

QSFP-DD MSA

QSFP-DD: Enabling 15 Watt Cooling Solutions

White Paper

Mark Nowell: Cisco

Attila Aranyosi, Vu Le, Jeffery J. Maki: Juniper Networks

Scott Sommers, Tom Palkert, Weiming Chen: Molex

3/12/2018

QSFP-DD-Enabling 15 Watt Cooling Solutions

1	Table of Contents	
2	Summary.....	3
3	Intro to QSFP-DD	3
3.1	QSFP-DD Hardware Specification.....	4
3.2	QSFP-DD Management Specification.....	5
3.2.1	Module Power Management State Machine	5
3.3	QSFP-DD module cage configurations	6
4	Equipment design options.....	8
4.1	Belly to Belly design	8
4.2	Stacked design.....	8
5	Thermal requirements.....	9
5.1	Required Module Power:	9
5.2	Industry experience and history:	9
5.3	QSFP-DD System design flexibility:	10
6	Thermal testing.....	10
6.1	Stacked Cage Thermal Testing	10
6.2	Belly-to-Belly Cage Thermal Testing	11
6.3	Test Setups and Conditions.....	11
6.3.1	Stacked cage test case:	11
6.3.2	Belly-to-Belly thermal test case:	12
6.4	Test Sample Preparation Details	13
6.5	Test Cases and Results	15
6.5.1	Temperature Rise, MSA Case vs Module Inside in 2x1 Stacked Cages.....	15
6.5.2	Clips, or Variable down Force set @ low and high for stacked cages	15
6.5.3	Prototype 2x1 stacked Cage Pressure Drops vs CFM	16
6.6	Stacked cage Raw Thermal Test Data	17
6.7	Belly-to-Belly thermal test results.....	18
7	Summary.....	19
7.1	Stacked cage thermal testing summary.....	19
7.2	Belly-to-Belly thermal testing summary	20
8	Conclusions.....	20

Figure 1: QSFP-DD module.....	4
Figure 2: Stacked connector inside the 2x1 cage	4
Figure 3: Surface mount connector in a 1x1 cage	5
Figure 4: QSFP-DD Reference Diagram.....	5
Figure 5: Module State Machine for controlling module power	6
Figure 6: QSFP-DD Stacked card cage	7
Figure 7: QSFP-DD single card cage	7
Figure 8: Copper cable and optical module in a QSFP-DD stacked card cage	7
Figure 9: Belly to Belly QSFP-DD card cages	8
Figure 10: Stacked QSFP-DD cages	9
Figure 11: Examples of QSFP-DD side-to-side airflow heat sinks and heat pipes	10
Figure 12: Stacked Cage Thermal Test Setup.....	11
Figure 13: Thermal Test Setup for Belly-to-Belly Configuration.....	12
Figure 14: Load Modules and Faceplate Venting for Belly to Belly Configuration.....	13
Figure 15: Heat Sources on Top and Bottom Side of PCB with Maximum Power Dissipation Capability per Region	13
Figure 16: Thermal Couple on module case near nose according to IA OIF-Thermal-01.0.....	14
Figure 17: Thermal Couple inside module case near center of HS pedestal	14
Figure 18: Pressure drop across the stacked cage vs the CFM flow rate	16
Figure 19: 400GE Pluggable Thermal Assessment.....	19
Figure 20: 1U switch with 36 stacked QSFP-DD cages used to generate thermal test data in Figure 18	20

2 Summary

The thermal performance of the QSFP-DD module is evaluated for use in a high performance data center environment. Thermal test data is presented and analyzed showing temperature rise vs airflow. Feasibility of networking system being able to effectively cool 15W QDFP-DD modules is shown.

3 Intro to QSFP-DD

The QSFP-DD module form factor is the industry’s smallest 400GbE module providing the highest port bandwidth density. This form factor leverages the industry’s manufacturing capability and cost structure that supports QSFP+ and QSFP28, the industry’s most prevalent form factor for 40GbE and 100GbE, respectively. QSFP-DD can support 36 ports of 400GbE in a single Rack Unit (RU) providing over 14Tb/s of bandwidth.

The QSFP-DD will support:

- 3m of passive copper cables
- 100m over parallel multimode fiber
- 500m over parallel single mode fiber
- 2 km and 10km over duplex single mode fiber
- WDM and Coherent designs
- Backward compatible with all QSFP based transceivers from 40G to 200G

3.1 QSFP-DD Hardware Specification

The Hardware specification covers the following items:

- a) Electrical interfaces including pad assignments for data, control, status and power supplies and host PCB layout requirements.
- b) Optical interfaces (including optical receptacles and mating fiber plugs for multimode and single-mode duplex and parallel fiber applications). Breakout cable applications are also specified.
- c) Mechanical specifications including dimensions and tolerances for the connector, cage and module system. Figure 1 shows a drawing of the module. Figure 2 shows a cutaway view revealing the stacked connector inside the 2x1 cage.



