PAM4:

The New Modulation Standard for High-Speed Ethernet Serdes

Introduction

When Draft 1.0 of the 802.3bs standard for 200Gbps and 400Gbps Ethernet was released in 2015, multi-level modulation was viewed as a potential substitute at high speeds for the widely used and well-characterized Non-Return to Zero (NRZ) modulation, common at both 10Gbps and 25Gbps. Soon after its release, however, the IEEE indicated a preferred modulation of four-level Pulse Amplitude Modulation (PAM4) for all single-lane serdes channels of 56Gbps and higher, though 25Gbps NRZ remains an alternative for specialized 400G SR16 implementations. The 56Gbps lanes are used for both 50 Gigabit Ethernet in its native format, and for lane aggregation in 100 Gigabit Ethernet, 200 Gigabit Ethernet, and 400 Gigabit Ethernet. The 802.3bs standard is slated for approval before the end of 2017, and PAM4 will be tapped as the workhorse of choice for the 56Gbps serdes physical layer. It is likely that developers will attempt 112Gbps single-lane serdes based on PAM4 as well, because higher-layer PAM modulation such as PAM8 and PAM16 is considered too difficult to implement in a cost-effective manner. Therefore, PAM4 is likely to remain a critical element of physical-layer Ethernet testing for the foreseeable future.

Line Coding and the Limits of NRZ

Through two decades of evolution of Ethernet from Fast (100Mbps) to Gigabit Ethernet, then to 10GbE and 100GbE, more attention traditionally has been paid to **block encoding**, in which data bits are grouped into blocks from MAC to PHY, prior to being sent to the PHY. **Line coding**, performed by a serializer to prepare for transmission on a physical medium, has relied on a simple two-level code scheme called **Non-Return to Zero (NRZ)**. In contrast to 10MbE's use of direct Manchester encoding, all Ethernet rates of 100Mbps and higher required a simple voltage change to turn a logic level in code space into a signal for transmission over a medium.

PAM4

The New Modulation Standard for High-Speed Ethernet Serdes

NRZ was chosen for its efficiency over alternatives such as **Return to Zero (RZ)** and **Encoded NRZ (ENRZ)**. Block codes often were chosen in the context of existing line codes – for example, the use of Manchester 8B/10B as the input block code for a serdes can maintain an accurate clock for NRZ. Nevertheless, NRZ was considered the obvious choice of line code for an Ethernet PHY design. To address NRZ requirements Spirent has the only quint speed series of products in the industry supporting 100/50/40/25/10GbE on the same test module. The flexibility to offer quint speed and support for NRZ across several ethernet speeds makes these test modules a popular choice.

For speeds of up to 10Gbps, a single signaling lane carrying Ethernet packets at the specified line rate was considered the easiest way to implement serdes functions. Beginning with 40GbE and particularly with the arrival of 100GbE, however, PHY designers looked to the aggregation of multiple serdes lanes as the most efficient way to design a transceiver—4x10GbE to realize a 40GbE transceiver, for example, or 10x10GbE/4x25GbE to implement 100GbE. In fact, the aggregation discontinuities between 40GbE and 100GbE gave rise to a new physical re-timing device called a **Gearbox**, unique to the remapping requirements of 40GbE and 100GbE PHY devices.

Once lane aggregation became the norm, the ability to retain NRZ as a line coding option beyond 28Gbps (25GbE plus overhead) became harder. Even when designers attempted a single-lane 56Gbps serdes based on NRZ, the channel losses encountered when attempting to support the Optical Internetworking Forum's electrical signaling rate for very-short-reach (VSR) lines were daunting. Lane aggregation added significantly to the problem. NRZ traditionally was seen as best suited for shorter distances, while PAM4 is preferred when running longer traces on PCBs. When NRZ was used in line coding for four multiplexed 28Gbps serdes in a 100GbE design, the crosstalk became problematic, requiring proprietary DSP solutions in some early implementations of 100G using 25G NRZ lanes. From the time the task force for 400GbE was launched in 2012, it was obvious a new multi-level concept for line coding was required.

