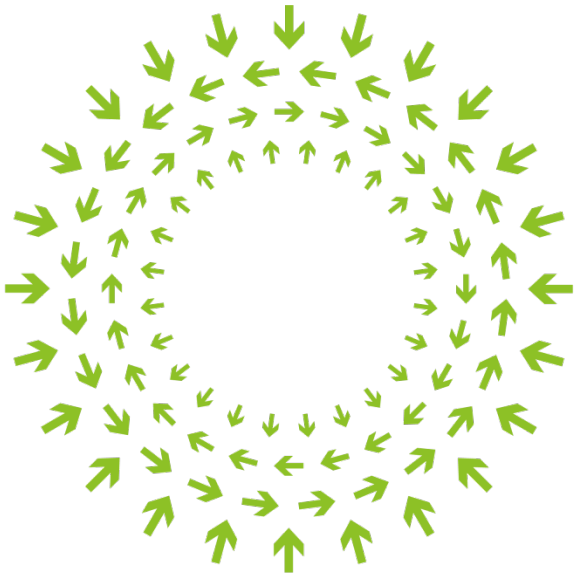

OPEN

Compute Project

Facebook: CWDM4-OCP

100G Optical Transceiver Specification

Authors:

Katharine Schmidtke, Technical Sourcing Manager, Facebook

Vincent Zeng, Supplier Quality Engineer, Facebook

Abhijit Chakravarty, Hardware Validation Engineer, Facebook

1. Revision History

Date/Version	Name	Description
1/9/17, v0.1	Katharine Schmidtke	Internal draft of CWDM4-OCP 100G optic specification

As of January 9th, 2017, the following persons or entities have made this Specification available under the Open Compute Project Hardware License (Permissive) Version 1.0 (OCPHL-P), which is available at http://www.opencompute.org/community/get-involved/spec-submission-process/.

Facebook, Inc.

2. Scope

This document defines a technical specification for 100G optical transceivers that are optimized for large scale data center applications. This specification is being submitted to the Open Compute Foundation.

3. Contents

4. Overview

This document defines a relaxed specification for 100G optical interconnects optimized for large scale data centers applications. This specification is currently being deployed in Facebook's switch fabric inside data centers and is being presented for submission into the OCP Networking Group to be shared with the OCP community.

5. Optical Interconnects in Data Center Applications

To meet current requirements and future bandwidth demands, data center operators are continuously pursuing ways to handle data more efficiently and at ever-faster speeds. Inside data centers, that means finding cost-efficient ways to connect together the switch fabric at higher data-rates. The diagram below (Figure 1) shows the layout of a typical Facebook data center built using our Fabric architecture. Details of this architecture are available here https://code.facebook.com/posts/360346274145943/introducing-data-center-fabric-the-next-generation-facebook-data-center-network/

