



Article 400GbE Technology Demonstration Using CFP8 Pluggable Modules

Yang Yue *, Qiang Wang, Jian Yao, Jason O'Neil, Daniel Pudvay and Jon Anderson

Juniper Networks, 1133 Innovation Way, Sunnyvale, CA 94089, USA; qiwang.thresh@gmail.com (Q.W.); jianyao@juniper.net (J.Y.); joneil@juniper.net (J.O.); dpudvay@juniper.net (D.P.); joanderson@juniper.net (J.A.)

* Correspondence: yyue@juniper.net; Tel.: +1-408-745-2000

Received: 3 October 2018; Accepted: 23 October 2018; Published: 25 October 2018



Abstract: In this article, we first review the current status of 400GBASE client-side optics standards and multi-source agreements (MSAs). We then compare different form factors for 400GE modules, including CFP8, OSFP and QSFP-DD. The essential techniques to implement 400GE, such as pulse amplitude modulation (PAM4), forward error correction (FEC) and a continuous time-domain linear equalizer (CTLE), are discussed. A 400GE physical interface card (PIC) in Juniper's PTX5000 platform has been developed, conforming to the latest IEEE802.3bs standard. To validate the PIC's performance, a commercial optical network tester (ONT) and the PIC are optically interconnected through two CFP8-LR8 modules. The CFP8-LR8 module utilizes eight optical wavelengths through coarse wavelength division multiplexing (CWDM). Each wavelength carries 50 Gb/s PAM4 signal. The signal transmits through 10 km of single mode fiber (SMF). The ONT generates framed 400GE signal and sends it to the PIC through the first CFP8 module. The PIC recovers the signal, performs an internal loopback, and sends 400GE signal back to the ONT through the second CFP8 module. The optical spectrum, eye diagram, receiver sensitivity, long time soaking results, and internal digital diagnosis monitoring (DDM) result are fully characterized. The pre-FEC bit error rate (BER) is well below the KP4 FEC threshold of 2.2×10^{-4} . After KP4 FEC, error-free performance over 30 km of SMF is achieved. In this way, we demonstrate both the interoperation between the PIC and the ONT, as well as the interoperation between the two CFP8 modules. This demonstration represents the successful implementation of the 400GE interface in the core IP/MPLS router.

Keywords: optical communication; fiber optics; client-side optics; 400G Ethernet; CFP8-LR8 transceiver

1. Introduction

Processing information and transmitting information are two fundamental functions in communication networks [1]. From a quantum physics perspective, particles can be classified as fermions or bosons. Electrons follow Fermi-Dirac statistics, while Bose-Einstein statistics apply to photons. Due to strong interaction, electrons are ideal for processing information. As there is minimal interaction between bosons, photons are ideal for transmitting information in different degrees of freedom including wavelength, time, amplitude, phase, polarization, mode, and space [2,3].

Figure 1 illustrates the architecture of a core router, which is widely used in today's communication networks. A typical router includes a routing engine (RE), routing control board (RCB), physical interface card (PIC), flexible PIC concentrator (FPC), and switch interface board (SIB). Information processing in the communication network is mainly performed by these functional units, which is in the electron domain. Moreover, some of Juniper's packet optics products on the PTX platform are also shown in Figure 1. Information transmission is mainly performed by these functional units, which is in the photon domain. The integrated photonics line card (IPLC) tightly

couples a 1 \times 2 wavelength selective switching (WSS), embedded bi-directional switch gain amplifiers, optical multiplexer/de-multiplexers, and optical supervisory channel for optical management into a single package that supports up to 64 ITU-T C-band wavelengths at 50 GHz spacing via an IPLC expansion line card. P3-15-U-QSFP28 and P2-PTX-5-100G-WDM are PTX Series PICs that support 15 QSFP28 and 5 CFP2 modules.



Figure 1. Typical architecture of core router and Juniper packet optical product. RE: routing engine; RCB: routing control board; PIC: physical interface card; FPC: Flexible PIC concentrator; SIB: switch interface board.