Figure 1: QSFP-DD module

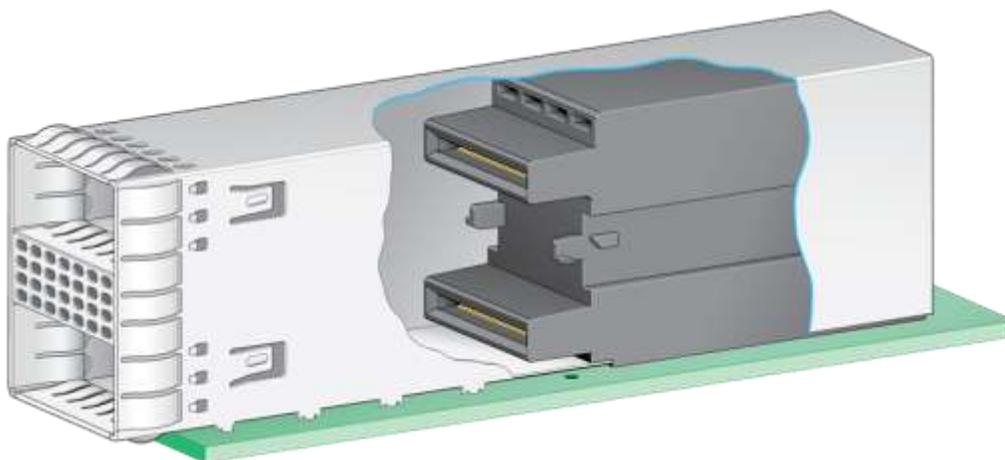


Figure 2: Stacked connector inside the 2x1 cage

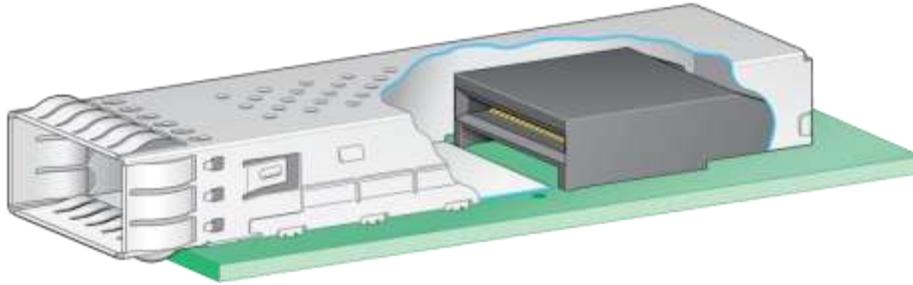


Figure 3: Surface mount connector in a 1x1 cage

A reference diagram showing the connection of a QSFP-DD module to a host system is shown in Figure 4. Retimers are used in both directions of data traffic.

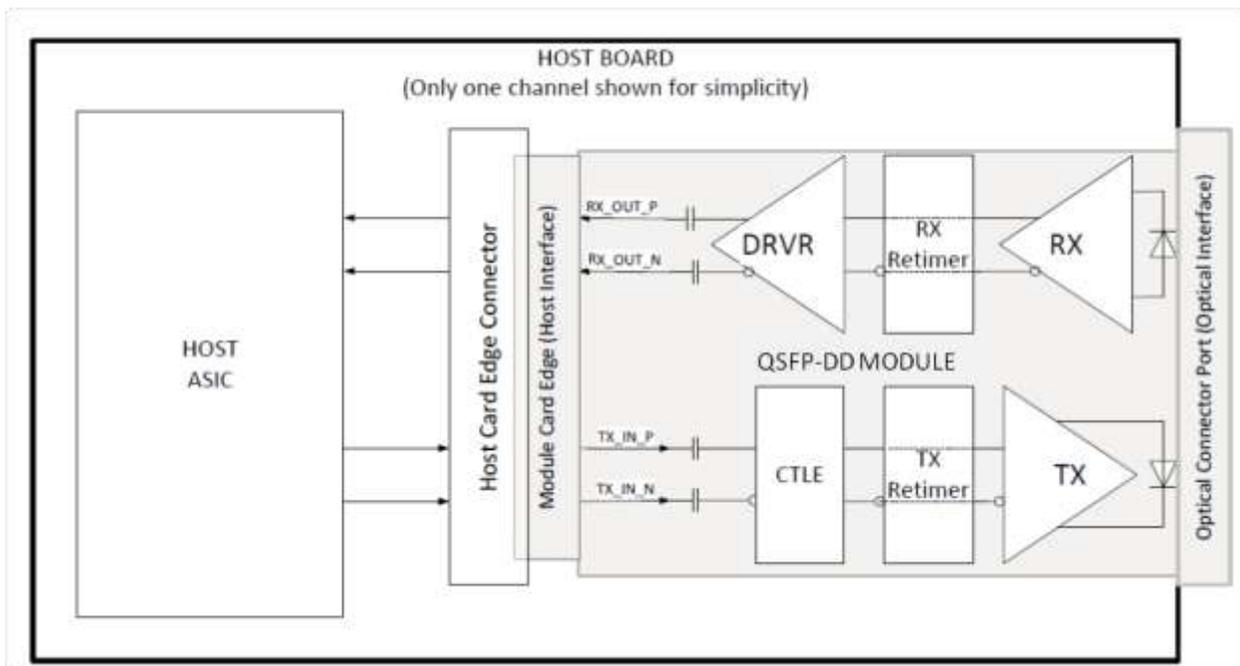


Figure 4: QSFP-DD Reference Diagram

3.2 QSFP-DD Management Specification

The QSFP-DD Management Specification defines the methodology and registers required for the host to control the QSFP-DD module power dissipation. Communication between Host and Module is done via a Two Wire serial Interface.

3.2.1 Module Power Management State Machine

The host controls the power dissipation of the module using the Module State Machine (see Figure 5). The purpose of the Module State Machine is twofold: to provide both a description of the bring-up process for the management interface as well as to identify the module-wide power mode. The State Machine is engaged after module insertion and power on. For the purposes of thermal management, the State Machine allows the Host to:

- 1) Reject a module if the host thermal management does not support the maximum power dissipation indicated by the module
- 2) Control the transition of the module from the low power state to the high power state
- 3) Partially power up the module. This would typically involve powering up a select number of lanes.