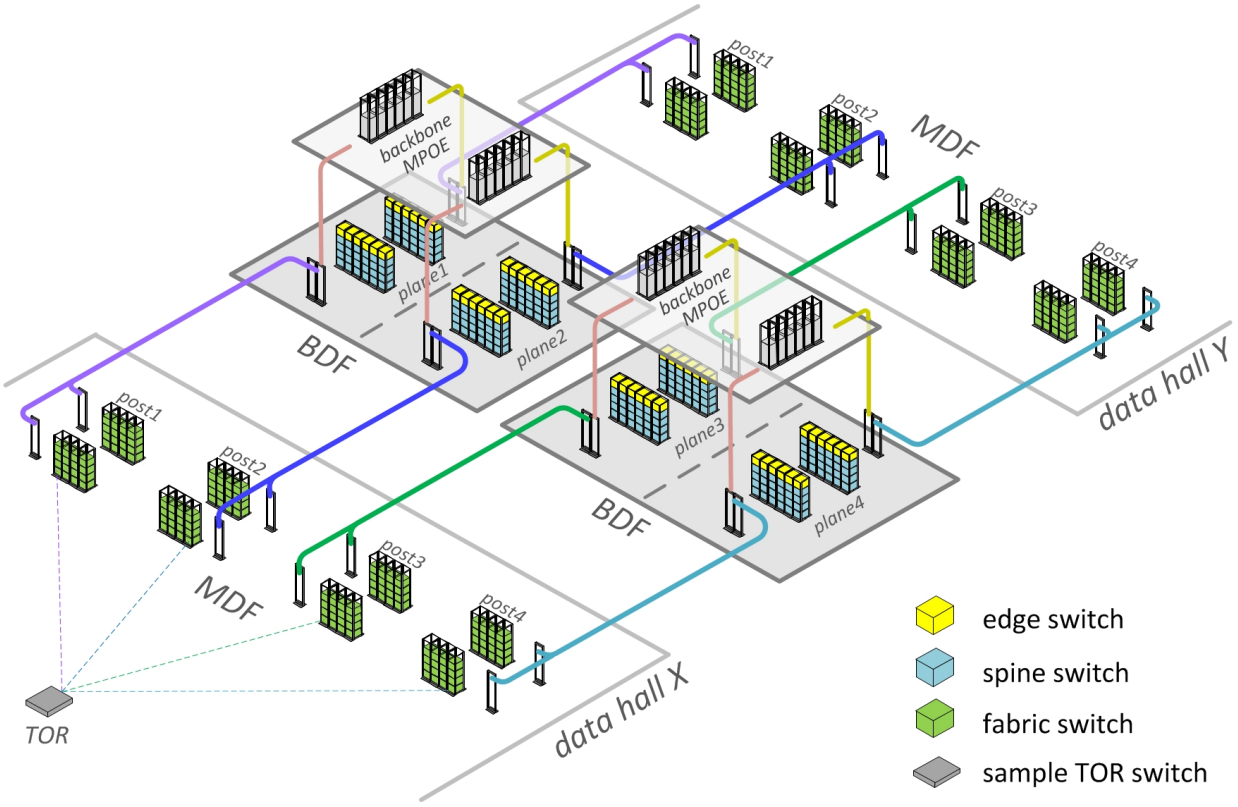

Figure 1: Schematic diagram showing optical cabling inside a Facebook data center

The networking interconnects are represented by colored lines between the different switching elements. Each connection is made up of a fairly simple optical link consisting of a minimal number of connections and relatively short link lengths below 500m. In the past, the switch fabric inside many hyper data centers was connected at data-rates of 40 Gbit/s or less over multi-mode optical fiber. As the data-rate requirements increased, new interconnects were needed which put a strain on the multi-mode fiber infrastructure deployed in data centers and there has been strong motivation to move to single-mode fiber.

For example, at 40Gbit/s the data center was cabled with OM3 multi-mode fiber. To reach 100m at 100 Gbit/s using standard multi-mode optical transceivers would require re-cabling from OM3 to OM4 multi-mode fiber. It is a large practical and financial challenge to change fiber at each generation of interconnect technology. Ideally, the fiber plant should remain installed for the whole life of the data center and support many technology innovation cycles. Therefore, there is strong motivation to move to single-mode fiber in data-centers to provide flexibility and ensure support of longer link lengths and many data-rate evolution cycles beyond 100 Gbit/s.

The amount of fiber deployed in hyperscale data centers can be very large, with tens of thousands of kilometers of multi-mode fiber typically deployed in each data center. In the past, 40G multi-mode links were carried on four 10G channels on separate parallel fibers. For a bi-directional link, four parallel fibers were needed in each direction. However, by multiplexing four different wavelengths onto a single fiber, the number of fibers for a bi-directional link can be reduced by a factor of four. Therefore, for a bi-directional link a single duplex pair of fibers is needed instead of eight separate parallel fibers.

The goal here is to identify an optical transceiver specification optimized to the requirements of data center applications at 100Gbit/s over duplex single-mode fiber. The performance requirements are relaxed compared to traditional telecommunication applications in terms of the link length, link budget, and temperature range, all the while maintaining reliability and high quality performance.

