Design Guide for Photonic Architecture

DRAFT DOCUMENT

Revision 0.5

January 2013

Revision History

Date	Revision Number	Modifications
December 2012	0.1	First Draft
January 2013	0.5	First posting to OCP Portal

Disclaimers

Table of Contents

1. Scope

This design guide is intended for adopters of the OCP Open Rack. The design guide provides an overview for implementing an intra-rack optical interconnect scheme that utilizes a New Photonic Connector (NPC) and embedded optical modules to deliver next generation data rate scaling and system architectural benefits which are described in more detail below.

2. Overview

When data center design and hardware design move in concert, they can improve efficiency and reduce the overall system power consumption. The Open Compute Project is focused on delivering this benefit by providing a set of technologies that reduces energy consumption, reduces cost, increases reliability, expands choice in the marketplace, simplifies operations for system implementers and reduces the maintenance overhead.

A Photonically Enabled Architecture envisions a system where the bandwidth density, line rate scalability and easier cable routing provides value to the data center built using Open Compute technologies. The photonically enabled architecture will provide the data path between the compute, storage, network, and IO functionality in the rack and throughout the data center.

2.1 Benefits of Optical Communication Links

Optical communication links have several advantages over the standard electrical signaling technology over copper cables and traces. The most significant advantage that optical communication links offer over electrical links is that the distance or reach of an optical link will be longer than an electrical link at a given data rate. This has several practical benefits; first, the optical links support longer communication paths in the data center, providing a path forward for future generations of traditional rack switch network topology common in data centers. Additionally, the fiber infrastructure is scalable with data rate, so a data center could be configured, built out and used for many component generations.

Optical cables are also thinner and lighter than equivalent bandwidth copper cables. A typical optical fiber has a diameter of 120 microns and is made from light-weight glass. Single fibers deployed in typical data center applications can carry >40Gbps over distances of hundreds of meters depending upon the optical technology used. Increasing the data rate per copper channel greatly shortens the reach, a physical property of the skin effect, while increasing the number of parallel connections per cable decreases the bend radius and increases the weight of the cable. The additional weight of copper cables can also become an issue for certain dense applications; there have been instances of data center deployments which required floor reinforcement in order to support the weight of the required copper cabling. Not only is the weight of the fiber much less than a copper cable, but the weight of the fiber is small compared to the fiber jacket, which defines the physical size of the fiber cable. Therefore increasing the number of fibers within the cable has little impact on the cable weight and bend radius.

Optical communication links are already widely deployed in data center applications; however the promise of advanced optical technologies and the impact that they can have on the rack based architecture are significant. The New Photonic Connector concepts described here allow rack designers the ability to integrate optics into the chassis in potentially new and innovative ways.

2.2 Benefits of the Data Center Connector

Optical communication links are currently widely deployed in the data center; however there are several different types of optical connectors which are used. The most common type of for shorter reach applications is called MTP/MPO (Multi-fiber Termination Push-On/Multiple fiber Push-On), an example of which is shown in Figure 2.1 below.

Figure 2.1 – MTP/MPO Connector

The MTP/MPO connector concept was first designed over 25 years ago. The connector was designed with what was then state of the art material and manufacturing techniques; however a mated MTP/MPO connection has over 20 moving parts which leads to a cost burden on the end user. It was also designed for the telecom market and thus is not optimized data center applications.

The New Photonic Connector, described in this Design Guide, is designed specifically for data center applications using contemporary materials and manufacturing techniques. Consequently it has fewer components, is more reliable and will be cheaper than an MTP/MPO connector. The New Photonic Connector is also designed to be smaller and more resilient to dust and other contaminants than the existing solution.

Figures 2.2, 2.3 and 2.4, included below, show different views of the constituent pieces of the New Photonic Connector solution for these data center applications. The concept relies upon a Receptacle which is attached to the panel of the rack mounted data center unit; the form factor of the compute platform is not critical since the New Photonic Connector concept will work with bladed or standard compute platform configurations. This Receptacle, shown below in Figure 2.2, provides a mechanical reference for the two mating ends of a New Photonic Connector mated pair, enabling the low loss optical connection. The New Photonic Connector mated pair is created with the optical plug from the internal fiber jumper and the optical plug from the external optical cable. A drawing of the external optical cable New Photonic Connector plug, which is 10.8mm by 4.7mm, is shown in Figure 2.3 below; this small, dense connector concept can support up to 4 rows of 16 fibers each for a total scalable solution of 64 optical fibers. This would result in a fiber density of greater than one fiber per square millimeter. Figure 2.3 also illuminates the simple nature of the New Photonic Connector concept with the limited number of moving parts; here there are two significant parts shown, the fiber ferrule (shown in yellow) and the ferrule housing (shown in black). The mechanical alignment tolerance is provided by the mechanical alignment pins, which are also shown in this figure. Also apparent is some of the design methodology to reduce the overall footprint of this connector, including the asymmetric nature of the latching mechanism. This figure shows a design with 24 fibers arranged in 3 rows of 8 each, which are represented by the lenses in the yellow fiber ferrule portion of the design. Figure 2.4 shows the mated New Photonic Connector cable pair including the Receptacle; this figure further illuminates the latching mechanism on the external New Photonic Connector fiber which provides positive latching in a robust manner for data center environments.