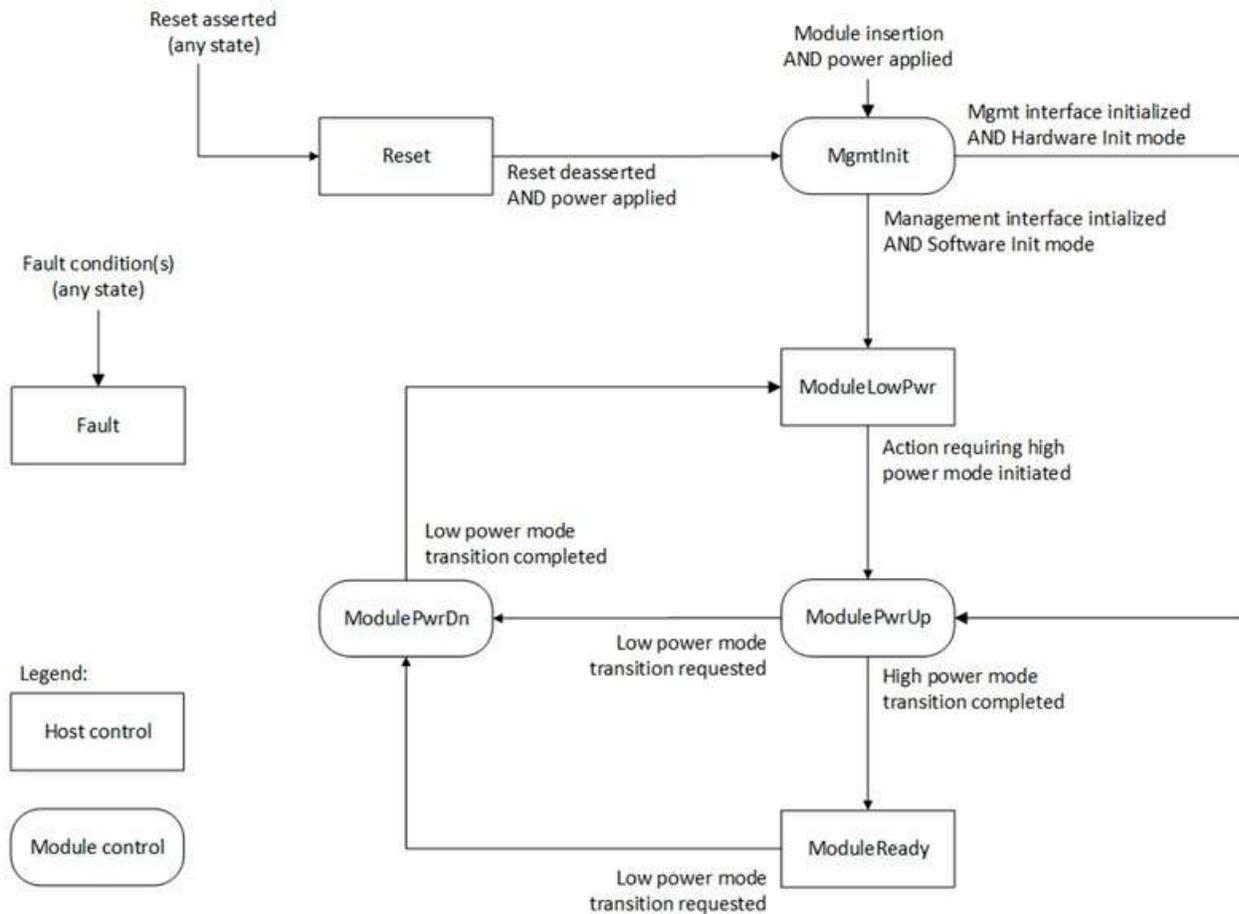


Figure 5: Module State Machine for controlling module power

3.3 QSFP-DD module cage configurations

For maximum flexibility in system design the QSFP-DD cages are defined for both single height and stacked configurations. (See Figure 6 and Figure 7). Figure 8 shows an image of QSFP-DD modules plugged into a stacked card cage.



Figure 6: QSFP-DD Stacked card cage



Figure 7: QSFP-DD single card cage

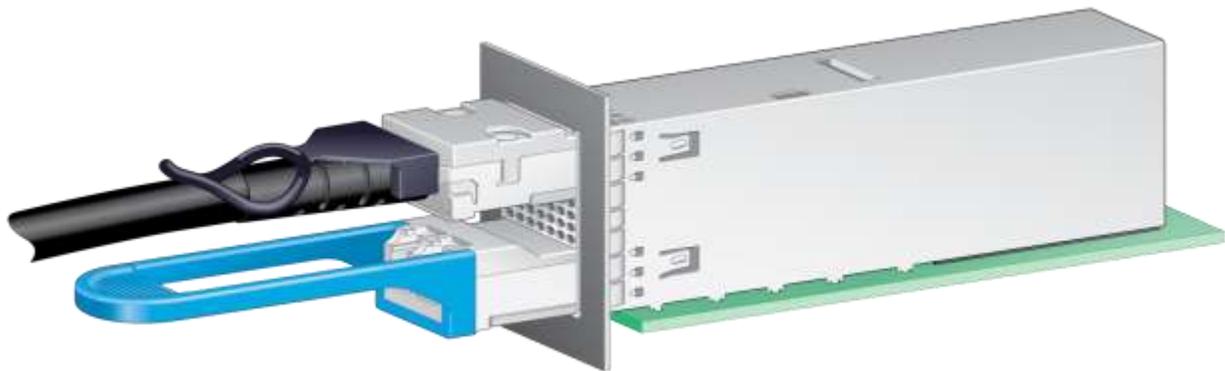


Figure 8: Copper cable and optical module in a QSFP-DD stacked card cage

4 Equipment design options

The thermal cooling capabilities of a system designed for pluggable modules is based on system design optimizations. The QSFP-DD form factor and cage designs provide a foundation upon which system designs can be built. System thermal designs for high density systems take advantage of different circuit board layouts, fan design, mechanical design features and airflow control to allow optimization of routing, module placement and airflow.

4.1 Belly to Belly design

Figure 9 shows an example of a 'belly to belly' layout which places the modules on opposite sides of the circuit board. In this design the air flows on both sides of the printed circuit board providing some advantages for cooling of the modules. It also gives better signal integrity of the high speed traces into the modules as compared to a stacked card cage. A challenge of this design can be the limited height of components on the printed circuit board.



Figure 9: Belly to Belly QSFP-DD card cages

4.2 Stacked design

Figure 10 shows an example of a stacked layout which places the modules on the same side of the circuit board. In this design the air flows on one side of the printed circuit board. This provides advantages for cooling of the host switch chip because it maximizes the heat sink height. Challenges of the stacked design include signal integrity of the high speed traces to the upper stacked card cage and cooling of the lower cage module.



Figure 10: Stacked QSFP-DD cages

5 Thermal requirements

5.1 Required Module Power:

The challenge of designing a piece of equipment with pluggable modules is that each socket must be capable of handling the maximum thermal load. The minimum thermal load would be represented by a direct attach copper cable; with MMF modules and 500m, 1km, 2km, 10km, 40km, and 80km SMF modules representing increasing power with reach. Based on a survey of all anticipated optical module types and reaches cooling at least 15W is necessary to support all possibilities up to and including the 80km reach DWDM pluggable modules.

5.2 Industry experience and history:

There is significant industry experience in building (and cooling) the dominant small form factor family of modules. This includes the smaller SFP single lane modules and the (QSFP-DD plug compatible) QSFP four lane modules. Both the SFP and QSFP modules are in high volume use in network switches today. This extensive experience in module, cage and system design has resulted in significant optimization of thermal designs. All of this experience is translatable to QSFP-DD, with further innovation providing support for 15W of dissipated power in 400GE products using either stacked or belly-to-belly cage designs.

5.3 QSFP-DD System design flexibility:

QSFP-DD's advantage is its system design flexibility that system designers can take advantage of to optimize the system thermal performance. Without this flexibility, it would greatly reduce the variations of system products that customers desire. The flat top design enables riding heat sinks and/or heat pipes to be optimized with numerous thermal innovations possible. Examples of this system design flexibility include range of port inlet, port exhaust and side-to-side cooling options. Figure 11 shows an example of QSFP-DD designs supporting side-to-side airflow and heat pipes.