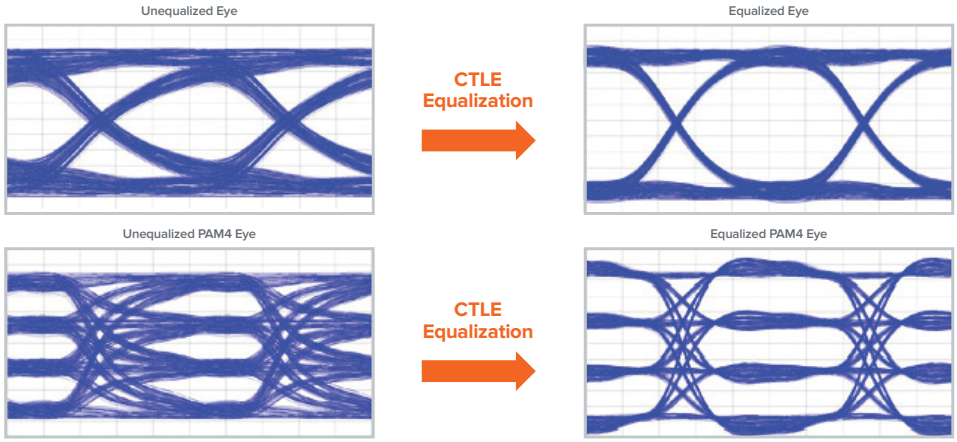


Figure 1: Unequalized vs. Equalized Eye

While PAM lacks the history in mainstream mixed-signal design that is offered by NRZ, it has been studied in various transmission systems for decades. In fact, NRZ is the equivalent of a two-level PAM, or PAM2. When four levels encode two data bits per unit interval, the bandwidth in the transceiver is effectively doubled. In the same way that quadrature amplitude modulation (QAM) in video transmission systems now extends to large constellations such as QAM-64 and QAM-128, initial Ethernet serdes designs have been attempted for PAM8 and PAM16. However, PAM4 provides faster signaling rates without significantly increasing design challenges. For a given speed, it has only half the Nyquist frequency of NRZ – 14GHz compared to 28GHz. Therefore, IEEE 802.3bs defined it as the line code of choice for 56Gbps lanes used in 200GbE and 400GbE. The initial studies on the next fastest serdes, 112Gbps, suggest PAM4 can be used at this speed as well, albeit not without some challenges requiring advanced signal processing solutions.

PAM4: Unexpected Complexity

Although PAM4 may have a longevity in Ethernet serdes design approaching that of NRZ, it would be wrong to underestimate the new design challenges encountered with the coding. In extremely short-reach applications (XSR) between chip and optical engine, acceptable eye-opening heights in optical eye diagrams can be achieved through the use of transmit finite-impulse response (FIR) filters. Once cabling distances hit the VSR range of chip to module, however, PHY designs must replace the common decision-feedback equalizers with continuous-time linear equalizers (CTLE). In most designs, forward error correction must be added, adding to latency and complexity of the PHY circuit. To address these issues Spirent adopted forward error correction parameters on its quad speed 400/200/100/50GbE test modules by offering Pre, Post and FEC block statistics.

Power dissipation represents a big step up. Even assuming the existence of 16nm FinFET CMOS process technology, industry simulations show a 56Gbps PAM4 serdes with FEC and CTLE in the receiver dissipates more than twice the power of a 28Gbps NRZ device.

PAM4 could be used in service-provider transport, once a new IEEE study group, proposed in July 2017, is established to consider PHYs for cable reach between 10km and 100km. The PHYs are being considered for privately-managed data center interconnect, but IEEE maintains ties with ITU and ATIS on mapping standards to OTN transport rates. Thus, a PHY developed for metro-reach private Ethernet transport could form the basis of a future metro-reach Carrier Ethernet. The first such PHYs would support 50Gbps and 100Gbps speeds, but discussions are already under way to extend them to the new 802.3bs rates of 200Gbps and 400Gbps. For distances greater than 40km, IEEE may consider new coherent-optics line coding and equalization methods, but for distances between 10km and 40km, PAM4 is almost certain to be the favored modulation.