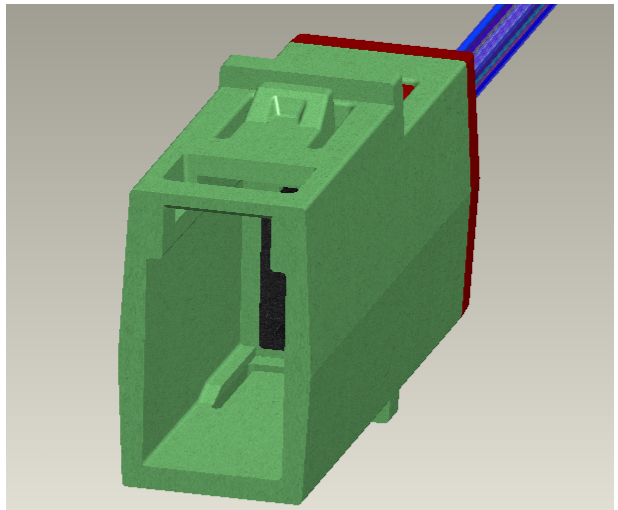


Figure 2.2: New Photonic Connector Receptacle - This Receptacle will be attached to the external face of the compute platform in the data center environment. This figure shows the external facing interface of the New Photonic Connector receptacle, with the internal jumper cable side already inserted.

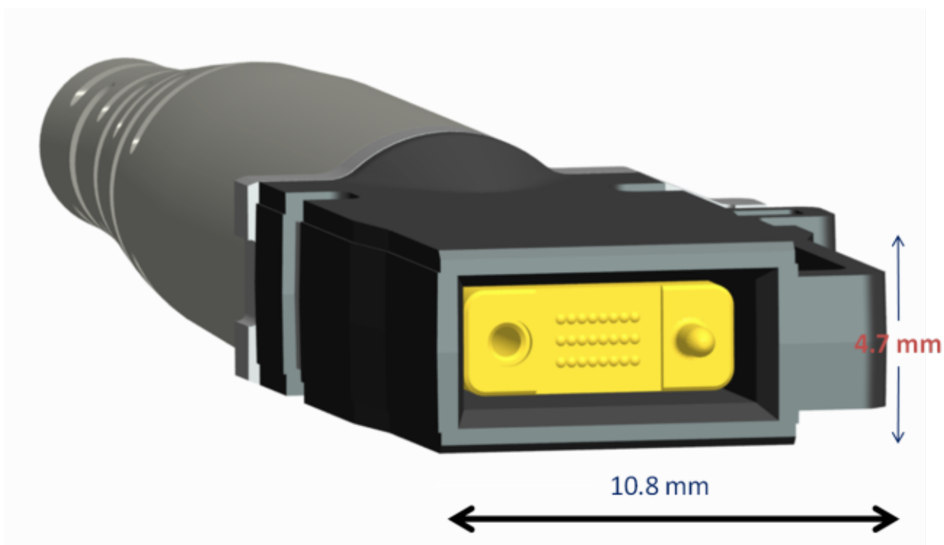


Figure 2.3: New Photonic Connector Plug – The New Photonic Connector Plug shown in this image is the termination of an external fiber cable. This view shows the mating surface of a 24 fiber connection arranged in 3 rows of 8; it further illustrates the small form factor and simplified design concept which has resulted in a small, dense, low cost connector.

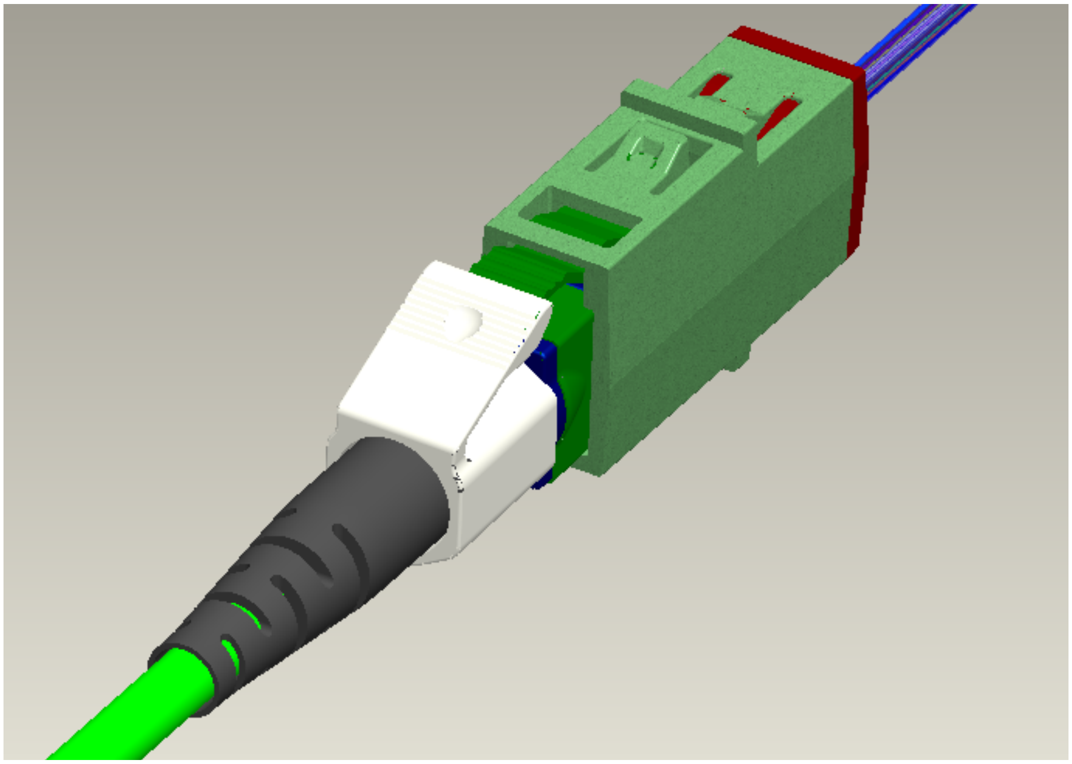


Figure 2.4 Mated Receptacles and Plug - Image shows further detail of the external fiber cable terminated with the New Photonic Connector plug, including the latching mechanism which provides positive latching and robust attachment of the cable in the data center environment.

