Progress in 100G Lambda MSA Based on 100G PAM4 Technology

Mark Nowell, Matt Traverso, Marco Mazzini, Kumar Lakshmikumar, Mark Webster, Peter De Dobbelaere
Cisco, 170 West Tasman Drive, San Jose, CA 95134-1706, USA
mnowell@cisco.com

Abstract: This talk will focus on the progress of the 100G Lambda MSA. Topics include: motivation in forming the group; market requirements for the technology; key technologies and results; and insights into next generation work.

1. Introduction

The rapid growth in global network traffic continues due to the combination of: growing number of users; the growing number of devices used to access the network; the improvements in access speeds (resulting in a better quality of experience); and the increase of higher bandwidth content (e.g. video). With the global IP traffic being estimated to continue to grow through 2022 at an annual CAGR of 26% [1] this puts considerable pressure on the underlying interconnect technologies to maintain progress in increasing speeds while reducing costs ($ per Gbps).

In today's network topologies this pressure is most keenly felt within the cloud-scale data center deployments where the scale and scope of both the networks and their facilities are unprecedented. The goal of these network operators is to minimize the overall power consumption (W/Gbps), minimize the costs ($/Gbps) and maximize the reliability.

Cloud-scale networks generally use a leaf-spine type network architecture to efficiently connect the vast number of servers within a facility. This architecture has many advantageous but results in significant interconnect density between servers to switches and between switches in the network with estimates of 70-85% of the traffic being considered East-West traffic (i.e traffic that stays within the facility vs ingress and egress traffic that leaves the facility [2].

Unsurprising, network planners are highly motivated to ensure the interconnect technology being used within these data centers are capable to meeting the performance requirements while at the same time lowering power and cost as much as possible.

2. Industry Standardization forums

A necessity of networking industry is commonality of specifications. This ensures interoperability and also enables multiple companies to make similar solutions resulting in economies of scale which both increases the market size and decreases the costs. The networking industry uses a number of different forums to develop these specifications. The IEEE 802.3 standards group has traditionally been where the foundational standards work is done on new technologies that are being brought to the market. For example, the seminal specifications for 10 Gigabit Ethernet (GbE), 25 GbE, 40 GbE, 50 GbE, 100 GbE, 200 GbE and 400 GbE were all defined in IEEE 802.3. The thoroughness of the IEEE specifications ensures interoperability and depending on the project and the scope of the new technologies within it, it can take a while to work through the process. IEEE 802.3 specifications also do not define implementation rather focusing on functional specifications.

As a consequence, the networking industry also heavily utilizes other specification writing forums where implementations need to be defined (e.g. pluggable module form factors such as QSFP-DD) or where additional interconnect specifications are needed by the industry, building on the foundational IEEE specifications.

The 100G Lambda Multi-source Agreement (MSA) , is one example where a number of companies saw the market interest and need to develop more fully a set of 100 GbE and 400 GbE interfaces to build on the early work with 100 Gb/s per wavelength PAM4 modulated single mode technology that was already defined in IEEE 802.3 at 500m reach for 400 GbE and 100 GbE.

3. 100G Lambda MSA

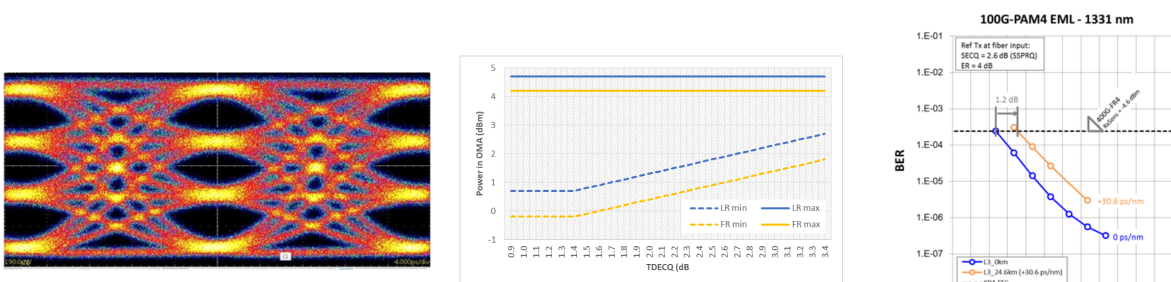
The motivation to accelerate the market adoption of interfaces based on 100 Gb/s PAM4 optical technology was to reduce cost and power which are both key requirements coming from the Data Center operators. Being able to reduce the number of devices as well as relax the wavelength spacing as a consequence had material impacts to the likelihood of hitting the cost, size and power targets as described in [3]. The strong industry momentum behind this activity was clear when 22 companies joined to initiate the group which has now grown to nearly 70 companies. The MSA [4] comprises the definition of an optical transmitter and receiver and supporting logical functions for coding of the data. The intent of the MSA is to create specifications to allow for interoperability between a number of implementation options resulting in industrial competition to deliver the lowest cost solution to the marketplace. As a result of the progress made by companies within the MSA, key technology building blocks have been brought to market and are being widely deployed today for 100 GbE at 2 and 10 km and 400 GbE at 2 and 10 km reaches.