Figure 11: Examples of QSFP-DD side-to-side airflow heat sinks and heat pipes

6 Thermal testing

The objective of the thermal tests are to characterize the QSFP-DD module and cage thermal performance under specified operation ranges, and to validate that QSFP-DD solutions are robust when subject to temperature extremes encountered during product operation. The purpose of the report is to document detailed airflow and thermal results using ΔT_{rise} as the key parameter for system design. Systems are designed from a thermal perspective to maintain the rise in temperature above the external ambient temperature (ΔT_{rise}) for components to be below a certain limit. For systems designed for QSFP-DD modules operating in cloud datacenters, the target thermal performance is to hold the ΔT_{rise} for the QSFP-DD case temperature to below 30°C.

6.1 Stacked Cage Thermal Testing

The ability to cool modules in both upper and lower slots is enable by heat sinks integrated into the 2x1 card cage. Testing was performed to determine the thermal performance of the module/cage/heatsink combination with high power optical modules. The module thermal testing was performed using 2x1 cages placed adjacent to each other representing a 1U switch. Significant thermal testing focus has been on the fixed 1RU system designs as they typically are the most challenging from a thermal design perspective. Fan space is limited and represents the hardest module-cooling thermal design. Modular systems designs with line cards that pull

out typically have larger fans and are able to provide higher airflow across the components. (Typically ΔT_{rise} for a QSFP-DD pluggable module is 5-7°C lower in a modular system as compared to a fixed design)

6.2 Belly-to-Belly Cage Thermal Testing

The belly-to-belly design uses heat sinks mounted on the top of surface mount cages. This design should provide optimum module airflow in a 1U switch design.

6.3 Test Setups and Conditions

6.3.1 Stacked cage test case:

This is a component level thermal test with two 2x1 QSFP-DD cages set side by side in each test. All thermal tests were conducted under room temperature ranged from 20°C to 22°C and at sea level. The airflow direction was front to back and the airflow ranges used in the tests were considered as typical for system designs. In some test cases, a force gauge/cell was used to set the heat sink down force at a specified value in order to reach consistent test results.

- 20 ~ 22°C ambient
- 1 atmosphere at sea level
- Two 2x1 QSFP-DD cages and four QSFP-DD thermal modules as loads
- Thermal module power profile programed as (0, 2.01, 1.61, 2.82, 0, 2.01, 1.61, 2.42) for total of 12.48 watts
- Total airflow at 3, 5, and 7 CFM
- One Thermocouple on module case near nose and one Thermocouple inside the module as shown in Figure 16 and Figure 17
- Tested production cage and heat sinks, clips with variable down forces
- Thermal interfaces between modules and heat sinks are dry with no Thermal Interface Materials

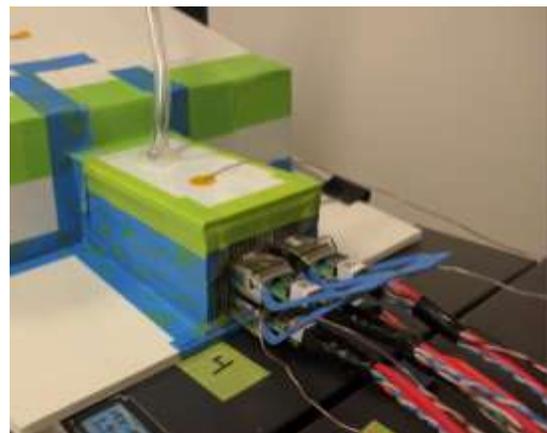
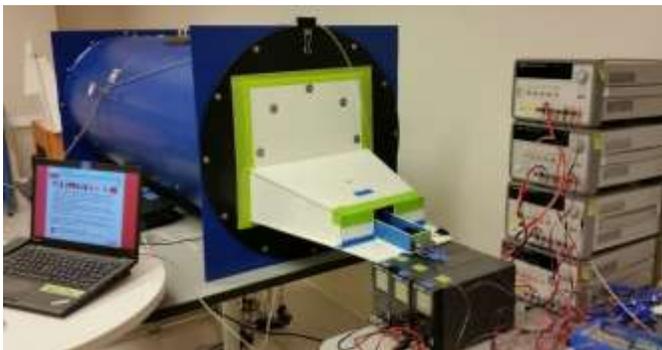


Figure 12: Stacked Cage Thermal Test Setup

6.3.2 Belly-to-Belly thermal test case:

This is a component level thermal test with two 1x2 QSFP-DD cages set on either side of a test board. All thermal tests were conducted at room temperature and the measured module temperatures were normalized to above-ambient temperature rises. The airflow direction was front to back and the airflow ranges used in the tests were considered as typical for system designs.

- 1 atmosphere at sea level
- Two 1x2 QSFP-DD cages and four QSFP-DD thermal modules as loads
- Thermal module power profile programmed for 14W, 15W with scaled power maps. One variant of the 15 watt power maps was done with heat dissipation biased to center/rear of module
- Total airflow at 4-6.4 CFM per port (achieved with counter-rotating fans at 1-2.5 in. H₂O pressure drop)
- One Thermocouple on module case near nose and one Thermocouple inside the module as shown in Figure 16 and Figure 17

