DesignCon 2018

Examining System Challenges When Implementing Next Generation Data Center Input/Output (I/O) Connectivity

Nathan Tracy, TE Connectivity
ntracy@te.com, 717-986-7546

Abstract

Abstract: As the networking industry moves to develop higher data rates to support next generation demands, there are also simultaneous demands to support higher Input/Output (I/O) port densities, higher I/O module power dissipation, improved signal integrity, maintain cable reach and satisfactory electromagnetic interference (EMI) performance. In this paper, we will discuss the comparative differences of three state-of-the-art port types including micro quad small form-factor pluggable (microQSFP), quad small form-factor pluggable double density (QSFP-DD) and octal small form-factor pluggable (OSFP). These new I/O ports address the conflicting performance demands in different ways and the comparative performance differences will be presented using a combination of simulation and measurement methods. A summary will be provided highlighting the relative ability of the different port types to meet next generation market needs.

Author(s) Biography

Nathan Tracy has over 30 years of experience in technology development, marketing, and business development for TE Connectivity (TE) in areas including RF and microwave, automotive, wireless, networking, and telecom. Currently, he is a Technologist on the system architecture team and the manager of industry standards, driving standards activities and working with key customers on new system architectures in TE's Data and Devices business unit.

Nathan is active in several industry standards and associations. He currently serves the OIF as a member of the board of directors and VP of marketing and is a regular attendee and contributor to the Ethernet Alliance, IEEE 802.3 and COBO. He is currently the chair of the microQSFP multi-source agreement (MSA) group and recently led the CDFP MSA group. He is active in a number of other industry MSAs and forums.

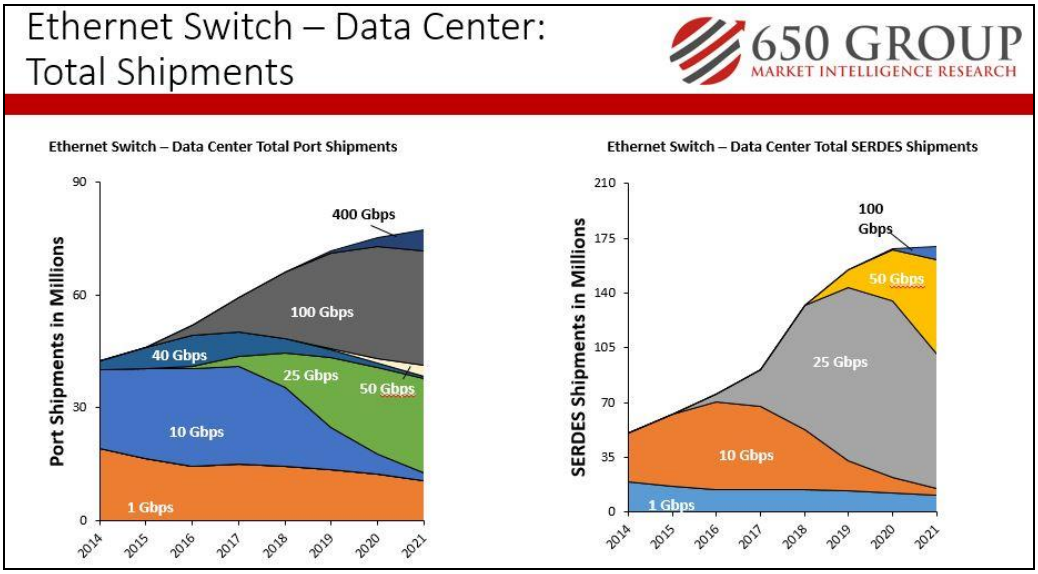
Nathan earned his Bachelor of Science Electrical Engineering Technology degree from the University of Massachusetts, Dartmouth.

1.0 Introduction

1.1 Electrical Bandwidth Density I/O Trends

Everyone has witnessed the seemingly unending growth in network and data center traffic across the industry, and it's the reason that industry forums like DesignCon exist: to discuss how these challenges will be technically addressed. Whether remembering the electrical signaling transition several years ago from 1 gigabits per second (Gbps) to 10 Gbps or more recently the jump from 10 Gbps to 25 Gbps, the trend continues with no end in sight.

The below figures show the current and projected data center switch port shipments. The graph on the left shows data center port shipments by port data rate. The graph on the right shows serializer/deserializer (SerDes) (silicon chip) shipments by data rate per SerDes port. Keep in mind that multiple SerDes "pairs" are used per data center I/O port. For example, four 25 Gbps SerDes pairs might be used per 100 Gbps data center port. Eight 50 Gbps SerDes pairs might be used per 400 Gbps and 2 per 100 Gbps data center port. The projected growth in numbers of ports and port rate demonstrates the market needs to be addressed with a technical solution.



Market data showing increasing port counts and port rates for data centers and SerDes chips (graph is used with the permission of 650 Group LLC)

In order to meet the network and data center operator demands, there is a requirement for increases in per channel data rates as well as a greater density of channels per port. In the standard 1RU enclosure, 48 ports of 10 Gbps (480 Gbps) used to be adequate, however, requirements have currently graduated to at least 32 ports of 100 Gbps (3.2 Tbps) with some applications having up to 36 ports.

Data rates that the industry currently use to deliver high levels of electrical bandwidth are 25 Gbps per port (x1) and/or 100 Gbps per port (x4) by using 25 Gbps channels. The current state of the art in the networking industry to provide these I/O ports effectively in a 1 rack unit (1RU) enclosure means using 48 ports of the single channel small form-factor pluggable (SFP) form factor or at least 32 ports of the quad channel small form-factor pluggable (QSFP) form factor as the electrical interfaces to the switch equipment. This results in a total of 48 or 128 electrical channels at 25 Gbps each in a 1RU enclosure.

These bandwidth density levels are shipping today, but cannot meet the future needs of data centers and network operators. A significant increase in bandwidth density is necessary for the next generation network performance.

Modular equipment showing various 1RU line cards with 36, 9, 48 and 32 ports

1.2 Next Generation Electrical Bandwidth Density I/O Challenges

When looking at industry projections, standards activities, and product roadmaps, the next level of increasing performance expectations will range from 64 ports of 100 Gbps and up to 36 ports of 400 Gbps. These are being enabled by silicon switch suppliers working to satisfy the network operators. The silicon roadmaps are well established and the chips are on the way to becoming reality. To achieve this next generation requirement of 64 ports of 100 Gbps and 32 ports of 400 Gbps with 25 Gbps electrical channels, we would require 256 and 512 channels respectively which creates a logistical and implementational impracticality for today's I/O form factors. Given the recent industry development of 50 Gbps Pulse Amplitude Modulation with 4 levels (PAM4) signaling technology, a more suitable solution would be to find a way to enable at least 256 channels of 50 Gbps. 256 channels of 50 Gbps is the implementation challenge being addressed here. This represents a doubling of the number of electrical channels from what ships today (128 channels x25 Gbps).

A brief overview of the challenges being addressed: As we double the number of signaling channels we worry first about the impact of signaling density on electrical performance. Cross-talk due to closer proximity of channels to each other is an obvious concern, but also the quality of the channel in terms of parameters such as return loss must also be considered. Additionally, the reach of the electrical channel is an important consideration. In the case of a switch chip connected to a plugged in optical module, this reach is determined by the host PCB and port's high speed signaling characteristics. In the case of a switch chip connected to a plugged-in passive direct attach copper cable, the wire gauge accommodated by the plug form factor is a dominant factor affecting the reach of the electrical channel. Once we meet the requirements of electrical signaling quality and reach, then the next challenge to be addressed is thermal management. By doubling the number of channels in a 1RU faceplate, we have forced the switch designer to have twice the number of pluggable electrical-to-optical conversion channels in the same space which significantly increases the power consumption and therefore the thermal dissipation density.