In 2017, standardization of the 400 Gigabit Ethernet (400GE) was ratified by the IEEE P802.3bs Task Force [4]. This paves the way for deployment of 400GE in the network. It is expected that the demand for 400GE will grow rapidly over the next couple of years [5]. Recently, multiple industrial line-side and client-side interoperability trials have been successfully demonstrated using 400GE CFP8 pluggable optical modules [6–8]. These trials demonstrate that the ecosystem for the 400GE application is mature. In this paper, we present a successful demonstration of the 400GE physical interface card (PIC) integrated in the core internet protocol/multi-protocol label switching (IP/MPLS) router, with the CFP8-LR8 modules acting as the optical front end.

IEEE defines the optical and electrical parameters for the physical media dependent (PMD) to guarantee interoperability between PMDs. The physical implementations of PMDs are defined by multi-source agreement (MSA) instead. Currently, there are still multiple MSAs defining pluggable optical modules. Figure 2 below shows the comparison between different form factors defined by multiple MSAs [9–16]. The most notable form factors for 400GE defined by MSAs are QSFP-DD, OSFP and CFP8. Currently the industry is converging behind QSFP-DD due to the high port density on the front panel, backward compatibility with QSFP28/QSFP+, and large ecosystem.



Figure 2. Form factors and electrical lanes for pluggable optical modules.

On-off-keying (OOK) has been the modulation format for 40 Gigabit Ethernet (40GE) and 100 Gigabit Ethernet (100GE). The IEEE 802.3bs has selected 4-level pulse amplitude modulation (PAM4) for 400GE transmission. Compared with OOK, the required baud rate of PAM4 could be reduced by a factor of two when keeping the desired net data rate constant. Thus, the transceiver implements PAM4 with fewer optical lanes/lower bandwidth components. This leads to lower cost, smaller power consumption and a denser footprint.

The following variants are defined in 803.bs: 400GBASE-DR4, 50 gigabit per second (Gb/s) 4 parallel single mode fiber transmission (PSM4) up to 500 m; 400GBASE-FR8, 25 Gb/s 8- λ wavelength division multiplexing (WDM) transmission up to 2 km; and, 400GBASE-LR8, 25 Gb/s 8- λ coarse WDM transmission up to 10 km. All interfaces use a 1300 nm window with minimum chromatic dispersion and a low-cost transceiver. For 8- λ WDM transmission, the assignment of the optical wavelength on the coarse wavelength division multiplexing (CWDM) grid removes the requirement of precise temperature control for the laser diode.

To improve sensitivity and increase the transmission distance, forward error correction (FEC) is widely used in today's client-side optics interface. IEEE has approved two FEC coding schemes, KR4 and KP4. KR4-FEC utilizes Reed-Solomon coding (528, 514). The net coding gain is 5.3 dB and the bit error rate (BER) threshold for the uncorrected code word (UCW) is 2.1×10^{-5} . This is widely used in the non-return-to-zero (NRZ) OOK modulation format. However, with the same symbol rate, the sensitivity penalty of PAM4 over NRZ OOK is $10 \times \log_{10} (1/3) = 4.77$ dB. In practice, there is further performance degradation due to the system nonlinearity. In order to close the link budget, KP4-FEC defined in IEEE 802.3bs clause 91, is widely used in PAM4 transmission [17]. KP4-FEC utilizes Reed-Solomon coding (544, 514). The net coding gain is 6.4 dB and the BER threshold for the UCW is 2.2×10^{-4} .

Figure 3 illustrates PAM4 eye diagrams without and with pre-compensation. As one can see, especially for the high baud rate PAM4 signal, its eye can easily be closed due to system bandwidth limitation. Consequently, it is critically important to implement pre-compensation techniques to improve the signal quality and open up the PAM4 eye. A continuous time-domain linear equalizer (CTLE) has been widely used to equalize the frequency-dependent loss for the electrical interface between the packet forwarding engine (PFE) and the pluggable optical module. CTLE is essentially a high-pass filter which inverts the frequency response of the trace on the printed circuit board (PCB). The typical response of CTLE is plotted in Figure 4. In Reference [4], the emphasis of the CTLE is defined in 1 dB steps. However, 0.5 dB steps may be required in large-scale deployment to improve granularity.



Figure 3. (a) Closed 4-level pulse amplitude modulation (PAM4) eye without pre-compensation, (b) open PAM4 eye with pre-compensation.