3. Rack Interconnect Scheme

3.1 Common Building Blocks

The optical interconnect is a modular composition of following components:

Optical Module	Optical transceiver capable of both transmitting and receiving optical data streams. Connected through the jumper cable to the New Photonic Connector Receptacle
Jumper Cable	The optical fiber which connects the Optical Module to the New Photonic Connector Receptacle. Having a module optical connector on one end and the New Photonic Connector internal optical plug on the other.
New Photonic Connector Plug	The fiber termination of an optical fiber which fits into a New Photonic Connector Receptacle and mates to another New Photonic Connector Plug. The internal version meant to terminate the jumper cable supports occasional connection and low cost, high volume assembly. The external fiber version supports a robust, multiple latching and engagement mechanism meant to be accessed by technicians in a standard data center environment.
Patch Panel with New Photonic Connectors	A platform component integrated into data center rack systems which provides fiber connections between New Photonic Connector Receptacles in a predetermined manner. The fiber connections provide ‘hard-wired’ connections between physical fiber locations located at pre-specified New Photonic Connector fiber ferrules, allowing the rack based system designer to enable a low cost, low latency within rack connection scheme
Photonically Enabled Architecture	A general terminology of the network topology that is technically feasible because of the advanced bandwidth density and scalability provided by photonics. The Photonics Enabled Architecture is applicable to both Ethernet and PCIe standards.
Mezzanine	A small board where lower speed electrical

Card	signals, either PCIe or Ethernet, are aggregated prior to transmission via the optical module. The mezzanine card also performs additional signaling conditioning or switching functions and route signals to and from the correct destinations.
Nodes or IT Modules	Providing the compute server, storage, or networking function for the data center.
ToR Switch	Top-of-Rack Switch providing communications and data transport interconnections for IT modules and other inter-chassis devices.

Table 3.1 - Definition of Terms used in this Design Guide

3.2 Interconnect Topology with a ToR Switch

One particular implementation of the Photonically Enable Architecture which is supported by the New Photonic Connector is shown below in Figure 3.1. In this implementation the New Photonic Connector cables are used to connect the compute systems arrayed throughout the rack to a Top of Rack switch. These intra-rack connections are currently made through electrical cabling, often using Ethernet signaling protocols at various line rates. The Photonically Enabled Architecture envisions a system where the bandwidth density, line rate scalability and easier cable routing provide value in this implementation model. One key feature of this architecture is that the line rate and optical technology are not dictated; rather the lowest cost technology which can support the bandwidth demands and provide the functionality required to support future high speed and dense applications can be deployed in this model consistent with the physical implementation model. This scalability of the architecture is a key value proposition of the design. Not only is the architecture scalable for data rate in the optical cable, but scalability of port count in each connection is also possible by altering the physical cabling and optical modules.

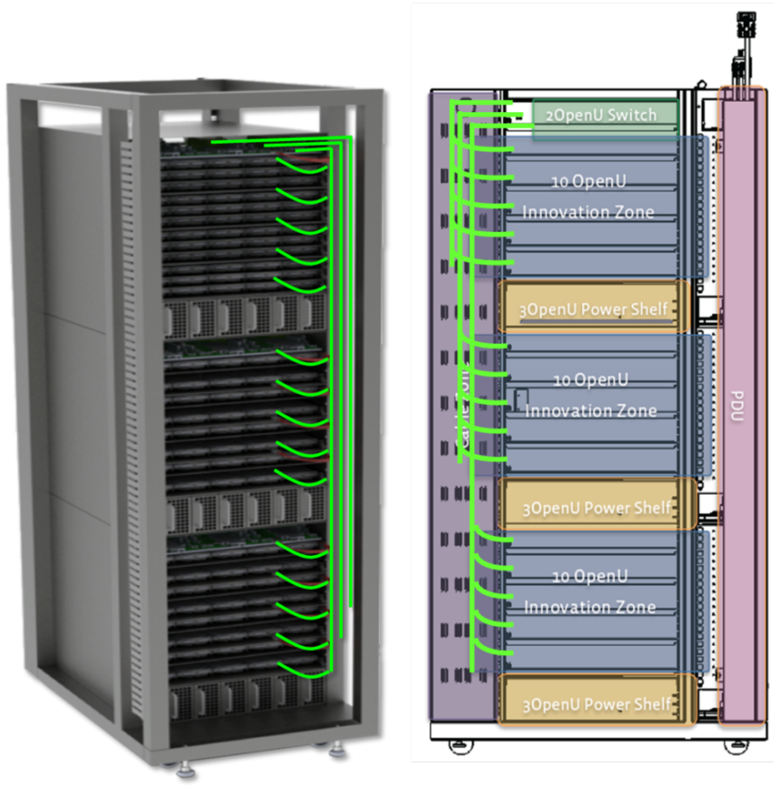


Figure 3.1: Open Rack with Optical Interconnect. In this architectural concept the green lines represent optical fiber cables terminated with the New Photonic Connector. They connect the various compute systems within the rack to the Top of Rack (TOR) switch. The optical fibers could contain up to 64 fibers and still support the described New Photonic Connector mechanical guidelines.