2.0 Next Generation Solutions

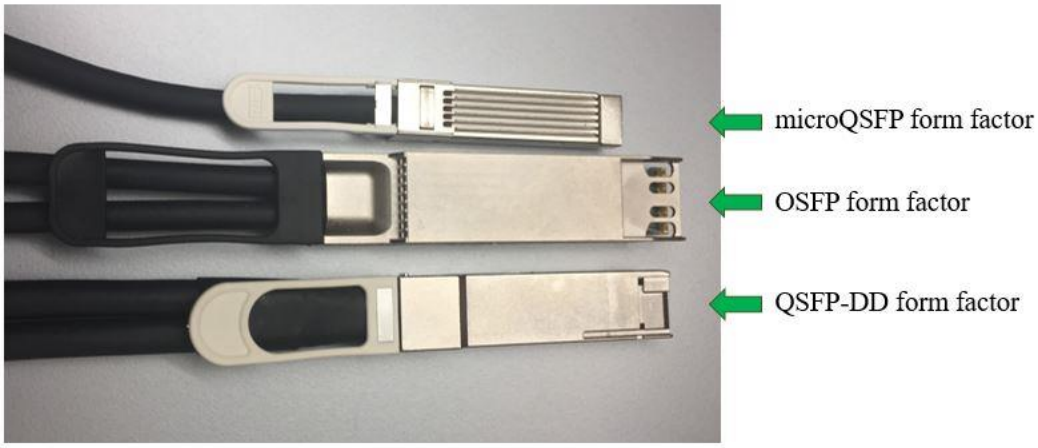
2.1 The Candidates

The task of fitting 256 channels of connectivity in the same 1RU space which traditionally housed 128 channels is currently being addressed. The industry has developed two different approaches to this density challenge: decrease the pitch of the connector contacts so that more contacts and connectors can be squeezed in to a smaller space, or add additional rows of overlapping contacts to each connector. This study looks at three I/O port implementations - microQSFP, OSFP and QSFP-DD, that utilize these connector density techniques to enable 256 channels of 50Gbps PAM4 signaling at the 1RU face plate. The three implementations have taken different approaches to solving this challenge, and therefore bring different design implications to the switch designer. This study focuses on the switch design considerations that are driven by I/O port choice.

As a major supplier to the networking and data center industry, TE Connectivity has firsthand knowledge on this subject as TE is a founding member of all three industry multi-source agreements (MSAs) that are defining these three new form factors and TE is providing all three to customers today.

2.2 Form Factor Overview

The three I/O ports being considered all have common attributes: they each define a pluggable form factor (module) and connector/cage assembly into which the module is plugged, allowing the switch user to add and subtract module capacity from the face plate of their switch.



2.2.1 microQSFP Form Factor

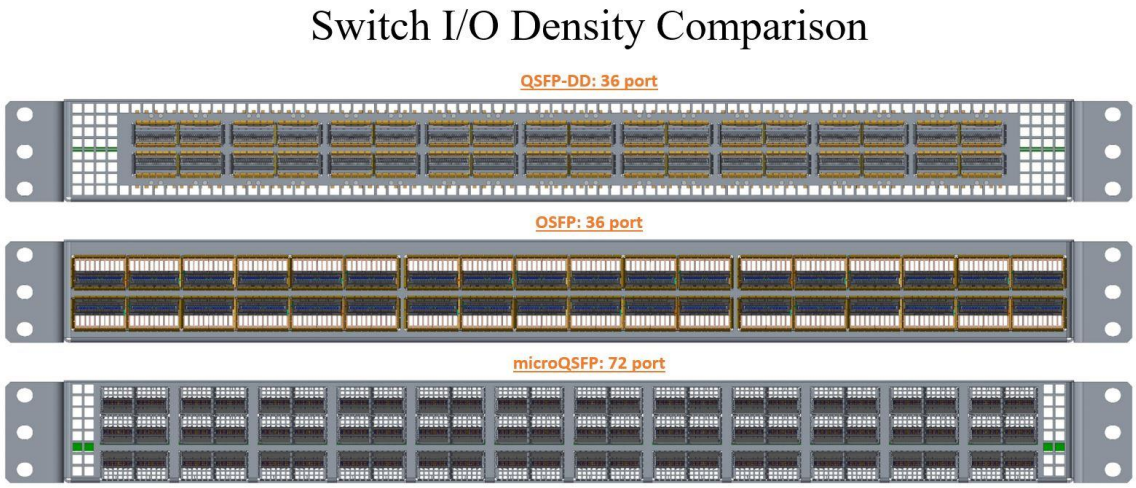
The microQSFP form factor consists of 4 channels implemented on a 0.6mm contact pitch vs. today's QSFP form factor which has a contact pitch of 0.8mm. This decrease in contact pitch allows the width of the module to be reduced such that 24 ports can fit across the width of a 1RU line card (same width as SFP modules). The design enables the 256-channel connectivity objective by allowing three-high port stacking within a 1RU face plate (see below figure). This allows a total of 64 ports per 1RU face plate with 4 channels per port. In fact, 72 ports can fit, but for purposes of this study, 64 is sufficient to meet the 256-channel requirement. To allow this three-high stacking port density, microQSFP developed a new solution for thermal management which increases the power capacity of each module by improving thermal dissipation. microQSFP is a new industry form factor and this connector/cage can offer backwards compatibility to SFP modules via a module adapter.

2.2.2 OSFP Form Factor

The OSFP form factor consists of 8 channels also implemented on 0.6mm contact pitch. The 8 channels result in a module width that allows 18 ports to fit across the width of a 1RU line card. OSFP allows stacking ports two-high, yielding 36 ports to fit into the 1RU space. Only 32 eight channel modules are required to address the 256 total channel objective. Similar to microQSFP, OSFP modules utilize a new system for thermal management to improve power capacity and thermal management due to the increase in channel density. The OSFP connector/cage, like microQSFP, is a new form factor and utilizes an adapter to provide backwards compatibility to QSFP modules.

2.2.3 QSFP-DD Form Factor

The QSFP-DD form factor also consists of 8 channels, but keeps the existing industry contact pitch of 0.8mm. This requires additional rows of overlapping contacts to realize an increase in contact density. The width of the QSFP-DD port also allows 18 ports to fit across the width of a 1RU line card. QSFP-DD ports can be stacked 2-high, yielding 36 ports in the 1RU space, again achieving the 256 channel objective with 32 ports. QSFP-DD utilizes conventional riding heat sinks on the connector/cage for module thermal management. Although QSFP-DD is a new connector/cage, it provides backwards compatibility with existing QSFP modules without the need for an adapter. This is because it uses the same contact pitch and module width as QSFP.