Figure 4. Typical response curve of the continuous time-domain linear equalizer (CTLE) as defined in Reference [4] in 0.5 dB step size.

2. 400GE Pluggable Module and Host Card

Figure 5 below shows the 400GE CFP8-LR8 pluggable optical transceiver and the block diagram of the physical interface card (PIC) hosting the 400GE pluggable module. The total data rate is 425 Gbit/s, due to the overhead of the FEC block and the 64B/66B mapping in the physical coding sublayer (PCS). The data are carried with 8 optical channels modulated using PAM4 signal running at 53.125 Gbit/s (8×50 G PAM4 WDM optical channels). The electrical interfaces (CDAUI-16) are 16 pairs of differential lanes running at 26.5625 Gbit/s (16×25 G electrical NRZ signals). The CFP8-LR8 uses the distributed feedback laser (DFB) as the transmitter and the PIN photodiode (p-type, intrinsic, n-type photodiode) as the receiver. The total power consumption is less than 16 W. Within the CFP8-LR8 module, a gearbox integrated circuit (IC) multiplexes two electrical lanes of NRZ signal into one optical lane of pulse amplitude modulation (PAM) signal. Gray mapping is implemented so that there is only a one bit difference between the adjacent levels to minimize the BER.



Figure 5. (a) Picture of 400GE CFP8-LR8 pluggable optical transceiver; (b) configuration of physical interface card (PIC). MAC: media access control, PCS: physical coding sublayer, FPGA: field programmable gate array, PFE: packet forwarding engine, CDAUI-16: 400G attachment unit interface, running at 16 lanes with 26.5625 Gb/s per lane.

Juniper's PIC hosts the PFE which forwards the IP packet based on the routing table generated from the routing engine (RE). The additional interfaces, like the bridge, the media access control (MAC) and the PCS layer are implemented with a field programmable gate array (FPGA). The control and management function is implemented in FPGA as well. One PIC can host three pluggable optical modules, giving a total capacity of the PIC of 1.2 Tbit/s.

Figure 6 shows the experimental setup to demonstrate the interoperation of CFP8-LR8 modules. We use an optical network tester (ONT, Viavi) to generate a CDAUI-16 signal. One CFP8-LR8 module

is inserted into the ONT to perform electrical-to-optical conversion. The optical signal goes through a variable optical attenuator (VOA) and is received by a second CFP8-LR8 module, which is hosted by a Juniper's PIC. Within the PIC, the 400GE signal is electrically looped back to the transmitter of the second CFP8-LR8 module. The signal either bypasses or goes through a single mode fiber (SMF) spool. The signal is then received by the CFP8-LR8 from the first module, and the ONT counts the BER.



Figure 6. (a) Experimental demonstration of the interoperation of CFP8-LR8 modules; (b) two CFP8-LR8 modules in Juniper's PIC and PTX5000 chassis.

3. Transmitter Performance

Figure 7 below shows the optical eye diagram obtained through Keysight's digital communication analyzer (DCA). The CDAUI-16 electrical signal is generated by a Viavi optical network tester (ONT) and passed through a pluggable connector. A clock signal derived from the ONT drives the DCA to obtain these optical eye diagrams. As seen, a clear 4-level eye diagram with wide eye opening indicates a good signal to noise ratio (SNR). One can also notice that there is a misalignment in the center of the eye, due to the nonlinear response of the modulator.



Figure 7. Optical eye diagrams of eight wavelength division multiplexing (WDM) channels using PAM4 modulation format. Top: channels 1 to 4; bottom: channels 5 to 8.

Figure 8 shows the optical spectrum from the CFP8-LR8 modules. The optical wavelengths are around 1310 nm so that the influence of chromatic dispersion is minimized. The wavelength assignment of these eight optical channels follows the CWDM designation. This facilitates the interoperation between the CFP8 modules. Also, the measured side mode suppression ratio (SMSR) is well above 40 dB, indicating the excellent jitter performance for the optical modules.

Figure 8. Optical spectra of two CFP8-LR8 modules: (a) first sample module; (b) second sample module. Top: spectrum at the output of the transmitter; Bottom: spectra of the individual lanes after going through a de-multiplexer.