6. 100G CWDM4-OCP Specification

The starting point for the 100G CWDM4-OCP specification optimization is the specification adopted by the CWDM4 MSA. Information on this specification can be found at http://www.cwdm4-msa.org/. This specification is based on a coarse wavelength grid (CWDM = 20 nm spacing), which for many different technology approaches does not require active cooling inside the optical module as compared to dense wavelength grids where cooling is required to keep the laser wavelengths stable.

The optimizations to the specification are in three main areas: reach, loss-budget, and temperature. Firstly, the simplicity of the optical link and its relative short distance allowed the reach to be reduced down from 2 kilometers to 500 meters based on a

reduced link loss of 3.5 dB rather than 5 dB. In addition, most data centers operate under a very predictable thermal environment and the air-inlet temperature to the switching equipment is very well-controlled. If the air-inlet temperature is well-controlled, so too is the thermal environment inside the switch. This means that the specification on the operating case temperature that the optical transceiver could be relaxed down from 0-70 ^{o}C to 15-55 ^{o}C. The specification relaxations are summarized in Table 1 below. This relaxed specification will interoperate with the standard MSA version of the specification over distances <500m.

Table 1: Specification Relaxations compared to MSA
Referenced to CWDM4 MSA http://www.cwdm4-msa.org

	CWDM4-OCP Relaxed Specification	CWDM4 MSA Base Specification
Reach	500 m	2000 m
Tx OMA	-5 dBm	-4 dBm
Tx OMA-TDP	-6 dBm	-5 dBm
Rx Sensitivity	-9.5 dBm	-10 dBm
Link loss	3.5 dB	5 dB
Operating Case Temperature	15-55 ^{o}C	0-70 ^{o}C

7. Ecosystem and Availability

The CWDM-OCP optical transceiver specification is supported by an ecosystem of network elements and optical transceiver suppliers..

7.1. Optical Transceiver Manufacturers

Examples of the optical transceiver suppliers that are supporting this specification are listed below along with the part number for the CWDM4-OCP module

Table 2: Optical Transceiver Suppliers supporting this specification include:

Manufacturer	CWDM4-OCP Part Number
Colorchip	C100 Q005 CWDM4 02B
Finisar	FTLC1152RGPL6

The CWDM-OCP version of the 100G optical transceivers are already being deployed and are commercially available from several suppliers. The CWDM4-OCP modules can be distinguished from other QSFP optical transceivers by the "OCP-green" colored pull tab as shown in Figure 2.

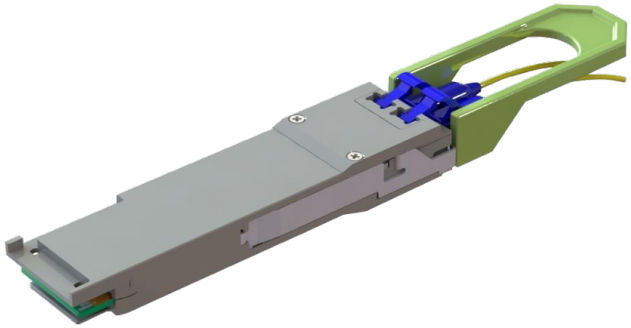

Figure 2: Example of commercially available 100Gbit/s CWDM4-OCP module with "OCP-green" pull-tab.

7.2. Switch Equipment Manufacturers

The ecosystem of switch vendors that support this specification include switches already contributed/accepted to OCP and are listed in Table 3.

Table 3: Switch platforms that support this specification include:

Supplier Name	Switch Description
Edgecore Networks	Wedge100-32X (Accepted)
Celestica	Backpack 100G – Part # D8020 (Contributed as of Q4 2016)

8. Summary

An optical transceiver specification for 100G duplex single-mode optical interconnects has been developed that is optimized for data center applications. This solution is already commercially available and part of an ecosystem of networking elements that is being successfully deployed in several Facebook data centers. It is being submitted to the Networking Track of the Open Compute Project to share with the rest of the data center community.