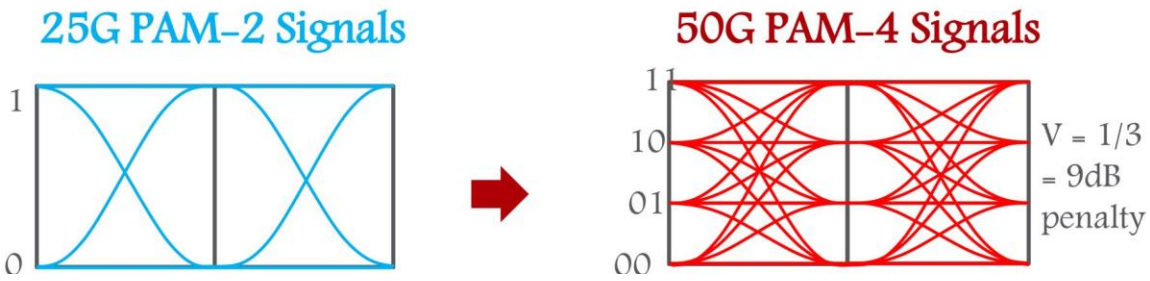
Comparison view of 36 ports of QSFP-DD, 36 ports OSFP and 72 ports of microQSFP

3.0 Electrical Comparison

3.1 Description

50 Gbps PAM4 signaling is based on 25 Giga*baud* transmission so the baud rate is the same as at 25 Gbps NRZ signaling. This means PCB insertion loss and thus channel reach will be similar to or slightly shorter than 25Gbps. Because the PAM4 modulation uses four signal levels (vs. NRZ's two signal levels) we lose 9dB of margin in signal-to-

noise ratio. The consequence of this is that the I/O port connector performance, specifically insertion loss, return loss and crosstalk, must be carefully controlled and be at least as good as 25 Gbps signaling and is preferred to be even better than 25 Gbps NRZ signaling to preserve channel objectives.



"eye diagrams" showing the comparison of NRZ (or PAM2) vs. PAM4 signaling to show the 9dB impact to signal to noise ratio in order to keep the baud rate the same.

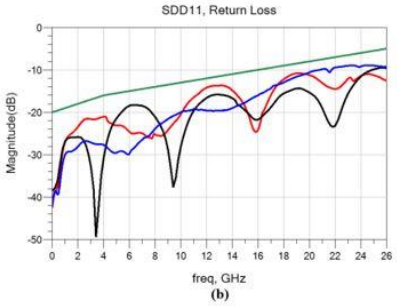
The three I/O port connector designs described above have differing electrical performance by virtue of their mechanical implementation. Both microQSFP and OSFP use a tighter contact pitch of 0.6mm so that all the contacts including both high speed differential contacts and side band contacts can fit into two rows on the host, mating to single rows on the upper and lower sides of the pluggable module PCB. This two-row implementation follows past industry convention and results in a simplistic connector design. QSFP-DD connectors have contacts on 0.8mm pitch and therefore add two additional recessed rows of contacts to fit both the high speed differential pairs and the side band contacts. This results in a four-row connector construction vs. the two rows of microQSFP and OSFP. The tighter contact pitch of microQSFP and OSFP IO port connectors requires care in terms of cross talk, but 0.6mm contact pitch has been demonstrated to provide "at least as good" performance to 0.8mm contact pitch. The extra contact rows required for QSFP-DD also correspond to higher cross talk concerns and the analysis has shown these extra rows can impact performance over a two-row implementation.

3.2 Comparison by Simulation

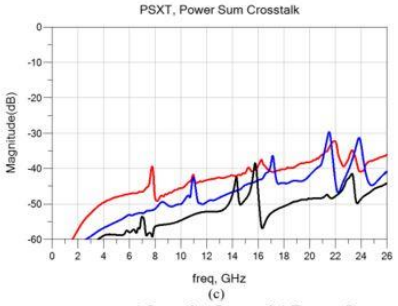
Electrical modeling allows the connector designer and the switch designer to simulate the signal integrity performance that will be achieved from the various connector designs. TE has simulated the connector and mated compliance board performance per the IEEE prescribed mated compliance board configuration and it is provided in the figures below.

Simulated Electrical Performance

(a)



(b)



(c)

	ICN FEXT (mV)	ICN NEXT (mV)	ICN Total (mV)
802.3bs Limits	4.2	1.5	4.4
OSFP	0.989	0.355	1.049
microQSFP	1.629	0.251	1.648
QSFP-DD	2.686	0.517	2.736

(d)

a.) Insertion Loss b.) Return Loss c.) PowerSum Crosstalk d.) ICN Table
OSFP (Black), microQSFP(Blue), QSFP-DD (Red)

Victim and Aggressors for simulated cross-talk

60	59	58	57	56	55	54	53	52	51	50	49	48	47	46	45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
G	Tx1+	Tx1-	G	Tx3+	Tx3-	G	Tx5+	Tx5-	G	Tx7+	Tx7-	G	SB	SB	SB	SB	G	Rx8-	Rx8+	G	Rx6-	Rx6+	G	Rx4-	Rx4+	G	Rx2-	Rx2+	G
G	Tx2+	Tx2-	G	Tx4+	Tx4-	G	Tx6+	Tx6-	G	Tx8+	Tx8-	G	SB	SB	SB	SB	G	Rx7-	Rxy+	G	Rx5-	Rx5+	G	Rx3-	Rx3+	G	Rx1-	Rx1+	G
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30

OSFP Pinout

38	37	36	35	34	33	32	31	30	29	28	27	26	25	24	23	22	21	20
G	Tx1-	Tx1+	G	Tx3-	Tx3+	G	SB	SB	SB	SB	SB	G	Rx4+	Rx4-	G	Rx2+	Rx2-	G
G	Tx2-	Tx2+	G	Tx4-	Tx4+	G	SB	SB	SB	SB	SB	G	Rx3+	Rx3-	G	Rx1+	Rx1-	G
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19

microQSFP Pinout

38	37	36	35	34	33	32	31	30	29	28	27	26	25	24	23	22	21	20
G	Tx1-	Tx1+	G	Tx3-	Tx3+	G	SB	SB	SB	SB	SB	G	Rx4+	Rx4-	G	Rx2+	Rx2-	G
G	Tx5-	Tx5+	G	Tx7-	Tx7+	G	SB	SB	SB	SB	SB	G	Rx8+	Rx8-	G	Rx6+	Rx6-	G
76	75	74	73	72	71	70	69	68	67	66	65	64	63	62	61	60	59	58

39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57
G	Tx6-	Tx6+	G	Tx8-	Tx8+	G	SB	SB	SB	SB	SB	G	Rx7+	Rx7-	G	Rx5+	Rx5-	G
G	Tx2-	Tx2+	G	Tx4-	Tx4+	G	SB	SB	SB	SB	SB	G	Rx3+	Rx3-	G	Rx1+	Rx1-	G
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19

QSFP-DD Pinout

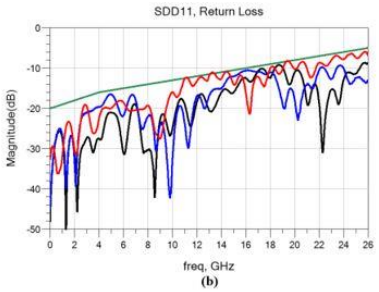
As mentioned above, due to some design similarity, microQSFP and OSFP offer a generally similar level of predicted performance, and one can observe the impact of the different construction of QSFP-DD connector in its results. All three connectors are shown to be compliant to the IEEE 802.3bs specification for 400Gbps, but there is more available margin in the performance of microQSFP and OSFP connectors. These simulations are used extensively to optimize the connector design prior to committing to expensive production tooling and allow design trade-off analysis to be performed. These connector performance simulations are also shared with switch developers so they

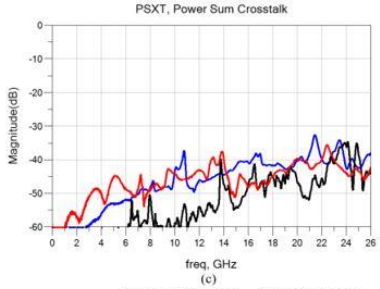
can take the connector and channel performance into account in their switch designs prior to the availability of actual product.