4. Link Performance

In this section, we first study the system performance in a back-to-back scenario, and then extend our experiment to tens of kilometers of fiber transmission.

4.1. Back-to-Back Performance

We first measure the performance of the 400GE CFP8-LR8 module in a back-to-back scenario. Here we bypass the SMF spool as shown in Figure 9. We first adjust the VOA1 to test the interoperation between the transmitter (Tx) of the first module to the receiver (Rx) of the second module. Meanwhile, the attenuation of VOA2 is set to zero so that there is minimum error introduced on the backward path from the PIC to the ONT. As seen, even with 14 dB attenuation between the transmitter and the receiver, the BER is still below the FEC's UCW threshold of 2.2×10^{-4} . For the BER performance of 16 ONT electrical lanes, we note that some lanes perform better than others. This is potentially because of the non-ideal combination of the PAM4 signal from two individual NRZ OOK signals.

Next, we adjust the VOA2 to test the interoperation between the Tx of the second module and the Rx of the first module. Meanwhile, the attenuation of VOA1 is set to zero so that there is minimum error introduced on the forward path from the ONT to the PIC. As seen in Figure 10, even with 12 dB attenuation between the transmitter and the receiver, the BER is still below the FEC's UCW threshold of 2.2×10^{-4} .

We also measure the receiver optical power (ROP) at the different values of attenuation. There is a difference of 2.3 dB between the launching powers of eight optical lanes. At 12 dB attenuation, the average ROP is -12.4 dBm.

Figure 9. Back-to-back performance for interoperation of two CFP8-LR8 modules. The signal is transmitted from the transmitter (Tx) of the first module to the receiver (Rx) of the second module: (a) Bit error rate (BER) vs. VOA1's attenuation for 16 individual electrical lanes; (b) total BER vs. VOA1's attenuation. The threshold for the uncorrected code word (UCW) is shown as the red line.

Figure 10. Back-to-back performance for interoperation of two CFP8-LR8 modules. The signal is transmitted from the Tx of the second module to the Rx of the first module: (**a**) BER versus VOA2's attenuation for 16 individual electrical lanes; (**b**) total BER vs. VOA2's attenuation. The threshold for UCW is shown as the red line.

4.2. Fiber Transmission

Furthermore, we measure the performance of the 400GE CFP8-LR8 module with the single mode fiber. The attenuation of VOA1 is set to zero so that there is minimum error introduced on the forward path from the ONT to the PIC. We adjust the fiber length and measure the BER. The result is shown in Figure 11. As seen, even with 30 km SMF between the transmitter and the receiver, the BER is still below the FEC's UCW threshold of 2.2×10^{-4} . This clearly demonstrates 400G Ethernet interoperation using two CFP8-LR8 modules.

Figure 11. Interoperation of two CFP8-LR8 modules through single mode fiber (SMF) spool. The signal is transmitted from the Tx of the first module to the Rx of the second module: (**a**) BER vs. fiber length for 16 individual electrical lanes; (**b**) total BER vs. fiber length. The signal is transmitted from the Tx of the second module to the Rx of the first module: (**c**) BER versus VOA2's attenuation for 16 individual electrical lanes; (**d**) total BER vs. VOA2's attenuation. The threshold for the UCW is shown as the red line.

5. Conclusions

In this article, we review the current status of the 400G Ethernet standard. The essential techniques to implement 400G client-side optics, like PAM4, FEC and CTLE, are discussed. The PIC in Juniper's PTX5000 platform has been developed with the CFP8-LR8 modules being the optical front-end. The optical signal transmits through 30 km SMF. The pre-FEC BER is well below the threshold for UCW and there is no post-FEC error. This demonstration represents the successful implementation of a 400G Ethernet interface in the core IP/MPLS router using the CFP8-LR8 pluggable modules.

Author Contributions: Conceptualization, Y.Y. and J.A.; methodology, Y.Y.; software, Y.Y. and Q.W.; validation, Y.Y.; formal analysis, Y.Y. and Q.W.; investigation, Y.Y., Q.W. and J.Y.; resources, Y.Y., Q.W., J.Y., J.O. and D.P.; data curation, Y.Y.; writing—original draft preparation, Y.Y. and Q.W.; writing—review and editing, Y.Y., Q.W., J.Y., J.O., D.P. and J.A.; visualization, Y.Y. and Q.W.; supervision, J.A.