3.1. Transmitter

While the reduced baud rate of PAM4 modulation relaxes the overall required bandwidth, the multi-level signaling imposes linearity requirements. The MSA does not assume a specific implementation for the optical transmitter. Data from integrated EA-DFB and silicon photonic optical transmitters have been presented. [3,5]

The TDECQ (Transmitter and Dispersion Eye Closure Quaternary) parameter [6,7] is a parameter governing the transmitter performance and provides a system level metric derived from estimating BER based on statistics of the transmitted waveform. An example waveform with good TDECQ is shown in Figure 1.

The transmitter must adhere to the OMA (Optical Modulation Amplitude) to insure that the optical signal to noise ratio is sufficient for the far end receiver. The transmitter optical power is constrained by both the OMA and the OMA minus TDECQ as shown in Figure 1.



a) b) c)

Fig. 1. (a) example SiPhotonics transmit waveform with ER = 9.85 & TDECQ = 1.45; (b) Transmit Power range over permitted TDECQ values; (c) EML transmission results with high dispersion (courtesy CIG)

3.2. Receiver

The optical receiver comprises a photodiode, a transimpedance amplifier (TIA), and a decision circuit. The decision circuit may be implemented in an analog slicer fashion or it may leverage an ADC (Analog to Digital Converter). The MSA assumes that the receiver implements an equalizer. The IEEE standard [8] defines a reference equalizer – 5 tap FFE T-spaced. This is a discrete linear equalizer intended to mitigate ISI and other equalizable impairments. However, linear equalizers do not distinguish signal and noise, so no SNR improvement is available and link improvement is offered via flattening the frequency response of the available bandwidth. [10]

Accordingly, while the required bandwidth for PAM4 modulation is relaxed, the sensitivity to noise is increased. Compared to PAM2, often referred to as NRZ, the eye height is 1/3 which works out to 4.8 dB optical penalty. This inherent reduction in signal strength imposes a need for a low noise optical front end. As shown in [9], a TIA with low input referred noise provides superior sensitivity and link BER performance.

In Figure 2, an example receiver performance is shown. The dotted lines between (a) and (b) in Figure 2 highlight that the receiver is performing at near the KP4 FEC limit of 2.4 e-4 BER. This example receiver has significant margin to the 100G Lambda MSA LR receiver sensitivity limit.