3.3 Comparison of Measurements

As a member of all three I/O port MSAs, TE has produced each connector. The measured connector performance is shown below and, as can be seen, the measured performance of the connectors generally follows the connector simulations.

Measured Electrical Performance





	ICN FEXT (mV)	ICN NEXT (mV)	ICN Total (mV)
802.3bs Limits	4.2	1.5	4.4
OSFP	0.574	0.696	0.902
microQSFP	2.144	0.612	2.229
QSFP-DD	2.416	xx	xx

(d)

a.) Insertion Loss b.) Return Loss c.) PowerSum Crosstalk d.) ICN Table
OSFP (Black), microQSFP(Blue), QSFP-DD (Red)

Victim and Aggressors for measured cross-talk

60	59	58	57	56	55	54	53	52	51	50	49	48	47	46	45	44	43	42	41	40	39	38	37	36	35	34	33	32	31
G	Tx1+	Tx1-	G	Tx3+	Tx3-	G	Tx5+	Tx5-	G	Tx7+	Tx7-	G	SB	SB	SB	SB	G	Rx8-	Rx8+	G	Rx6-	Rx6+	G	Rx4-	Rx4+	G	Rx2-	Rx2+	G
G	Tx2+	Tx2-	G	Tx4+	Tx4-	G	Tx6+	Tx6-	G	Tx8+	Tx8-	G	SB	SB	SB	SB	G	Rx7-	Rx7+	G	Rx5-	Rx5+	G	Rx3-	Rx3+	G	Rx1-	Rx1+	G
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30

OSFP Pinout

38	37	36	35	34	33	32	31	30	29	28	27	26	25	24	23	22	21	20
G	Tx1-	Tx1+	G	Tx3-	Tx3+	G	SB	SB	SB	SB	SB	G	Rx4+	Rx4-	G	Rx2+	Rx2-	G
G	Tx2-	Tx2+	G	Tx4-	Tx4+	G	SB	SB	SB	SB	SB	G	Rx3+	Rx3-	G	Rx1+	Rx1-	G
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19

microQSFP Pinout

38	37	36	35	34	33	32	31	30	29	28	27	26	25	24	23	22	21	20
G	Tx1-	Tx1+	G	Tx3-	Tx3+	G	SB	SB	SB	SB	SB	G	Rx4+	Rx4-	G	Rx2+	Rx2-	G
G	Tx5-	Tx5+	G	Tx7-	Tx7+	G	SB	SB	SB	SB	SB	G	Rx8+	Rx8-	G	Rx6+	Rx6-	G
76	75	74	73	72	71	70	69	68	67	66	65	64	63	62	61	60	59	58

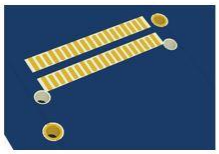
39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57
G	Tx6-	Tx6+	G	Tx8-	Tx8+	G	SB	SB	SB	SB	SB	G	Rx7+	Rx7-	G	Rx5+	Rx5-	G
G	Tx2-	Tx2+	G	Tx4-	Tx4+	G	SB	SB	SB	SB	SB	G	Rx3+	Rx3-	G	Rx1+	Rx1-	G
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19

QSFP-DD Pinout

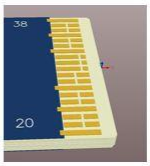
3.4 PCB Considerations Due to Connector Definition

An additional consideration for the three I/O port designs is that the different connector structures also impact the PCB layout and the resultant routing on both the host board in the switch and the card edge PCB in the module that mates to the connector. The two-row structure of the microQSFP and OSFP host allows lower cost and improved performance of these routing escapes, while the QSFP-DD additional recessed two rows do create incremental PCB complexity and performance implications. The PCB electrical performance effects are captured in the mated compliance boards and are included in the data shown above. The below images show the comparative challenges of routing the host PCB and module card edge PCB.

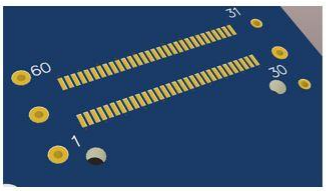
Relative comparison of host footprint and module card edge PCB
Complexity to route out signals



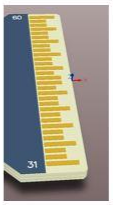
microQSFP host footprint



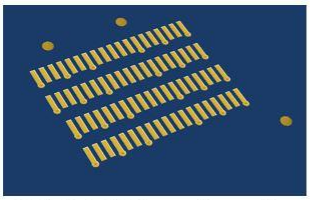
microQSFP card edge PCB



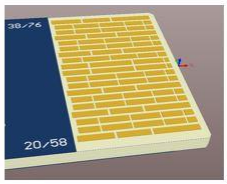
OSFP host footprint



OSFP card edge PCB



QSFP-DD host footprint

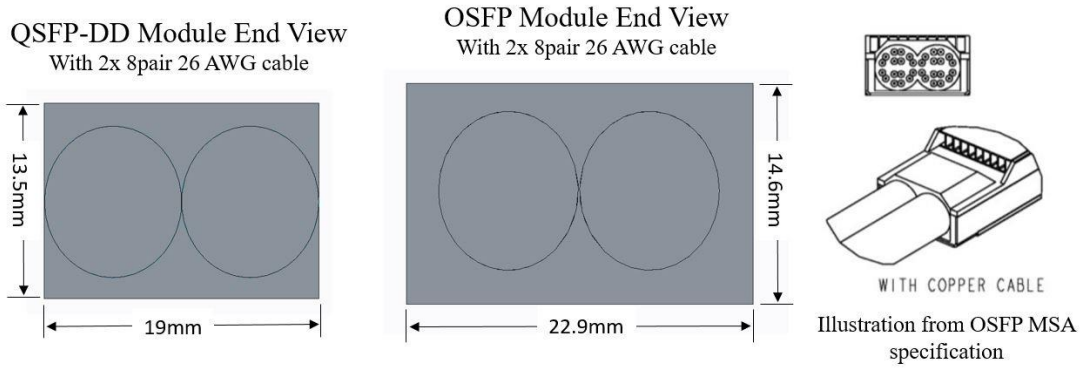


QSFP-DD card edge PCB

microQSFP and OSFP host footprints show the two row style, QSFP-DD host footprint shows the four row style. microQSFP and OSFP module card edge show the single row (top and bottom) mating array, QSFP-DD module card edge shows the two row (top and bottom) mating array. Since I/O connectors are usually ganged side by side, there is limited room on the sides of each footprint to expand routing to the side

3.5 Direct Attach Copper Reach Considerations

As was mentioned in Section 1.2, one of the types of plug implementation is passive direct attach copper cables where the reach of the copper cable is critically important to end users due to copper cables' lower cost per connection than optical modules. The key to cable reach is the wire gauge (diameter) of the cable being used because a larger diameter provides lower loss which enables longer reach. Copper cables are specified at 5m in the IEEE specifications operating at 25 Gbps per channel and are currently targeted to reach 3m at 50 Gbps PAM4 per channel; both reaches are based on assumption of 26 AWG (American Wire Gauge) cable. There are two primary constraints to fitting larger wire gauges in these form factors. The first is fitting the diameter of the wire into the height and width of the front of the module, and the second is terminating the wires to the module PCB inside the module while still achieving the necessary signal integrity (due to packaging challenges such as cross-talk and return loss). microQSFP cables have demonstrated a method of packaging four channels of 26 AWG cable into the form factor and the OSFP form factor has also demonstrated eight channels of 26 AWG cable, however the height and width of the QSFP-DD module front and module PCB area is smaller and does not easily accommodate 26 AWG. The result is that microQSFP and OSFP will always enable longer reach passive cable assemblies than QSFP-DD. This is important to end users because it means that for longer reaches they must use more expensive optical modules.