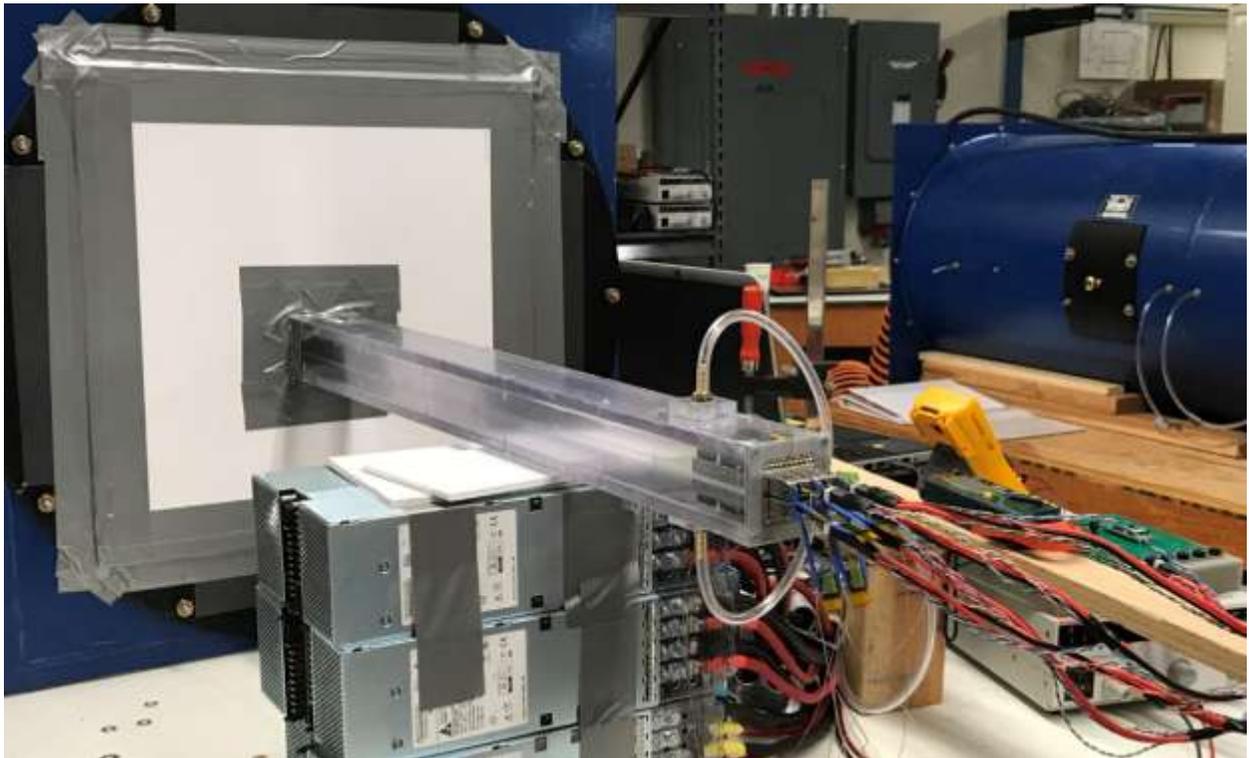


Figure 13: Thermal Test Setup for Belly-to-Belly Configuration

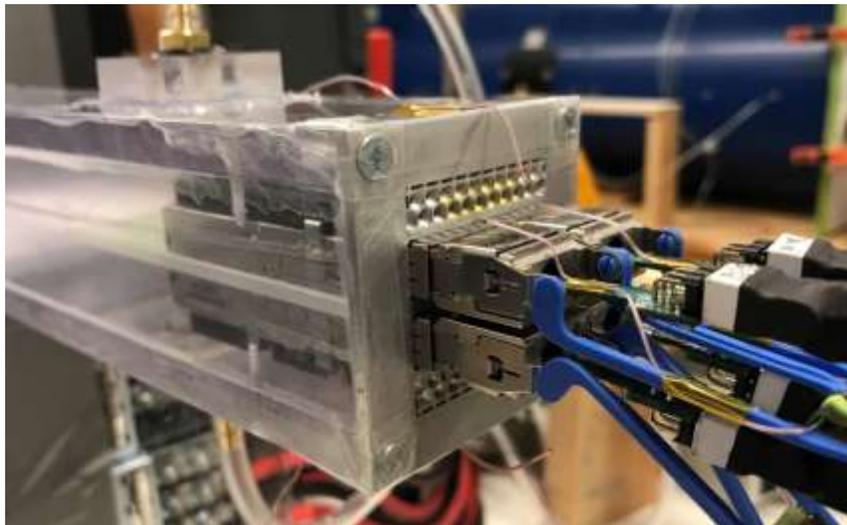


Figure 14: Load Modules and Faceplate Venting for Belly to Belly Configuration

6.4 Test Sample Preparation Details

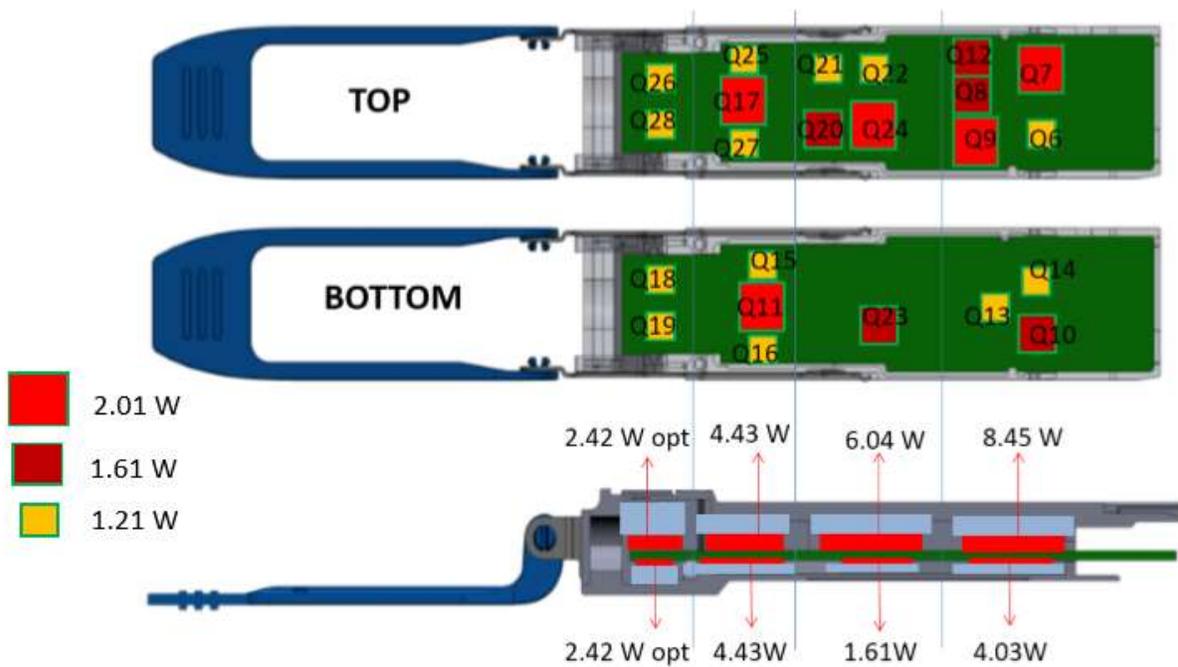


Figure 15: Heat Sources on Top and Bottom Side of PCB with Maximum Power Dissipation Capability per Region