One key advantage of the optically enabled architecture is that it supports disaggregation in the rack based design of the various system functionality, which means separate and discrete portions of the system resources may be brought together. One approach to disaggregation is shown below in Figure 3.2; in the design shown here the New Photonic Connector optical cables are still connecting a computing platform to a Top of Rack switch, but the configuration of the components has been altered to allow for a more modular approach to system upgrade and serviceability. In this design the computing systems have been configured in 'trays' containing a single CPU die and the associated memory and control, while communication is aggregated between three of these trays through a Silicon Photonics module to a Top of Rack switch. The Top of Rack switch now communicates to the individual compute elements through a Network Interface Chip (NIC) while also supporting an array of Solid State Disk Drives (SSD's) and potentially additional computing hardware to support the networking interfaces. This approach would allow for the modular upgrade of the computing and memory infrastructure without burdening the user with the cost of upgrading the SSD infrastructure simultaneously provided the IO infrastructure remains constant. Other options for the disaggregated system architecture are of course also possible, potentially leading to the disaggregation of the memory system as well.

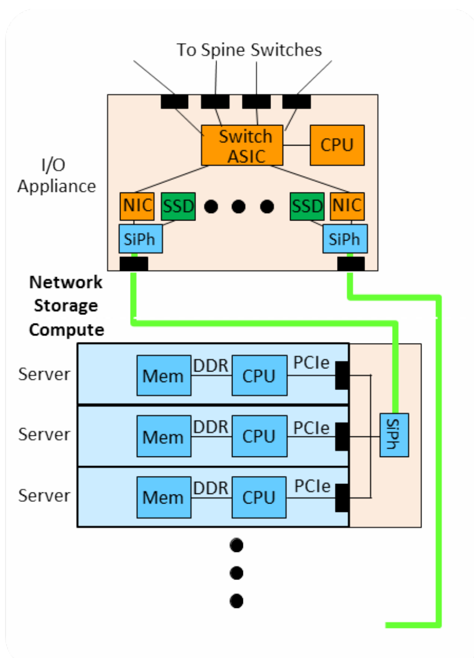


Figure 3-2: Disaggregated Photonic Architecture Topology with a ToR Switch. This design shows 3 compute trays connected through a single New Photonic Connector enabled optical cable to a Top of Rack (TOR) switch supporting Network Interface Chip (NIC) elements, Solid State Disk Drives (SSD's), Switching functionality and additional compute resources.

3.3 Interconnect Topology with Distributed Switch Functionality

The Photonically Enabled Architecture which is supported by the New Photonic Connector cable and connector concept can support several different types of architectures, each with specific advantages. One particular type of architecture, which also takes advantage of the functionality of another Intel component, an Intel Switch Chip, is shown in Figure 3.3, shown below. In this architecture the Intel Switch Chip is configured in such a way as to support both aggregation of data streams to reduce overall fiber and cabling burden as well as a distributed switching functionality.

The distributed switch functionality supports the modular architecture which was discussed in previous sections. This concept allows for a very granular approach to the deployment of resources throughout the data center infrastructure which supports greater resiliency through a smaller impact from a failure event. The concept also supports a more granular approach to upgradability and potentially could enable re-partitioning of the architecture in such a way that system resources can be better shared between different compute elements.

In Figure 3.3 an example is shown of 100Gbps links between compute systems and a remote storage node. Both PCIe and Ethernet networking protocols may be used in the same rack system, all enabled by the functionality of the Intel Switch Chip (or Device). It should be understood that the components in this vision could be swapped dynamically and asymmetrically so that improvements in bandwidth between particular nodes could be upgraded individually or new functionality could be incorporated as it becomes available.

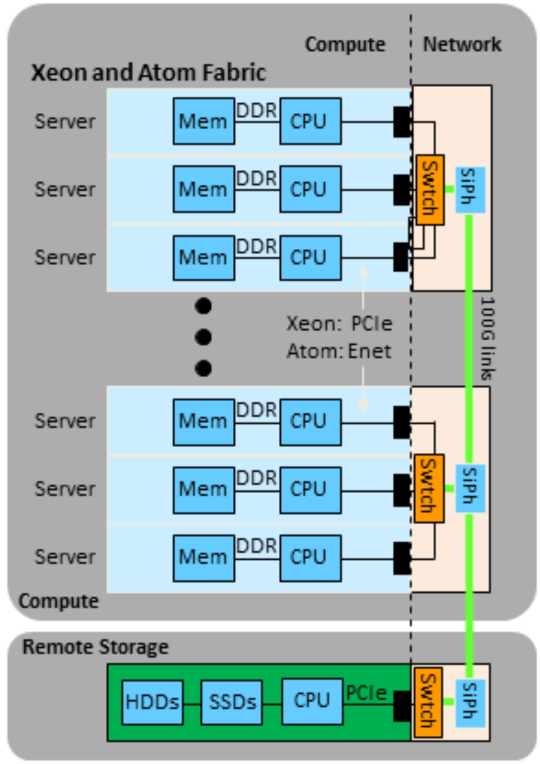


Figure 3.3: An example of a Photonically Enabled Architecture relying upon the New Photonic Connector concept, Silicon Photonics and the Intel Switch Chip (or Device). In this example the switching between the rack nodes is accomplished in a distributed manner through the use of these switch chips.

3.4 Interconnect to Compute and IO Nodes

The Distributed Switch functionality which was described in the previous section relies upon a collection of switching nodes interconnected through a high bandwidth, reduced form factor cable to reduce the impact of the cabling and interconnects on the system. Shown in Figure 3.4 below is one particular implementation of this scheme envisioned as part of this Photonically Enabled Architecture. In this case three compute ‘trays’ are connected with a lower speed electrical interconnect, based on either PCIe or Ethernet, to a mezzanine board where the network traffic is aggregated. In this aggregation step various signal conditioning or switching functions may be enabled in order to route the signals to and from the correct destinations. The non-local network traffic is then sent through a Silicon Photonics module through a New Photonic Connector cable solution to the final destination, which could consist of a ToR switch, a spline switch, or an adjacent node in a distributed switching architecture.