2x 8 pair 26 AWG cable bundles shown against the front surface area of QSFP-DD and OSFP modules. The QSFP-DD module will be more challenged on copper cable reach than microQSFP and OSFP due to available volume

4.0 Thermal Comparison

4.1 Background on Module Thermal Management

The pluggable modules that will be plugged into the microQSFP, OSFP and QSFP-DD connector/cages will include significant increases in componentry and power dissipation from today's 100 Gbps modules due to higher speeds, different modulation and higher density. Looking back in time, 40 Gbps QSFP modules all operated within a 3.5W power envelope. When the industry transitioned to 100 Gbps this power envelope increased significantly, and in some cases power dissipation exceeded 5W for longer reach optical pluggable modules. This required improvement and optimization of the QSFP cages and

riding heat sinks to achieve improved transfer of heat away from the module and into the air stream cooling the module and the rest of the heat dissipating equipment in the switch. One of the challenges that equipment developers face today is the balancing act of maximizing the number of pluggable modules that can be fit into a 1RU face plate (i.e. aggregate bandwidth) with the need to allocate a significant amount of faceplate area to remain open for airflow perforations to cool both the modules and the active electronics that are on the switch printed circuit board. It is well known in the industry that power dissipation is increasing significantly with every new generation of equipment due to the greater bandwidth density being delivered. The importance of equipment thermal design now impacts virtually all components in some way, including I/O ports.

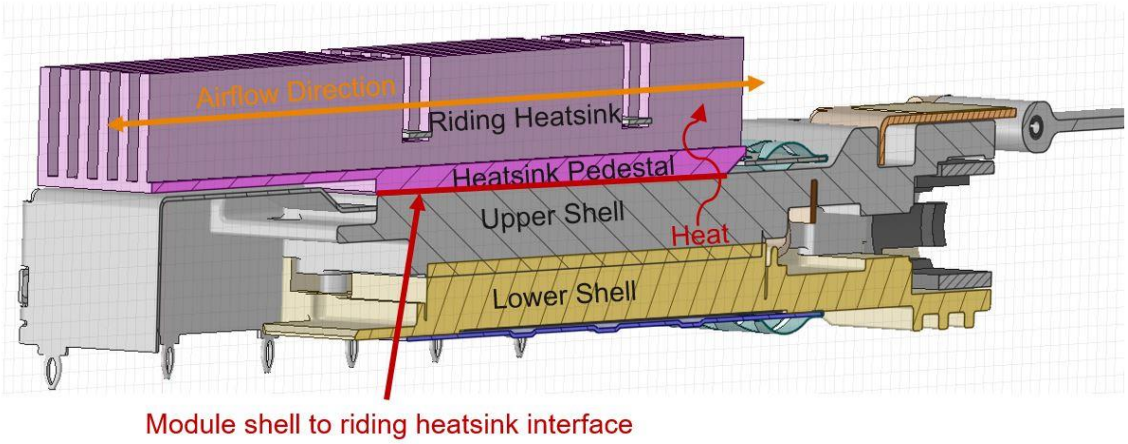
The challenges of thermal management of the above mentioned 5W QSFP modules for 100 Gbps focused the I/O port suppliers on the issue of module thermal management and drove thermal design optimization at connector industry leaders. Leading connector companies including TE have made a significant investment in thermal design, modeling and testing expertise to meet this growing need. The technology utilized in a riding heat sink interface has been improved, and alternative methods to better handle module power dissipation have been developed.

Because of the equipment suppliers' need to be able to "plug and play" various optical pluggable modules in their equipment, the riding heat sink was developed as a functional way to manage the thermal energy in the pluggable module. This module-to-riding heat sink interface is a sliding surface to allow insertion and removal of pluggable modules. The performance of the actual thermal transfer from the module to the heat sink is dependent on many characteristics including module and heat sink flatness and roughness along with the heat sink pressure against the module. The pressure is managed by a spring in the I/O port assembly. Of course, the heat sink design itself is critical too (fin geometry, etc.). Due to cost pressures in the industry, there are practical limits to the flatness, roughness and the pressure that can be specified as well as the resultant degradation of flatness and roughness due to wear over a lifetime of module insertion and removal. This "wear" on the heat sink is especially concerning when an earlier generation module (such as 100 Gbps) with a rougher surface is inserted against the newly specified smooth/flat heat sink. This may degrade the performance of the new I/O port significantly. Further investigation in this area is in progress. The bottom line is that riding heat sinks may have a limit to the increased performance that can be delivered given the pluggable module environment and usage trends.

4.2 Module Heat Removal – Riding Heat Sink vs Integrated Sinks

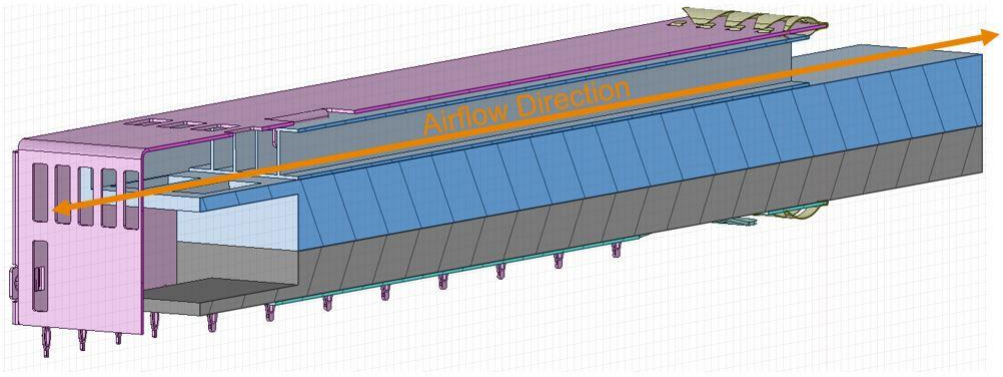
When considering the overall thermal management design of a pluggable module into a piece of networking equipment, a significant portion of the overall thermal resistance between a module's power dissipating components and the air stream passing by the heat

sink is dictated by the module to heat sink interface and its associated thermal resistance.



Cross section view of module and riding heat sink showing thermal resistance at the interface

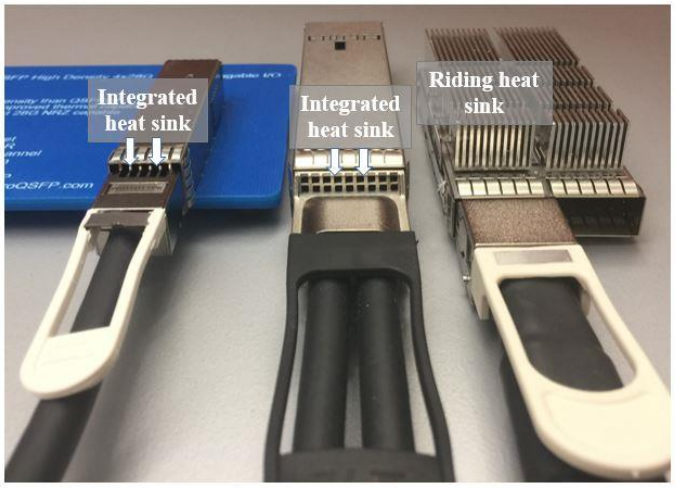
microQSFP and then OSFP attacked this problem by integrating the heat sink directly into the module, eliminating the thermal resistance of a riding heat sink interface from the overall design. This has enabled a whole new level of thermal performance.