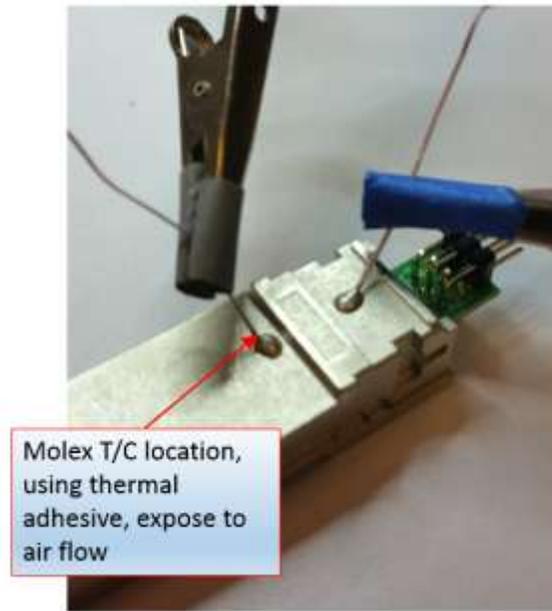


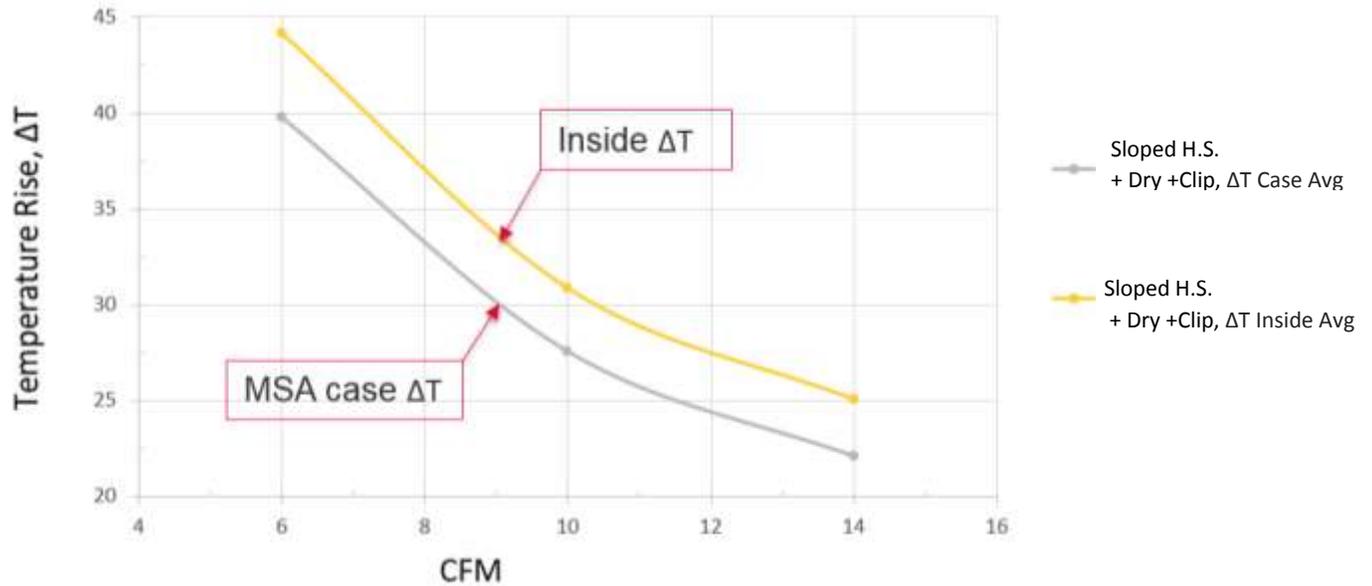
Figure 16: Thermal Couple on module case near nose according to IA OIF-Thermal-01.0



Figure 17: Thermal Couple inside module case near center of HS pedestal

6.5 Test Cases and Results

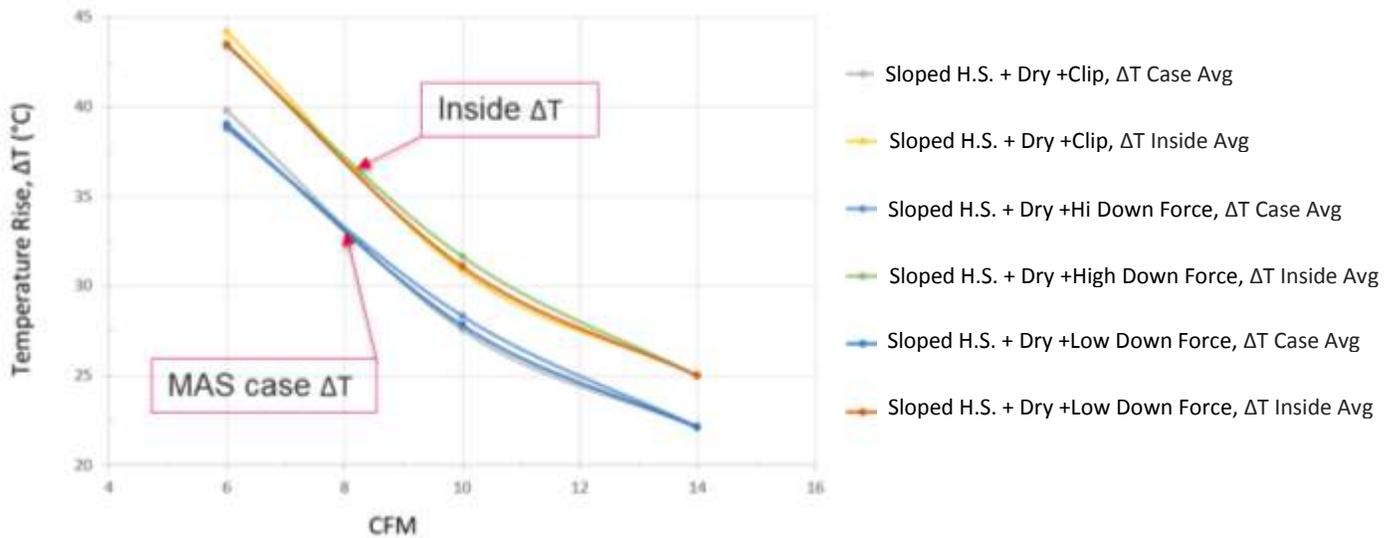
6.5.1 Temperature Rise, MSA Case vs Module Inside in 2x1 Stacked Cages



Note: The reported case and inside temperatures are the average from four modules under test

Observation: Module inside temperature rise near hot spot area are 3~5°C hotter than MSA case temperature rise depending on the CFM

6.5.2 Clips, or Variable down Force set @ low and high for stacked cages



Note: The reported case and inside temperatures are the average from four modules under test

Observation: Results for top heat sink using clip, or pre-set down force set at low or high were about the same

6.5.3 Prototype 2x1 stacked Cage Pressure Drops vs CFM

Figure 18 shows the measured pressure drop across the stacked cage vs the CFM flow rate.