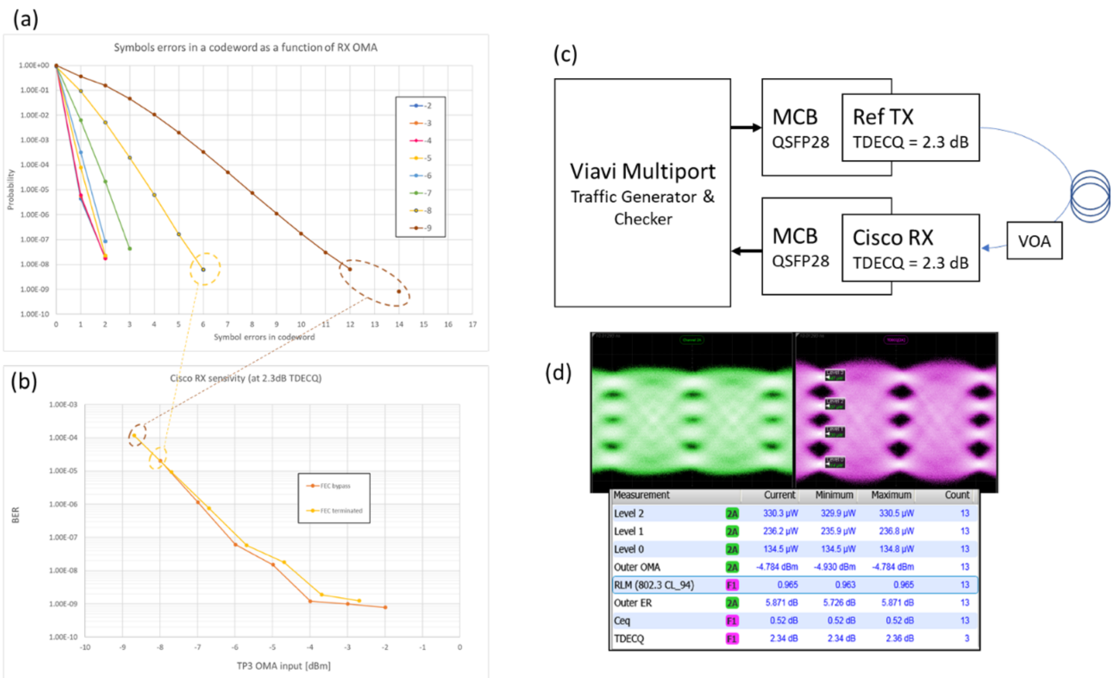


Fig. 2. (a) Symbol codewords plotted for different OMA levels; (b) Receiver sensitivity and how it corresponds to codeword errors; (c) Experimental setup with FEC terminated either in Viavi or in module; (d) Characteristics of reference transmitter

4. Summary

When there is a clear market need to accelerate a technology's broad adoption, organizations such as the 100G Lambda MSA are created to bring multiple companies together with common goals even though they may have different competing sub-component technologies. Through significant analysis, testing and review, a set of specifications gets developed that allows the industry to efficiently develop, deploy and then compete. While the MSA scope is limited to the current 100 Gb/s technology, this work has been foundational to the development of what comes next be it higher baud rates or shortwave VCSEL technology.

5. References

[1] Cisco VNI Global IP Traffic Forecast, 2017–2022.

[2] Cisco Global Cloud Index, 2016–2021.

[3] D. Lewis, M. Nowell, J. Maki, K. Tamura, "100 Gb/s Per Lane Optical PHYs for 2 km and 10 km for 100 GbE and 400 GbE", [Online]. Available: http://www.ieee802.org/3/cfi/1118_1/CFI_01_1118.pdf. November 2018.

[4] QSFP-DD Multi-source Agreement, [On-line] http://www.qsfp-dd.com

[5] M. Mazzini, "100GBASE-DR: 'GoldenEye'", [Online]. Available http://www.ieee802.org/3/cd/public/May18/mazzini_3cd_01a_0518.pdf

[6] J. King, D. Leyba, and G. D. Le Cheminant, "TDECQ (transmitter dispersion eye closure quaternary) replaces historic eye-mask and TDP test for 400 Gb/s PAM4 optical transmitters," in Proc. Opt. Fiber Commun. Conf. Exhib., Mar. 2017, pp. 1–3.

[7] "IEEE 802.3cd 50 Gb/s, 100 Gb/s, and 200 Gb/s Ethernet Task Force,"[Online]. Available: http://www.ieee802.org/3/cd/index.html

[8] "IEEE P802.3bs 200 Gb/s and 400 Gb/s Ethernet Task Force," [Online]. Available: http://www.ieee802.org/3/bs/index.html

[9] K. Lakshmikumar, A. Kurylak, M. Nagaraju, R. Booth, J. Pampanin, "A process and temperature insensitive CMOS linear TIA for 100 Gbps PAM-4 optical links", Proc. IEEE Custom Integr. Circuits Conf., pp. 1-4, Apr. 2018.

[10] S. H. Hall and H. L. Heck, Advanced Signal Integrity for High-Speed Digital Designs, John Wiley & Sons, 2009.