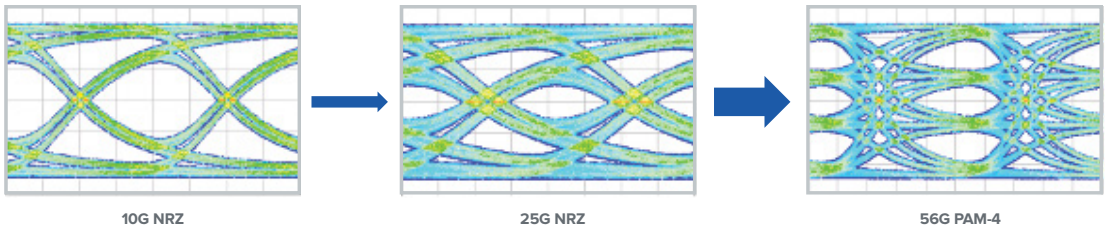


Figure 2: Complexity of 10G to 56G PAM-4

The electrical interface for 112Gbps serdes is in early stages of being discussed within OIF. IEEE indicates that there will be no immediate 800GbE or 1.6TbE follow-on to the 200/400GbE standard. Instead, faster PHYs and serdes will be considered in a special discussion group, the 802.3 New Ethernet Applications or NEA ad-hoc group. NEA is working closely with OIF, which is developing a series of common electrical interfaces, or CEIs, for 112G. Previous generations of OIF CEI standards began the smallest-reach links at chip-to-chip or chip-to-module, but a special group in OIF is developing an even shorter-reach system-in-package interface to link logic and driver chips. Such a link could be used in complex ASICs or FPGAs as a substitute for 2.5D packaging.

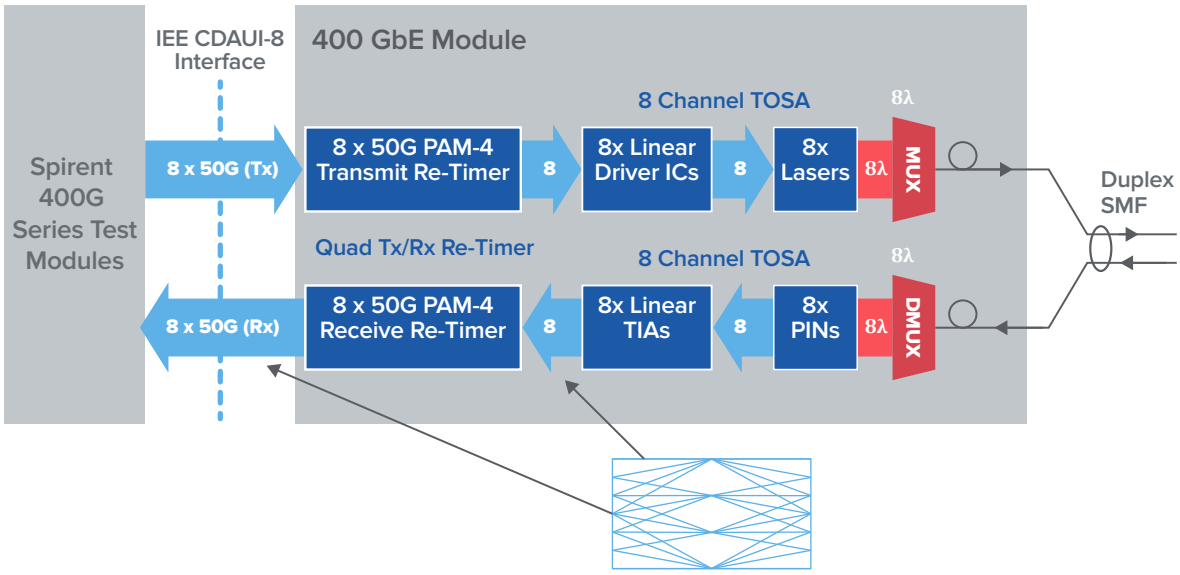
Reach	Media	25G	50G	NG 100GE	200GE	400GE
CR (3 m/5 m)	Copper	P802.3by 25GBASE-CR 25GBASE-CR-S (25G NRZ)	P802.3cd 50GBASE-CR (50G PAM-4)	P802.3cd 100GBASE-CR2 (2x50G PAM-4)	P802.3cd 200GBASE-CR4 (4x50G PAM-4)	-no objective-
SR (100 m)	MMF	P802.3by 25GBASE-SR (25G NRZ)	P802.3cd 50GBASE-SR (50G PAM-4)	P802.3cd 100GBASE-SR/SR2 (2x50G PAM-4)	P802.3cd 200GBASE-SR4 (4x50G PAM-4)	P802.3bs 400GBASE-SR16 (16x25G NRZ)
DR (500 m)	SMF	-no objective-	-no objective-	P802.3cd 100GBASE-DR (1x100G PAM-4)	P802.3bs 200GBASE-DR4 (4x50G PAM-4)	P802.3bs 400GBASE-DR4 (4x100G PAM-4)
FR (2 km)	Duplex SMF	-no objective-	P802.3cd 50GBASE-FR (50G PAM-4)	-no objective-	P802.3bs 200GBASE-FR4 (4x50G PAM-4 CWDM)	P802.3bs 400GBASE-FR8 (8x50G PAM-4 LAN-WDM)
LR (10 km)	Duplex SMF	P802.3cc 25GBASE-LR (25G NRZ)	P802.3cd 50GBASE-LR (50G PAM-4)	-no objective-	P802.3bs 200GBASE-LR4 (4x50G PAM-4 LAN-WDM)	P802.3bs 400GBASE-LR8 (8x50G PAM-4 LAN-WDM)
ER (40 km)	Duplex SMF	P802.3cc 25GBASE-ER (25G NRZ)	-no objective-	-no objective-	-no objective-	-no objective-