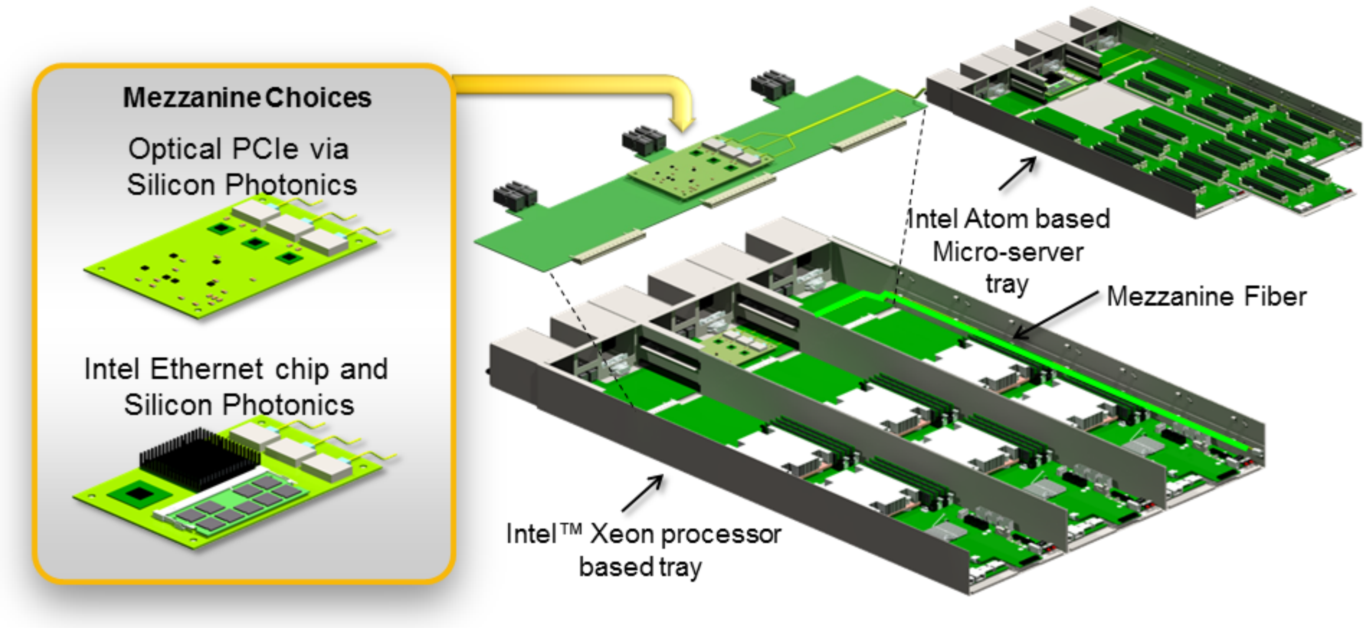


Figure 3.4: An example of a Photonically Enabled Architecture in an Open Compute mechanical arrangement using a Mezzanine Fiber – In this concept the New Photonic Connector cable concept is used to enable a reduced cable burden, and front panel access, through the use of silicon photonics modules and the modular architectural concepts which were discussed earlier.

4. Mechanical Considerations

The receptacle dimensions for a single receptacle are shown below in Figure 4.1. The receptacle is designed to fit into a faceplate that has a typical thickness range of 1.5-1.7 millimeters. The receptacle can be designed such that it can be plugged into the hole from either the front or the rear of the panel; that is the inside or the outside of the compute platform. It is designed to be plugged in only once and a special tool is required to remove it once it is seated in the faceplate cut-out.

