



# 100Gb/s CWDM4 Optical Interconnect at Facebook Data Centers for Bandwidth Enhancement

Abhijit Chakravarty, Katharine Schmidtke, Vincent Zeng, Srinivasan Giridharan, Cathie Deal and Reza Niazmand  
Facebook Inc., 1601 Willow Road, Menlo Park, CA 94025

## ABSTRACT:

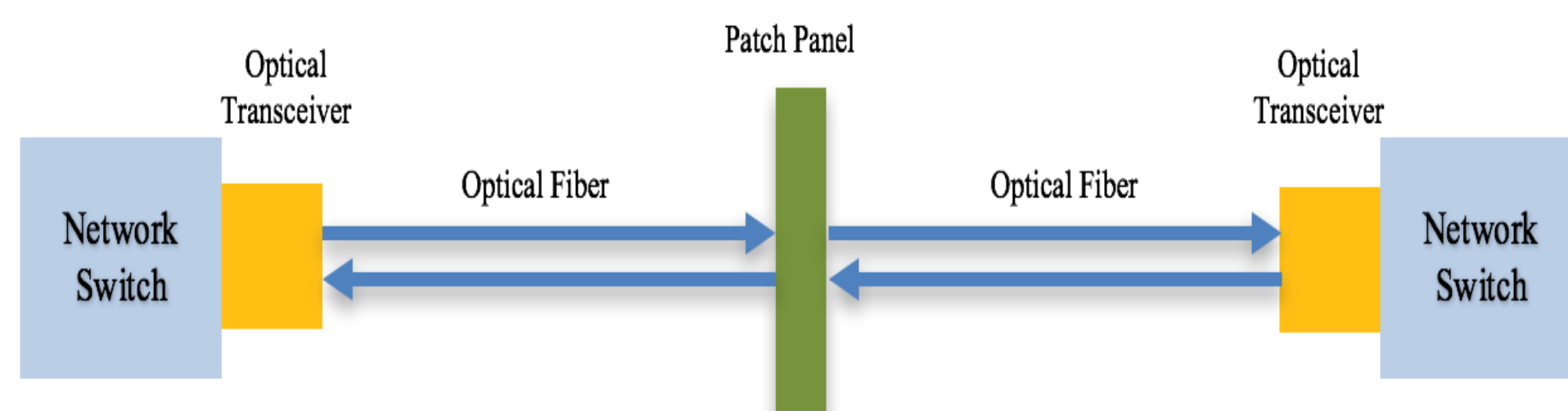
Facebook has developed 100G data centers from the ground-up by fine tuning optical technologies, optimizing link-budget, limiting operating temperatures and ultimately improving manufacturability. 100G-CWDM4 is an effective technology to enable connectivity over duplex single-mode fiber.

## INTRODUCTION:

As the Facebook (FB) community grows and the user experience becomes richer, more immersive, and more interactive with capabilities such as live video, 360 video, and virtual reality, the demand to handle more data efficiently increases exponentially for the network. This drives the need to advance to newer interconnect technologies at a rapid pace. To achieve this, FB is continually experimenting and intensively validating the advantages and disadvantages for various approaches in the real-world FB infrastructure. From the available 100G technologies, FB selected 100G CWDM4-OCF [1,2] and modified it to meet our data center operating requirements. This paper describes the details behind the analysis of FB's next generation optical interconnects.

## OPTIMIZING FOR DC ENVIRONMENT:

The typical data center consists of a combination of network switches, patch panels, optical fiber and optical transceivers. At the end of each link an optical transceiver converts electrical signals to optical, and transmits the signal over optical fiber. On the other end, an identical transceiver will convert the optical signal to electrical and transmit it to a switch port. To span the length of datacenters, patch panels are used to bridge between two network domains and longer distances as shown in fig. 1.



“Fig. 2.(a) Supply voltage (b) QSFP28 case temperature”

Distances in DC are shorter than 500 meters. This allows the reach of the optical link to be reduced from 2 kilometers to 500 meters, based on a reduced link loss budget of 3.5 dB rather than 5 dB. Furthermore, the FB DCs operate under a highly regulated thermal environment. Consequently, the air-inlet temperature to the switch gear is well controlled and monitored. This allows the operating thermal requirements of transceivers to be reduced from 0°C to 70°C (industry standard) to 15°C to 55°C. By allowing parameters such as reach, operating temperature, and link budget to be fine-tuned to the FB infrastructure, we can optimize for transceiver performance. Further, we need to ensure interoperability with the standard MSA specified transceivers operating within 500 meters. The specification relaxations are summarized in Table 1. This relaxed specification [2] ensures that optical transceivers will interoperate with the standard MSA version of the specification over distances shorter than 500 meters.