*100GBASE-SR is SWDM2 of 2x50G PAM-4

Figure 3: State of standard PMDs: 25G to 400G

Already, NEA is concluding that any future 112Gbps serdes will be based on PAM4, not higher-order PAM. The use of PAM8 or PAM16 would require much more expensive FEC, resulting in high latency, and in large designs that could not be integrated into a single ASIC. Mixing two levels of PAM in one system would require new converter chips, while mixing any FEC other than the approved Reed-Solomon (544.514) would necessitate FEC termination in the transceiver design.

At the same time, however, designing a single-lane 112Gbps serdes based on PAM4 will be far harder than many OEMs anticipate. Insertion losses may require a more advanced circuit-board material for chip-to-module implementations, and at some point, OEMs may decide a transceiver based on on-board-optics is simpler. By default, a system based on the lane aggregation of 56Gbps PAM4 serdes may be the ideal physical underpinning for 50GbE, 100GbE, 200GbE, and 400GbE ports—as well as proprietary ports at 800Gbps based on a 16x50 aggregation scheme.



PAM-4's Impact on Test

In early 2017, Spirent launched the first test system in the market for 200G, based on the use of 4x50GbE PAM4 technology. The use of lane aggregation based on PAM4 also will impact network emulation platforms dedicated to impairment testing for longer-reach high-speed Ethernet, such as Spirent's Attero-100G.

Longer term, knowledge of PAM4 will be important for PHY-to-MAC emulation in any Spirent service-provider tool. Dedicated links longer than 10km may be limited to private DCI over the next two to three years, but full seven-layer OSI testing between OTN and traditional Ethernet may be a logical follow-on to today's Carrier Ethernet, as future IEEE standards emerge.

NRZ has been synonymous with Ethernet line coding for more than 20 years, since the early days of 100-Mbit Fast Ethernet. Today, the Ethernet chip and system industry is undergoing a fundamental change to PAM4. This line code is not only the basis of a 50Gbit Ethernet link that forms the basis of aggregated lanes for faster speeds. It also will be the most likely line code for future 112Gbps serdes, the building block for single-lambda 100Gbps lanes that still are in the very early stages of design.

Contact Us

For more information, call your Spirent sales representative or visit us on the web at www.spirent.com/ContactSpirent.

www.spirent.com

Americas 1-800-SPIRENT
+1-800-774-7368 | sales@spirent.com

US Government & Defense
info@spirentfederal.com | spirentfederal.com

Europe and the Middle East
+44 (0) 1293 767979 | emeainfo@spirent.com

Asia and the Pacific
+86-10-8518-2539 | salesasia@spirent.com

Rev B | 09/18