Funding: This research received no external funding.

Acknowledgments: The authors gratefully acknowledge the vigorous encouragement and strong support for innovation from Domenico Di Mola at Juniper Networks.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- 1. Kaminow, I.; Li, T.; Willner, A.E. *Optical Fiber Telecommunications*, 6th ed.; Academic Press: San Diego, CA, USA, 2013; ISBN 9780123969606.
- Bozinovic, N.; Yue, Y.; Ren, Y.; Tur, M.; Kristensen, P.; Huang, H.; Willner, A.; Ramachandran, S. Terabit-scale orbital angular momentum mode division multiplexing in fibers. *Science* 2013, 340, 1545–1548. [CrossRef] [PubMed]
- 3. Winzer, P.J. Making spatial multiplexing a reality. Nat. Photonics 2014, 8, 345–348. [CrossRef]
- 4. IEEE P802.3bs 400 Gb/s Ethernet Task Force. Available online: www.ieee802.org/3/bs/ (accessed on 1 September 2018).
- Rokkas, T.; Neokosmidis, I.; Tomkos, I. Cost and Power Consumption Comparison of 400 Gbps Intra-Datacenter Transceiver Modules. In Proceedings of the 2018 International Conference on Transparent Optical Networks (ICTON), Bucharest, Romania, 1–5 July 2018.
- Nelson, L.E.; Zhang, G.; Padi, N.; Skolnick, C.; Benson, K.; Kaylor, T.; Iwamatsu, S.; Inderst, R.; Marques, F.; Fonseca, D.; et al. SDN-Controlled 400GbE end-to-end service using a CFP8 client over a deployed, commercial flexible ROADM system. In Proceedings of the 2017 Optical Fiber Communications Conference and Exposition (OFC), Los Angeles, CA, USA, 19–23 March 2017.
- 7. Birk, M.; Nelson, L.E.; Zhang, G.; Cole, C.; Yu, C.; Akashi, M.; Hiramoto, K.; Fu, X.; Brooks, P.; Schubert, A.; et al. First 400GBASE-LR8 interoperability using CFP8 modules. In Proceedings of the 2017 Optical Fiber Communications Conference and Exposition (OFC), Los Angeles, CA, USA, 19–23 March 2017.
- 8. Nelson, L.E. Advances in 400 Gigabit Ethernet Field Trials. In Proceedings of the 2018 Optical Fiber Communications Conference and Exposition (OFC), San Diego, CA, USA, 11–15 March 2018.
- 9. CFP-MSA. Available online: http://www.cfp-msa.org/ (accessed on 1 September 2018).
- 10. SFF Committee. Available online: http://www.sffcommittee.com/ie/ (accessed on 1 September 2018).
- 11. OSFP. Available online: http://osfpmsa.org/index.html (accessed on 1 September 2018).
- 12. QSFP-DD. Available online: http://www.qsfp-dd.com/ (accessed on 1 September 2018).
- 13. OIF: Optical Internetworking Forum. Available online: http://www.oiforum.com/ (accessed on 1 September 2018).
- 14. Cole, C. Beyond 100G client optics. IEEE Commun. Mag. 2012, 50, s58-s66. [CrossRef]
- Yue, Y.; Wang, Q.; Maki, J.J.; Zhang, B.; O'Neil, J.; Vovan, A.; Anderson, J. Latest Industry Trend in Pluggable Optics. In Proceedings of the 2017 International Conference on Optical Communications and Networks (ICOCN), Wuzhen, China, 7–10 August 2017.
- Isono, H. Latest standardization trends for client and networking optical transceivers and its future directions. In Proceedings of the 2018 optoelectronics, photonic materials and devices conference (SPIE OPTO), San Francisco, CA, USA, 27 January–1 February 2018.
- 17. Chagnon, M.; Lessard, S.; Plant, D.V. 336 Gb/s in Direct Detection below KP4 FEC Threshold for Intra Data Center Applications. *IEEE Photonics Technol. Lett.* **2016**, *28*, 2233–2236. [CrossRef]

© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).