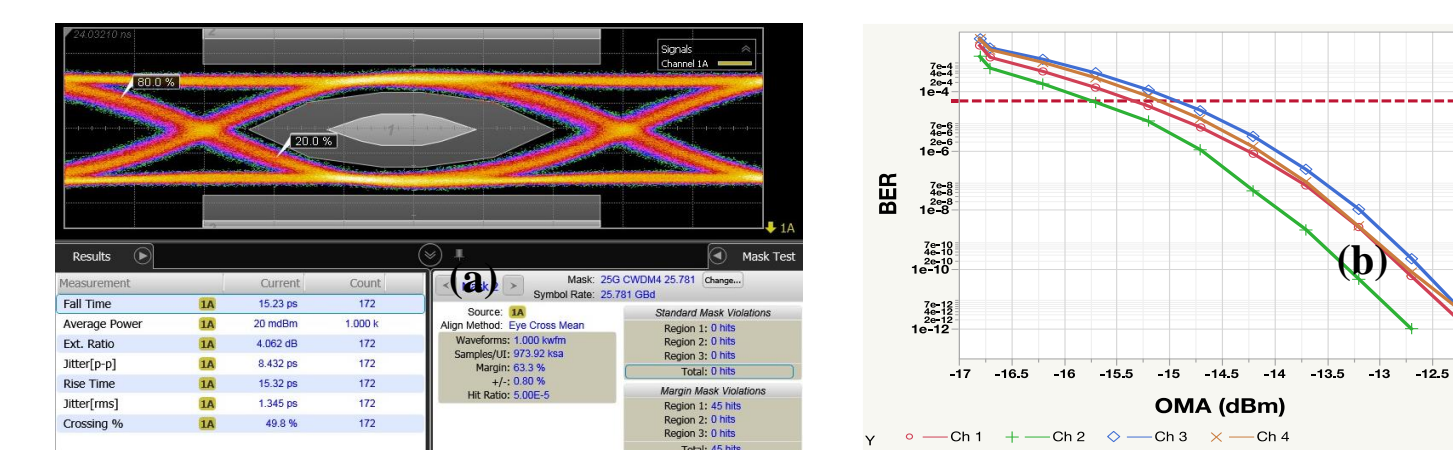
“Table 1 – CWDM4 MSA Spec vs. FB Relaxed Spec”

Optimized Parameters	CWDM4 MSA	FB Relaxed Spec
Reach	2000m	500m
Link loss	5dB	3.5dB
Operating Case Temperature	0°C to 70°C	15°C to 55°C

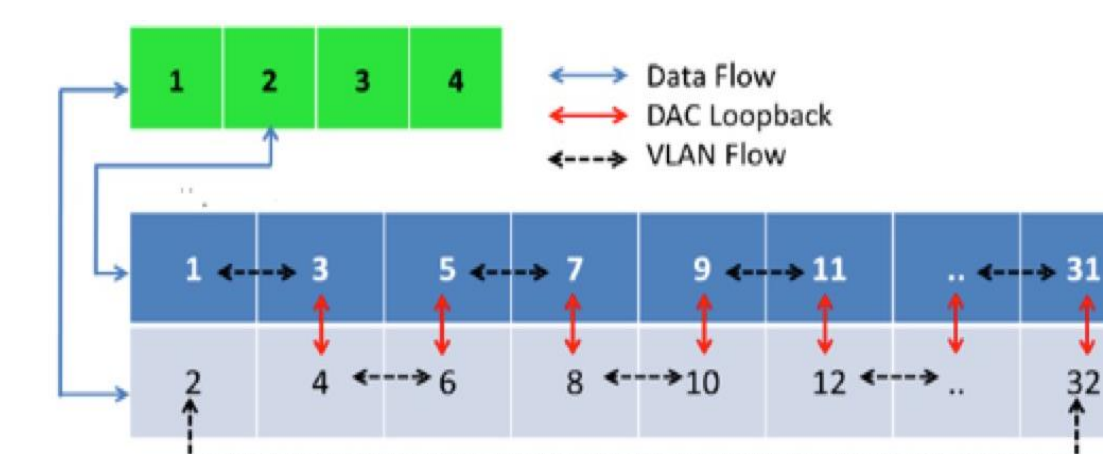
## FB VALIDATION TEST METHODOLOGY:

To ensure that the transceivers would perform reliably once installed, we tested at the component level, system level and in the data center environment. The component level evaluation tests were performed at 25.78 Gb/s per channel to characterize both transmitter (average optical power, extinction ratio, jitter, mask margin, OMA, etc.) and receiver (sensitivity) performance over different temperatures and different fiber lengths as shown in fig. 2 (a) and (b). After component level tests, we connected these modules to our switch platforms to perform system level traffic tests. Fig. 3 shows 100Gbps Ethernet traffic configuration through 32 concatenated CWDM4 ports to perform packet loss tests using different frame sizes and traffic loads.

By stressing the system at different temperatures, we were able to accelerate failure modes and screen out critical problems well before mass deployments. We encountered a variety of technical challenges such as laser mode widening, side mode suppression ratio (SMSR), temperature calibration and PCB delamination, however, the FB testing and validation ensured the deployment of 100G CWDM4 transceivers on time with less infant mortality failure modes. From simple errors of EEPROM coding to complex laser issues, we resolved design challenges before the technology saw live traffic in our data centers.



“Fig. 2. 100G CWDM4 component validation data”



“Fig. 3. 100G CWDM4 system level test structure”

## CONCLUSION:

Our data centers are moving from 40G multimode to 100G single-mode technology to accommodate application demands for higher bandwidth. FB introduced the relaxed spec for 100G CWDM4 in 2015, this approach enabled the industry to take this technology from concept to reality in less than a year. We have started deploying this technology at scale to support our actual production application traffic throughout our data centers.

## References:

- [1] Abhijit Chakravarty, Katharine Schmidtke, Srinivasan Giridharan, John Huang, and Vincent Zeng, “100G CWDM4 SMF Optical Interconnects for Facebook Data Centers”, Conference on Lasers and Electro-Optics, OSA Technical Digest, paper STu4G.1, 2016.
- [2] Katharine Schmidtke, Vincent Zeng, Abhijit Chakravarty “Facebook: CWDM4-OCF, 100G Optical Transceiver Specification”, Open Compute Project © Facebook, 2017.