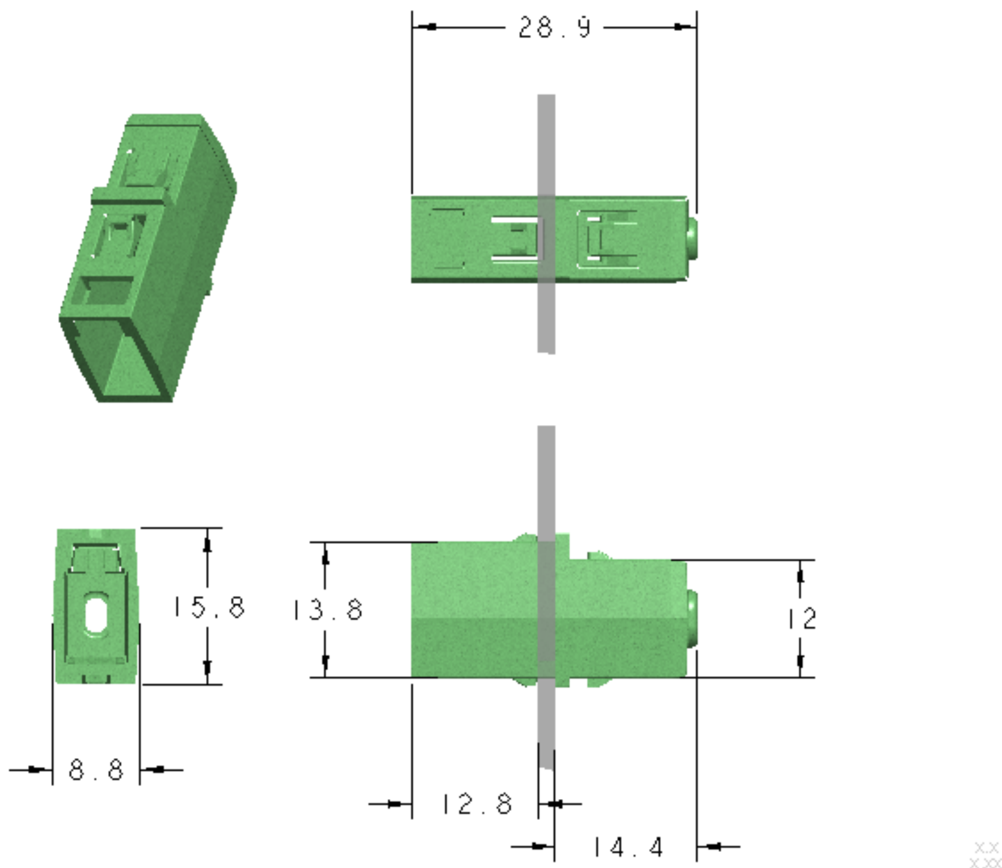
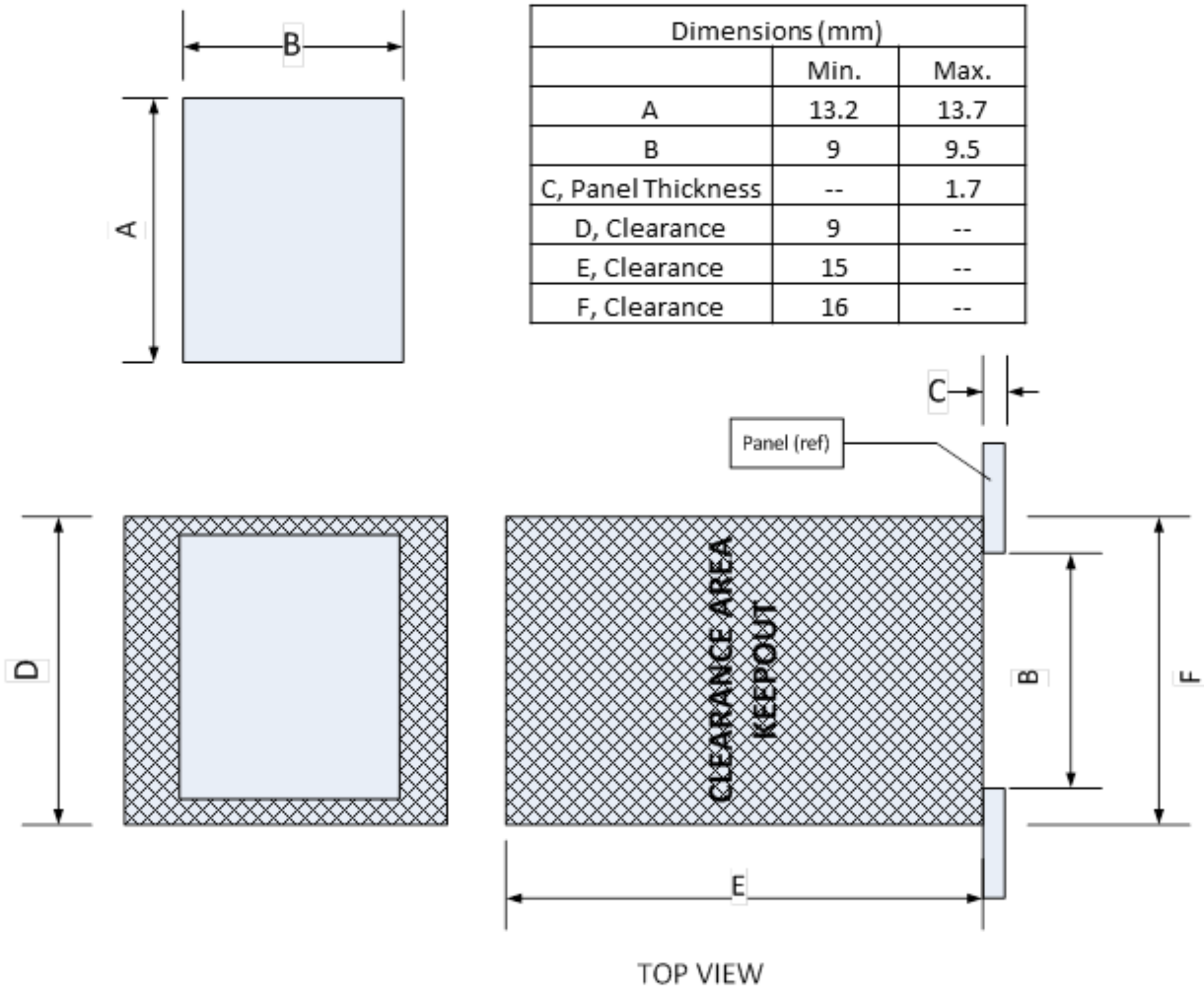


Figure 4.1 Single Receptacle Dimensions – This drawing of the New Photonic Connector Receptacle illustrates the outer dimensions of the Receptacle and some of the significant features of the design.

Figure 4.2 below shows the dimensions for the faceplate cutout to accept the receptacle. As mentioned above, the faceplate thickness must be less than 1.7mm. In addition to the faceplate cutout area there is a keep out area requirement for the space behind the faceplate which is required to accommodate the depth of the receptacle which protrudes into the rack. Clearance space behind the cutout is also needed for ingress or egress of the fibers from the jumper cable which connects into the receptacle.



Dimensions (mm)		
	Min.	Max.
A	13.2	13.7
B	9	9.5
C, Panel Thickness	--	1.7
D, Clearance	9	--
E, Clearance	15	--
F, Clearance	16	--

Figure 4.2: New Photonic Connector Receptacle Faceplate Cut Out Dimensions: The simple square cut out is all that is required to support the insertion of the New Photonic Connector Receptacle from the rear of the faceplate.

4.1 Multiple Receptacle Placement Specifications

The New Photonic Connector was designed to support deployments with multiple connectors in a single face plate to address those applications where density is a strong requirement, such as for Top of Rack switching applications. The clearance requirements surrounding each New Photonic Connector Receptacle cut-out ensure mechanical integrity as well as clearance for human access to each Receptacle in a dense application where occasional reconfigurability is required. The clearance dimensions to support the New Photonic Connector receptacle are shown below in Figure 4.3 below.