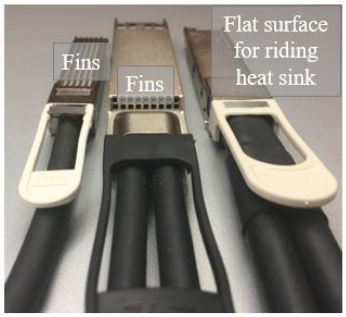


Cross section view of module with integrated heat sink eliminates interface thermal resistance and shows additional air flow through the module/cage

An important added benefit of this integrated fin type of design is that as air flows through the cage and module integrated heat sink, that air is now also available to the equipment designer to cool down-stream electronics. By gaining airflow from the I/O port, the equipment designer can revisit the balancing act of module face plate allocation vs. airflow hole perforation. This enables the designer to address the growing thermal management challenges or consider adding more modules for additional bandwidth.

Comparative view of microQSFP integrated heat sink,
OSFP integrated heat sink, and QSFP-DD riding heat sink

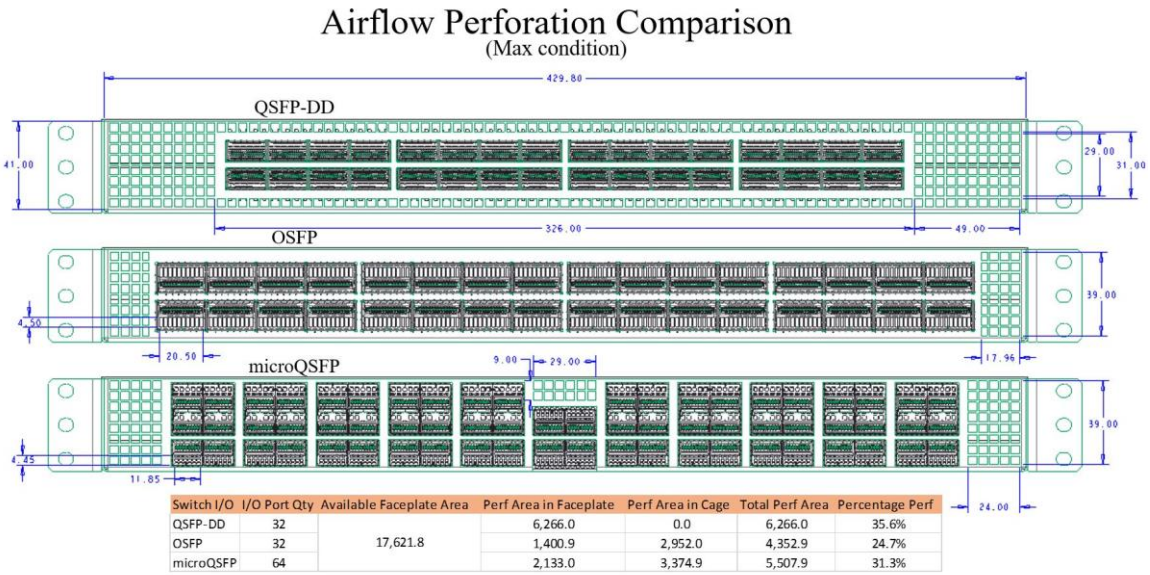




microQSFP form factor in cage | OSFP form factor in cage | QSFP-DD form factor in cage

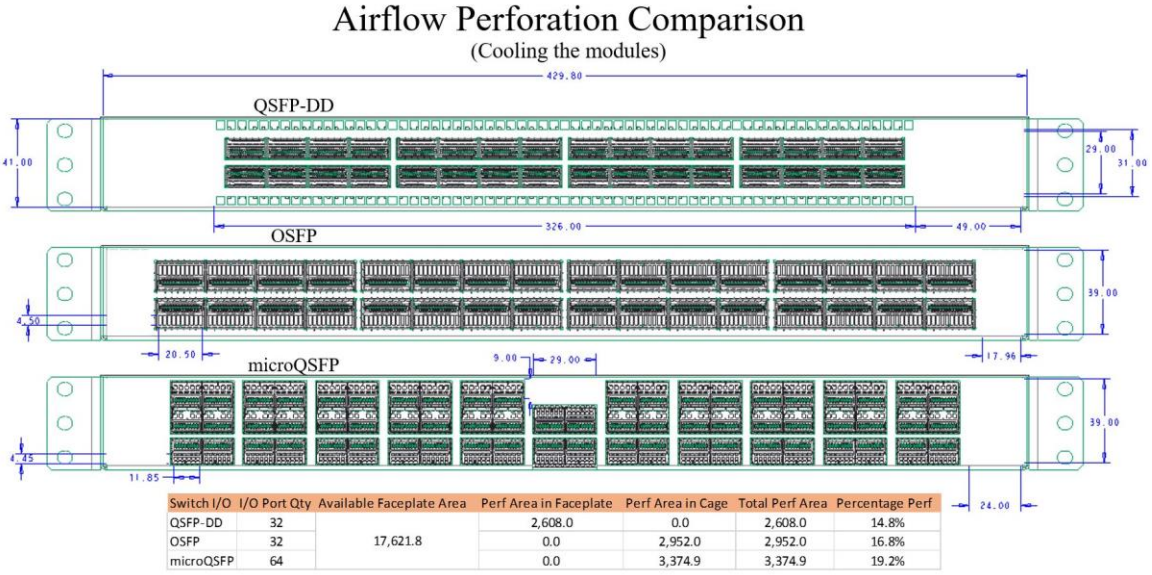
microQSFP form factor | OSFP form factor | QSFP-DD form factor

The following graphic shows the calculated face plate area available for cooling down-stream electronics in a 1RU enclosure for 256 channels of IO density, comparing QSFP-DD, OSFP and microQSFP. Unfortunately, this ignores the reality that air will bypass the modules and take the path of least resistance. Modeling and testing has shown that 12W modules installed in this implementation will not be cooled adequately. In the second set of images, the design focuses the airflow over the modules and then the perforation calculation provides a very different result. When we provide the necessary module airflow, then we see that the microQSFP provides the greatest airflow, followed by OSFP, and then QSFP-DD with the most restrictive airflow. An actual equipment design will probably be somewhere between the two examples, but the point is made that with 12W modules, the air must be forced over/through the modules or reliable performance will not be obtained due to excessive temperatures.