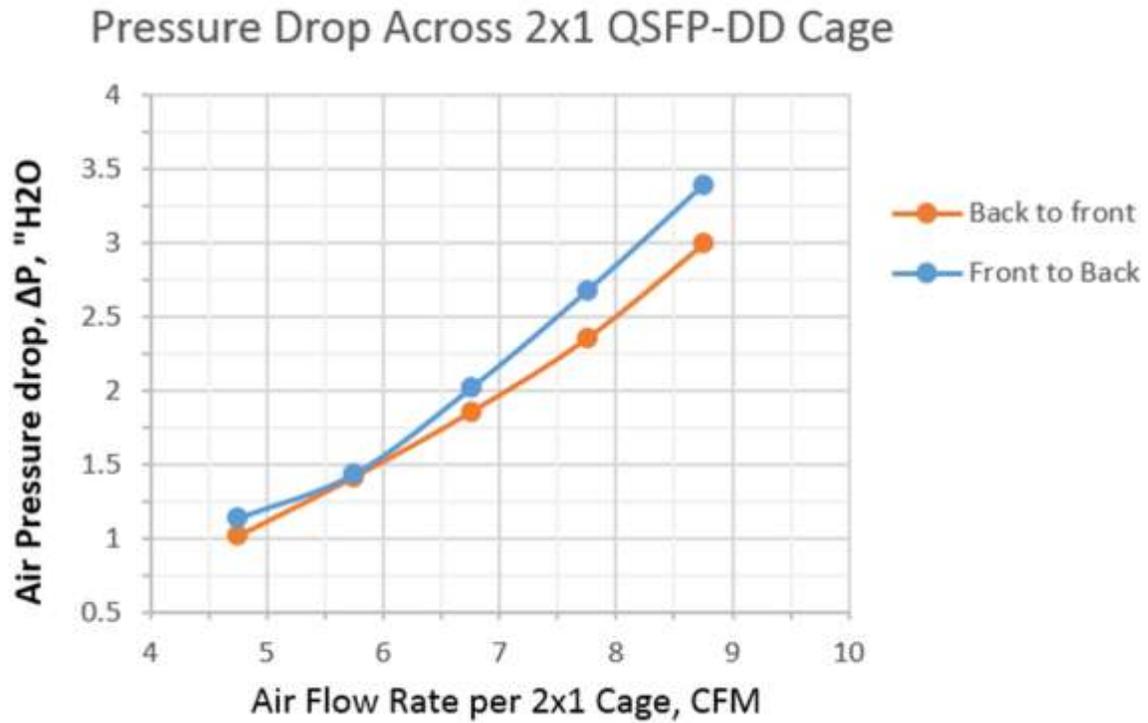


Figure 18: Pressure drop across the stacked cage vs the CFM flow rate

6.6 Stacked cage Raw Thermal Test Data

Sloped H.S. + Dry +Clip, ΔT Case Avg

Sloped H.S. + Dry +Clip, ΔT Inside Avg

		Top				Bottom				Avg. Temperature		Temperature Rise	
		Left		Right		Left		Right		Case	Inside	Case	Inside
CFM	Tamb	Tcase	T Inside	T case Avg	T Inside Avg	ΔT Case Avg	ΔT Inside Avg						
3	19.5	56.2	60.7	54.6	59.6	63.6	68.1	62.6	66.3	59.3	63.7	39.8	44.2
5	20.3	45.9	49.2	44.6	48.4	50.6	54.3	50.3	52.7	47.9	51.2	27.6	30.9
7	21	41.5	44.5	40.8	44.1	45.2	48.6	44.9	47	43.1	46.1	22.1	25.1

Sloped H.S. + Dry +Hi Down Force, ΔT Case Avg

Sloped H.S. + Dry +High Down Force, ΔT Inside Avg

		Top				Bottom				Avg. Temperature		Temperature Rise	
		Left		Right		Left		Right		Case	Inside	Case	Inside
CFM	Tamb	Tcase	T Inside	T case Avg	T Inside Avg	ΔT Case Avg	ΔT Inside Avg						
3	20.3	56.2	61	54.7	59.8	62.6	67.3	62.5	66.5	59.1	63.7	38.8	43.4
5	20.5	47.1	50.3	45.7	49.5	51.4	55.1	55.1	53.6	48.8	52.1	28.3	31.6
7	20.2	41	43.7	39.9	43.3	44.4	47.5	44.1	46.2	42.4	45.2	22.2	25.0

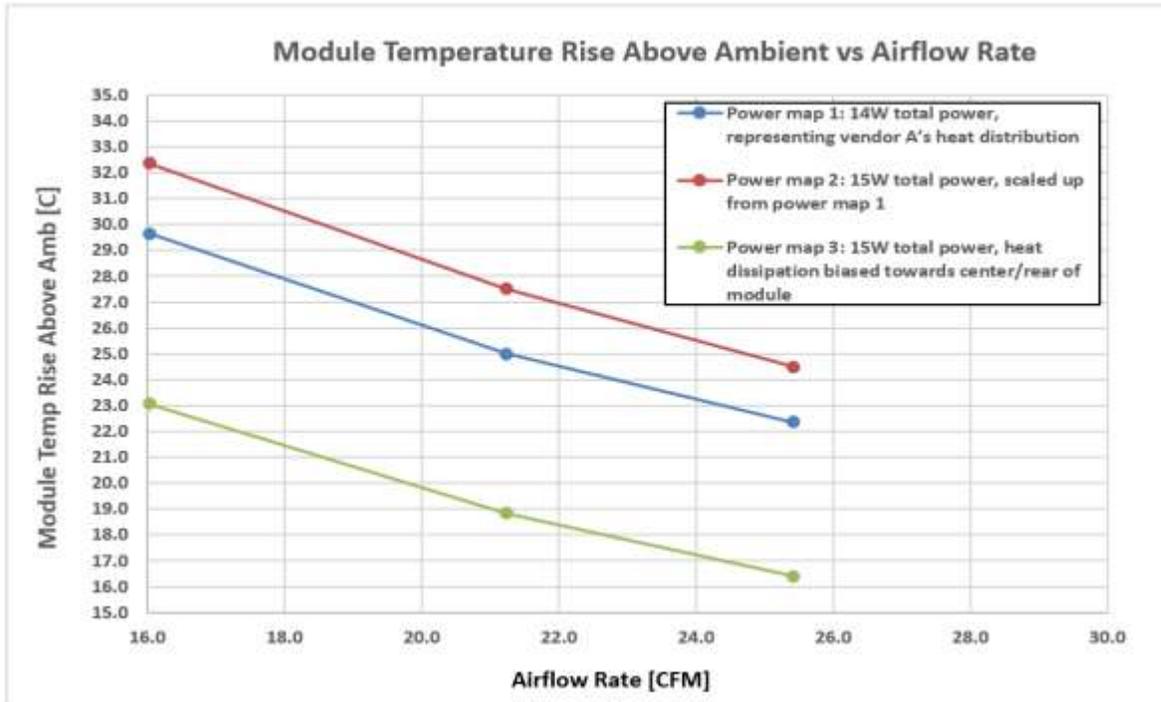
Sloped H.S. + Dry +Low Down Force, ΔT Case Avg