It is the responsibility of the faceplate (housing) designer to design and confirm the faceplate (housing) will support the New Photonic Connector hardware can meet the performance as specified in ANSI/TIA-568-C.3 Clause 5.2.1.2. The example given below in Figure 4.3

and Table 4.1 is shown for illustrative purposes and is included here for reference purposes only. It is a representative example, based upon the dimensions required to maintain structural integrity of the faceplate when using mild steel 16 gauge (~1.52 mm) or aluminum 14 gauge (~1.63 mm). Under these assumptions, the dimensions given are minimum values, the maximum values will be driven by the panel size and the targeted New Photonic Connector density target.

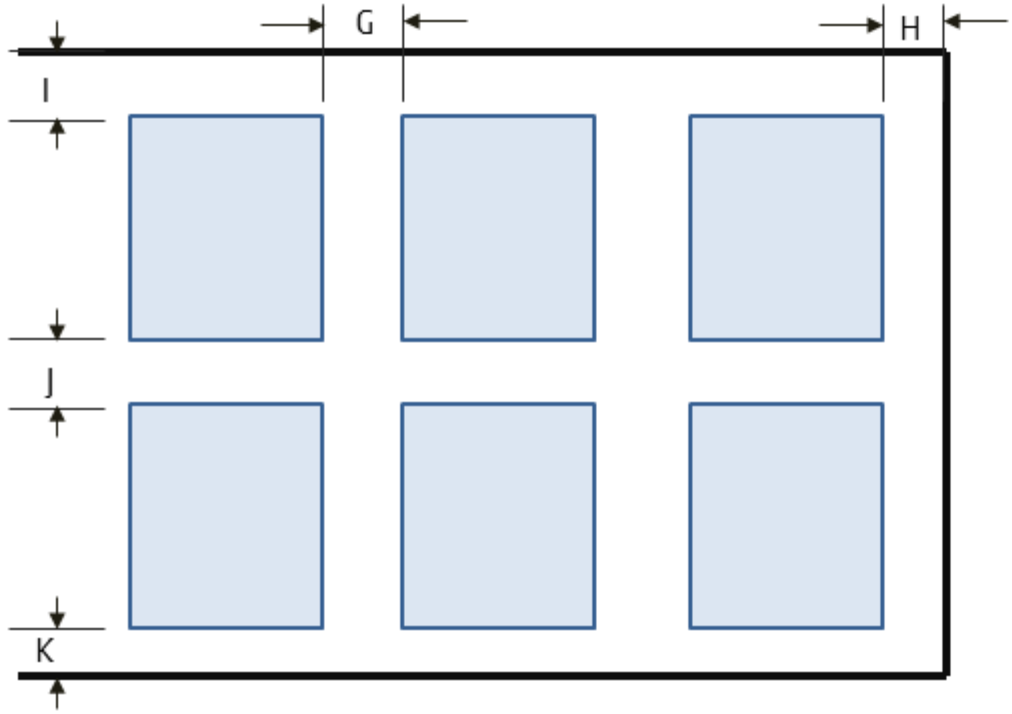


Figure 4.3: Multiple Receptacle Mechanical Spacing Requirements – These clearance requirements support very high bandwidth density requirements for TOR switch applications for instance.

Dimensions (mm)		
	Min.	Max.
G	5.0	--
H	5.0	--
I	8.0	1.7
J	8.0	--
J, Finger Clearance	13.5	--
K	8.0	--

Table 4.1: The minimum clearance values as shown in Figure 4.3 assuming that the face plate is 16 gauge (1.52mm) steel or 14 gauge Aluminum (1.63mm) and compliance with ANSI/TIA-568-C.3 Clause 5.2.1.2 is targeted.

5. New Photonic Connector Connector Layout Considerations

As has been previously described, the New Photonic Connector has been designed in order to support a wide array of data center applications, spanning the spectrum of single compute platforms to very high end High Performance Computing applications requiring the densest of deployments. In this section of the Design Guide examples of a typical configuration and maximum high density configuration are shown.

5.1 Typical Configuration

In a typical data center configuration a top of rack switch could have 48 ports which must be supported at the faceplate. One example of how such a configuration might be realized is shown in Figure 5.1 below. In this configuration the New Photonic Connector Receptacles are all arranged close together along a single horizontal axis; one of the strengths of the New Photonic Connector concept is that the placement of the New Photonic Connector Receptacles and optical connections can be configured in such a manner so as to support other requirements of the platform, such as air flow or cable routing. For instance these New Photonic Connector Receptacles could be placed in two rows of three at either edge of a full rack width single 'U' module to allow for more clearance for airflow through the center of the module. The only impact of this on the platform would be the configuration of the internal jumper cables and the externally routed New Photonic Connector enabled cables, otherwise the high speed data transfer network would perform identically.