Switch I/O	I/O Port Qty	Available Faceplate Area	Perf Area in Faceplate	Perf Area in Cage	Total Perf Area	Percentage Perf
QSFP-DD	32	17,621.8	6,266.0	0.0	6,266.0	35.6%
OSFP	32		1,400.9	2,952.0	4,352.9	24.7%
microQSFP	64		2,133.0	3,374.9	5,507.9	31.3%

Comparative view of QSFP-DD, OSFP, and microQSFP 256 channel implementation showing the maximum perforation condition



Switch I/O	I/O Port Qty	Available Faceplate Area	Perf Area in Faceplate	Perf Area in Cage	Total Perf Area	Percentage Perf
QSFP-DD	32	17,621.8	2,608.0	0.0	2,608.0	14.8%
OSFP	32		0.0	2,952.0	2,952.0	16.8%
microQSFP	64		0.0	3,374.9	3,374.9	19.2%

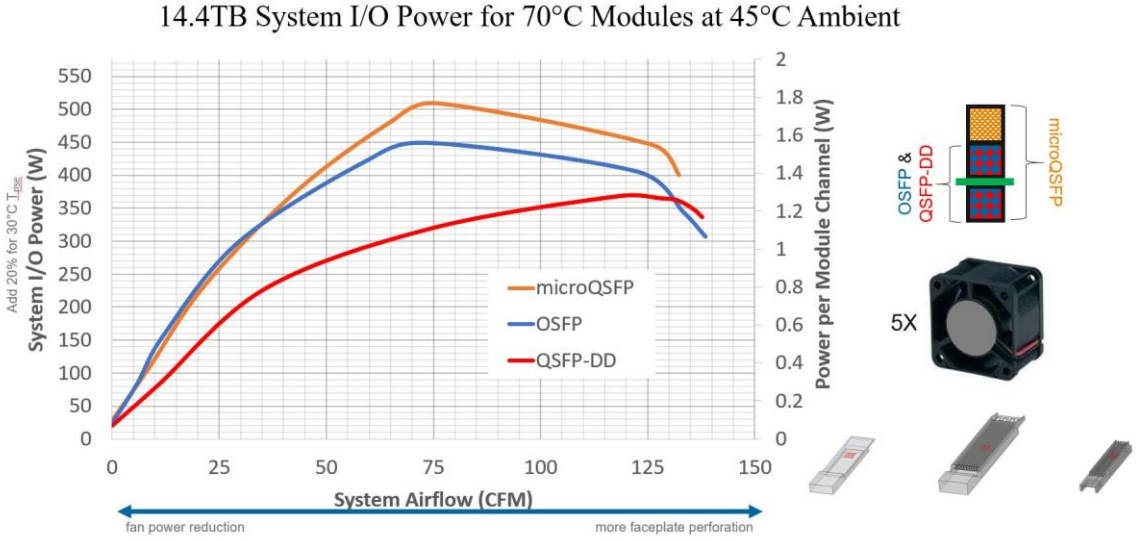
Comparative view of QSFP-DD, OSFP, and microQSFP 256 channel implementation showing the perforation condition necessary for maximum module cooling

As mentioned above, the increase in module power has resulted in a renewed focus on thermal design by TE as well as development of thermal modeling tools and validation test procedures and environments. This enables TE to rapidly predict a design's thermal performance for a customer's given equipment enclosure design and thermal sources as well as the ability to consider variations in module and heat sink designs. These new tools, experience, and expertise have enabled continued optimization and improved performance to be brought forward for pluggable connector/cage environments whether

they are integrated heat sinks like microQSFP and OSFP or riding heat sink such as QSFP-DD.

4.3 Simulated Comparative Thermal Performance

Below, we provide simulated performance for an equipment enclosure with either 72 microQSFP modules, 36 OSFP modules and 36 QSFP-DD modules. Unlike the above analysis, in this case the enclosure face plate perforation was optimized for each form factor to provide the best-case environment to enable a side by side by side comparative analysis. This analysis determines the comparative performance of the three form factors for this enclosure/fan environment. In all cases the simulation was done with the worst-case module monitoring point (unique for each form factor) held to a maximum case temperature of 70-degree C and the module power and airflow were varied over a full range to determine the performance envelope. The results are plotted in total system I/O power and power per module channel to account for the fact that microQSFP has four channels and OSFP and QSFP-DD have 8 channels.



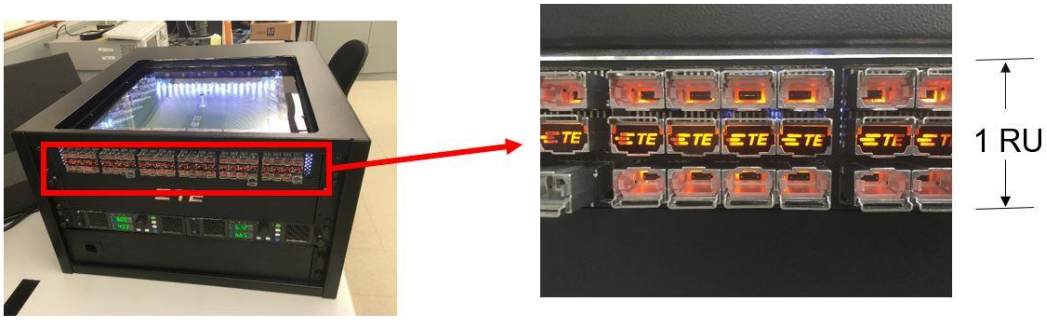
Comparative power per module channel for microQSFP, OSFP and QSFP-DD 288 channel implementations

It is shown that microQSFP provides the highest thermal efficiency of the three, with OSFP offering similar performance and QSFP-DD with its riding heat sink providing the least performance. Most importantly, it is established that all three designs have the improved performance that is required for 400 Gbps pluggable modules.

4.4 Measured Thermal Results

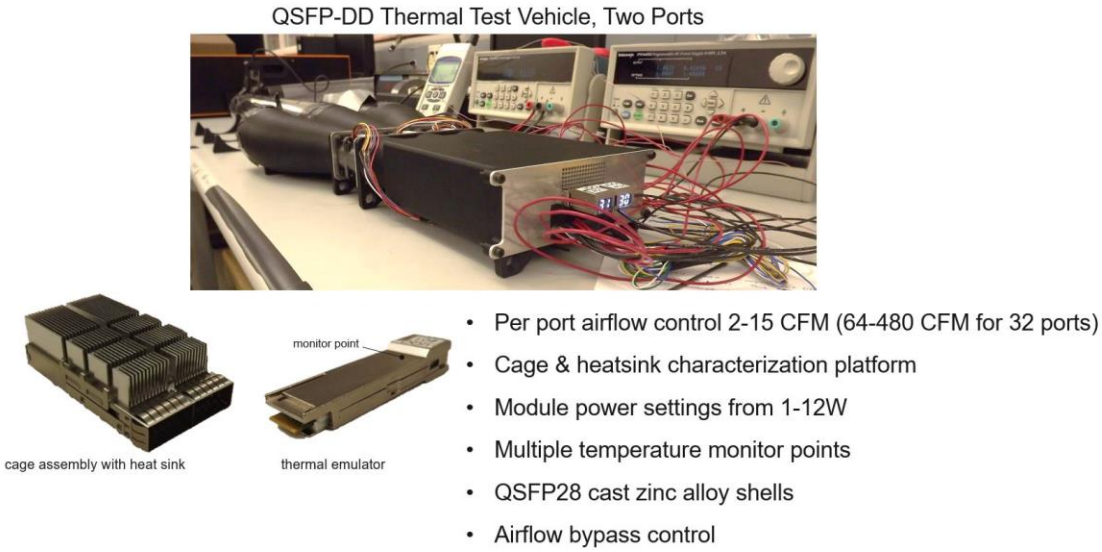
TE has made measured thermal tests on both microQSFP and QSFP-DD form factors. Module results correlate well to the simulations that have been performed. Test vehicles used at TE for thermal measurements have varied from single and dual port test chambers up to full 1RU enclosures. For microQSFP we have validated a thermal capacity of

greater than 1.9W per channel (7W per four channel module), and for QSFP-DD we have scaled measurement data to 1.5W per channel (12W per eight channel module) when used in a belly-to-belly implementation. OSFP has been confirmed to greater than 1.9W per channel (15W per eight channel module) based on modeling and correlation to microQSFP measurements due to the similar cooling mechanism (integrated module fins).