Sloped H.S. + Dry +Low Down Force, ΔT Inside Avg

		Top				Bottom				Avg. Temperature		Temperature Rise	
		Left		Right		Left		Right		Case	Inside	Case	Inside
CFM	Tamb	Tcase	T Inside	T case Avg	T Inside Avg	ΔT Case Avg	ΔT Inside Avg						
3	20.6	57.4	62.1	55.6	60.6	62.7	67.1	62.7	66.4	59.6	64.1	39.0	43.5
5	20.9	47.3	50.7	45.8	49.7	50.7	54.2	51	53.3	48.7	52	27.8	31.1
7	20.7	41.9	44.8	40.5	44.1	44.2	47.4	44.6	46.5	42.8	45.7	22.1	25.0

6.7 Belly-to-Belly thermal test results

The belly-to-belly thermal testing was performed with both 14W and 15W module power dissipation. One variant of the 15 watt power maps was with heat dissipation biased to center/rear of module. The results show that biasing of the power dissipation to the center/rear of the module can result in significant improvement in thermal performance.



7 Summary

7.1 Stacked cage thermal testing summary

- Temperature rises from using clips, pre-set at low or pre-set at hi, resulted very close temperature results, an indication of proper clip force from the clip design
- Module average case temperature rise is around 21~ 22°C when airflow per 2x1 cage is about 7 CFM. The temperature rise plots indicated that the module case temperatures rise can be less than 20°C if the airflow is more than 8 CFM per 2x1 cage
- In most cases, the bottom modules in 2x1 cages will run 2 ~ 4°C hotter than top modules under the CFM ranges tested

Figure 19 is an example of the measured curves of QSFP-DD modules in a 36 port 1RU system design. The ΔT_{rise} vs power graph confirms the ability of the QSFP-DD module/cage combination to support the required 15W with significant margin to the max target of $\leq 30^\circ\text{C}$ temperature rise.

Figure 20 shows the 1U switch design used to generate thermal test data curves in Figure 19.

For comparison, the results from a similar 32 port 1RU system designed with OSFP modules are shown. However, it is noted that these results are based on a 1st generation design for the OSFP cage and, while there is not much system thermal design flexibility with these modules, optimizations may be possible.

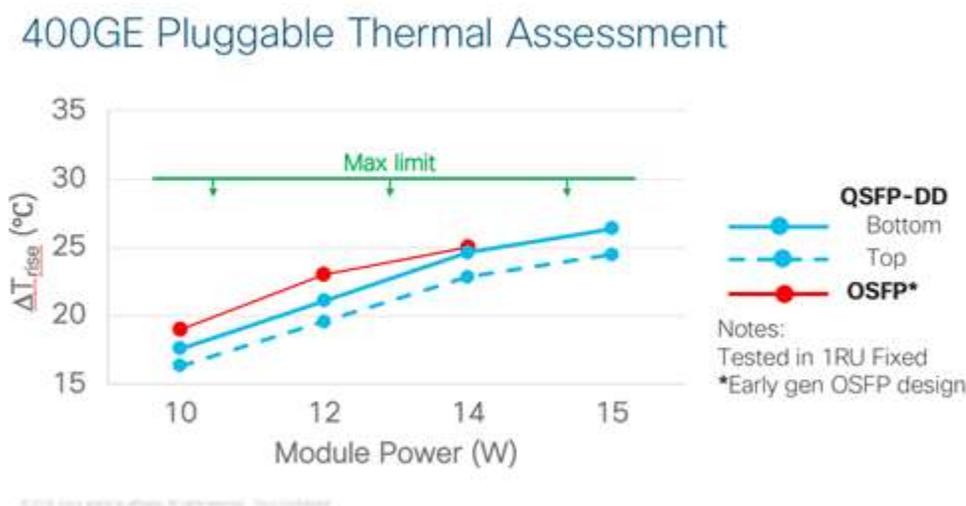


Figure 19: 400GE Pluggable Thermal Assessment



Figure 20: 1U switch with 36 stacked QSFP-DD cages used to generate thermal test data in Figure 19

7.2 Belly-to-Belly thermal testing summary

- In a system with front-to-back airflow and off-the-shelf heat sinks the case temperature of a 15W QSFP-DD module can be kept below 70°C in a 46°C ambient environment with 6.4 CFM per module airflow. This is achievable with counter rotating fans at 2.5" H₂O pressure drop.
- At 40°C ambient air temperature the maximum module power dissipation increases to 18W.
- Optimization of the thermal environment and/or custom heat sinks (taller fins and/or higher fin density) can increase maximum module power dissipation to beyond 18W.

8 Conclusions

QSFP-DD represents a flexible, low cost solution that leverages years of past system and module/cage thermal experience. It offers the greatest flexibility to the industry to customize their products as they see fit.

It has been demonstrated that the QSFP-DD module/cage design in both stacked and belly-to-belly configurations can support thermal loads beyond 15W as needed for next generation high bandwidth switching applications.

All testing in this report was based on 1RU system designs and demonstrated that measured ΔT_{rise} values had significant margin against targets necessary for deployment. The improved airflow available in modular systems provides a further $\sim 5\text{-}7^\circ\text{C}$ ΔT_{rise} which is further encouraging since the typically higher power modules, such as 400ZR modules for 80km DWDM, would be deployed in modular systems.