero | Electronic Cooling Solutions | aromero@ecooling.com |

Appendix B: About the OIF

Launched in 1998, the OIF is the first industry group to unite representatives from data and optical networking disciplines, including many of the world's leading carriers, component manufacturers and system vendors. The OIF promotes the development and deployment of interoperable networking solutions and services through the creation of Implementation Agreements (IAs) for optical, interconnect, network processing, component and networking systems technologies. The OIF actively supports and extends the work of standards bodies and industry forums with the goal of promoting worldwide compatibility of optical internetworking products. Information on the OIF can be found at <http://www.oiforum.com>.