1RU enclosure with 72 ports of microQSFP thermal modules operating at 7.5W per module

microQSFP 72 port 1RU enclosure operating at 7.5W per module



QSFP-DD two port thermal test vehicle

Through TE's simulation and measurement work we have investigated the key variables required to optimize the cooling system whether it is a riding heat sink or integrated module heat sink. Variables such as heat sink pressure (for riding heat sinks), module and heat sink flatness and roughness, fin geometry and fin materials are all going to be critical determinants in the thermal management of these 200 Gbps and 400 Gbps modules and system equipment enclosures. In fact, these variables are so important that it may be necessary to optimize them for each different equipment design that is brought to market at these module power levels.

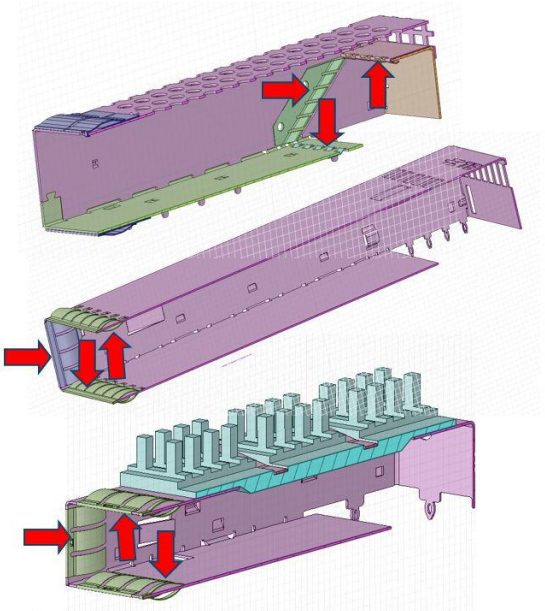
5.0 EMI Considerations

Another aspect of pluggable modules that should not be left out of the evaluation for these new form factors is that of EMI. Pluggable module ports create huge holes in high performance equipment and an undesired consequence is the potential for electromagnetic radiations that exceed those allowed by regulations. I/O ports and pluggable modules have included features since their beginning to ensure these ports do not exceed allowed limits set by regulations. microQSFP, OSFP and QSFP-DD continue this practice but in slightly different ways. QSFP-DD was specifically established to provide backwards compatibility to legacy QSFP modules and therefore is limited in the ability to innovate the EMI design because it must accommodate legacy features originally designed for operation years ago at lower data rates. Since microQSFP is a new form factor, it was not burdened by legacy features and could take a new approach, offering approximately 10dB improved performance over other pluggable form factors. See measurement data below. OSFP EMI utilizes a more conventional EMI containment approach and offers performance on par with QSFP-DD.

microQSFP integrates EMI features at the back of the cage to allow easy airflow

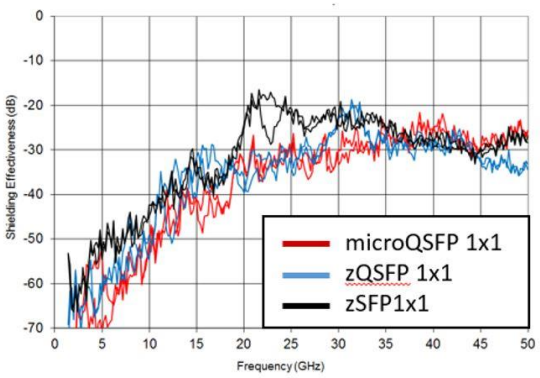
OSFP integrates EMI at the front of the cage similar to QSFP

QSFP-DD integrates EMI at the front of the cage so that legacy modules can be used



Cut-away views of the form factor cages

Comparative view of EMI containment methodologies



Comparative EMI measurements of microQSFP and standard SFP28 and QSFP28

microQSFP EMI performance compared to QSFP28 and SFP28 demonstrating the improvements possible when legacy features are not required

6.0 Summary

This analysis used simulation tools to predict performance of electrical signal integrity, thermal and EMI performance of three new form factors being developed for next generation higher density 100 Gbps, 200 Gbps and 400 Gbps modules. These tools have enabled design optimizations and tradeoffs for performance and cost evaluation. They have also provided an early input for equipment developers to consider the implications of the three form factors and the effects on equipment design. Subsequent measurement of product has established correlation to the performance predicted by the simulation tolls.

Assessment of the three, next generation I/O ports:

microQSFP form factor:

Pros: The microQSFP I/O port has been shown to have excellent thermal performance of greater than 1.9W per channel (7.5W per four-channel module) via integrated cooling fins, good signal integrity with a total modeled ICN of 1.6mV, ability to accommodate 26 AWG copper cable and excellent EMI performance. The form factor has demonstrated the ability to provide up to 288 channels in a 1RU enclosure.

Cons: It is a new to market form factor that can only provide backwards compatibility to SFP modules with the use of an adapter module.

OSFP form factor:

Pros: The OSFP I/O port also uses integrated cooling fins to provide excellent thermal performance of greater than 1.9W per channel (15W per eight-channel module), excellent signal integrity with a total modeled ICN of 1.0mV, ability to support 26 AWG copper cable and good EMI performance. It can support up to 288 channels in a 1RU enclosure.

Cons: It is a new to market form factor that can only provide backwards compatibility to QSFP modules with the use of a QSFP-to-OSFP adapter.

QSFP-DD form factor:

Pros: The QSFP-DD I/O port provides direct backwards compatibility to QSFP modules. QSFP-DD can support densities of up to 288 channels in a 1RU enclosure.

Cons: QSFP-DD's smaller form factor and riding heat sink thermal management solution limit thermal performance to 1.5W per channel (12W per eight-channel module) and the connector and PCB complexity results in a total modeled ICN of 2.7mV. The reduced module volume makes it difficult to accommodate larger wire gauge and will result in shorter cable assembly reaches.

7.0 Implications to Equipment Developers:

It has been demonstrated that there are three pluggable form factor solutions that achieve the overall bandwidth density of at least 256 electrical channels in a 1RU form factor. Some have more operating margin (or performance "up-side") and others have less.

Direct backwards compatibility can be achieved with QSFP-DD, but does not come without performance "costs" or burdens. Some are easily observable differences and some are small, incremental differences but small differences are important in these next generation systems. These "costs" may include the use of extra re-timer chips to extend channels or incorporation of internal cables (both of which degrade thermal performance), use of higher performance fans, or fewer modules per 1RU equipment to allow more airflow. In addition, it might be simply margin for operation at broader temperature ranges.

In cases where backwards compatibility is not required or is limited, the extra performance margin of the microQSFP and OSFP I/O ports can be leveraged for this generation and future generations to provide an optimized system while only burdening the backwards compatible ports (that have adapters). In those cases that do need some backwards capability, an adapter seems like a logical choice to get good performance for the legacy use case while delivering the best possible performance where it is needed when looking forward.

The three I/O port options discussed in this analysis provide different performance and use-case trade-offs for equipment developers. The key issue is to understand the impact of the trade-offs and what design and performance aspects are most important to your customers and end users.

8.0 Contributors/Sources

Eric Briant, TE
Alan Bucher, TE
Bruce Champion, TE
David Herring, TE
Richard Long, TE
Joel Myers, TE
Megha Shanbhag, TE
Nikhil Shankar, TE
Alex Sharf, TE
Linda Shields, TE
Kevin Stauffer, TE
Mike Tryson, TE
Alan Weckel, 650 Group
microQSFP MSA, www.microQSFP.com
OSFP MSA, www.osfpmsa.org
QSFP-DD MSA, www.qsfp-dd.com