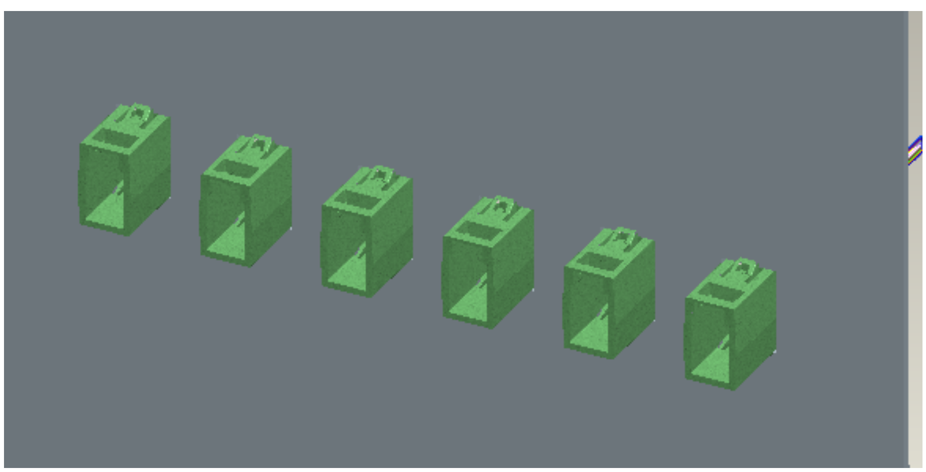


Figure 5.1: A typical 48 port New Photonic Connector Receptacle configuration at the face plate of a TOR switch for instance. Other configurations of the Receptacles are possible and readily achieved.

5.2 Maximum Density Configuration

Configurations using the dense-packed New Photonic Connector Receptacle design in order to achieve maximum fiber density are also conceived of, as explained above. One particular implementation of this concept is shown below in Figure 5.2, where 6 New Photonic Connector Receptacles are shown connected to the face plate of a single compute system. This system achieves a fiber density increase of 42% from the single Receptacle design, and illustrates some of the benefits of using the New Photonic Connector Receptacle design in a dense-pack application.

The design achieves this increase by placing all 6 of these New Photonics Connector Receptacles side-by-side into a single cut-out in the face plate. This single, wide cutout will be 54.0-54.5mm when loaded with the 6 Receptacles. Referring to Figure 4.3 and Table 4.1; the concept drives the dimension labeled as ‘G’ to zero, while all other dimensions will remain the same.

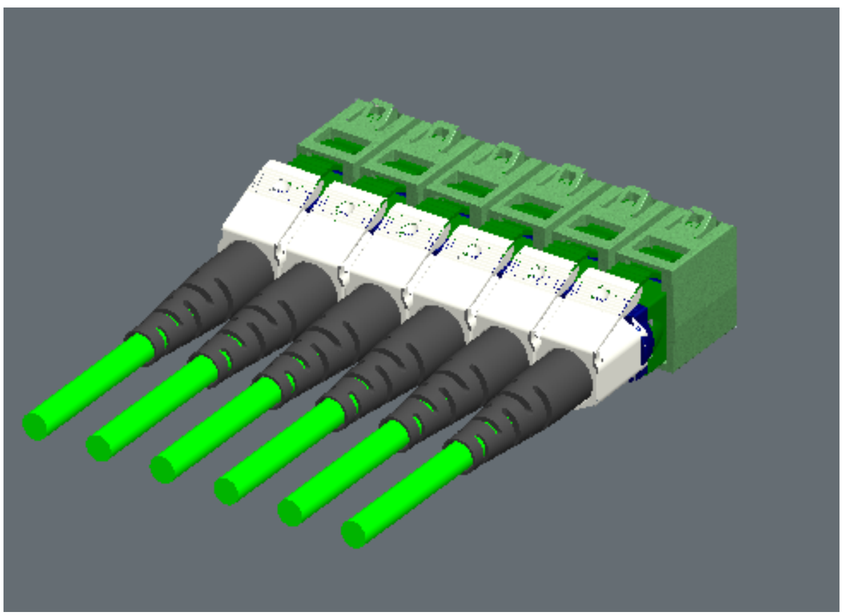


Figure 5.2: Dense-packed New Photonic Connector Receptacle Configuration – This figure illustrates the fiber density improvement possible when dense-packing New Photonic Connector Receptacle concept. For the same 6 fiber connections the New Photonic Connector Receptacle concept results in an improvement in fiber density of 42 %.

6. Environmental Requirements

6.1 Receptacle Physical Force Specifications

Figure 6.1 illustrates the maximum allowable force which can be placed on the Receptacle while it is mated to the faceplate.

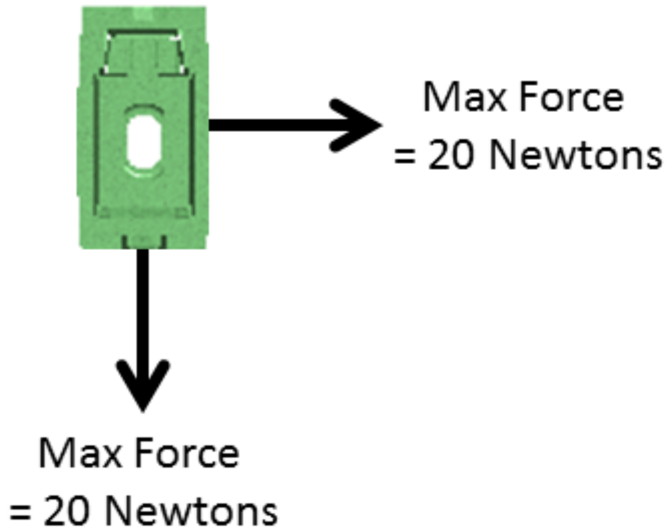


Figure 6.1: An illustration of the maximum force which can be applied to the New Photonic Connector Receptacle in both the 'X' and 'Y' directions when mated to the compute system face plate

6.2 Thermal Specifications

The connector shall meet the following environmental requirements:

- Ambient operating temperature range: 0°C to +70°C
- Operating and storage relative humidity: 10% to 90% (non-condensing)
- Storage temperature range: -40°C to +70°C
- Transportation temperature range: -55°C to +85°C (short-term storage)

7. Reference Documents

The following specification is available on the Open Compute website:

- Open Rack specifications and documents (http://opencompute.org/projects/open-rack/)
- Open Rack Hardware specifications (http://opencompute.org/projects/intel-motherboard/)