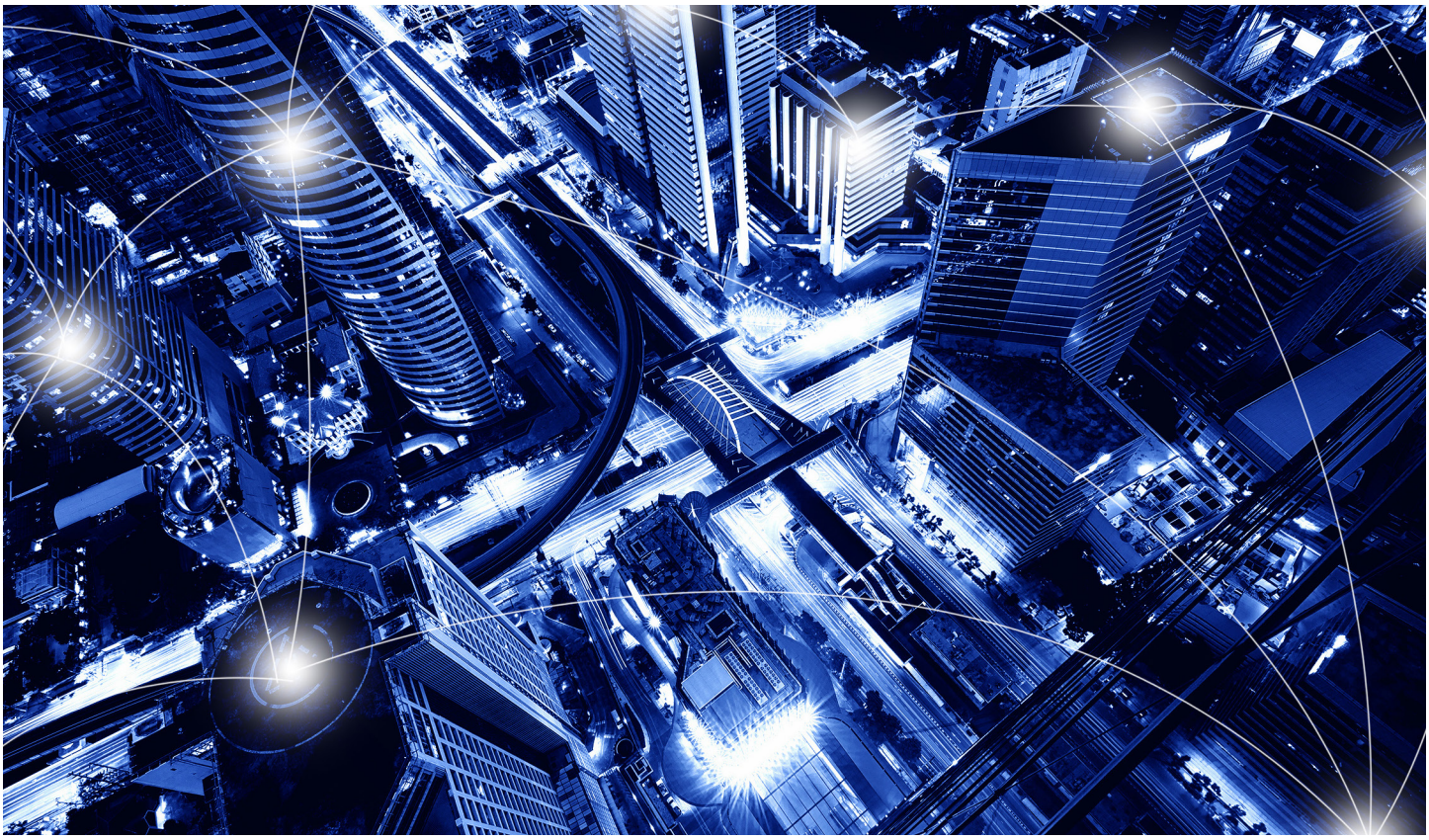
White Paper



FEC for Ethernet:

A New Path to Flexibility

Introduction

The IEEE's task force on 200/400 Gigabit Ethernet has issued a standard that specifies a single Forward Error Correction code, Reed-Solomon (544,514, 10), based on the use of PAM4 line coding. This might appear to be a limiting factor for FEC implementations of the future, except that the group also has proposed a new sublayer that allows flexible use of end-to-end FEC or segment-by-segment FEC without underlying changes to Ethernet's PHY or MAC layers. For the first time in Ethernet's history, developers can experiment with proprietary FEC while retaining compatibility with Ethernet standards.

The "extender sublayer," or CDXS, sits between the media-independent interface (MII) and physical attachment unit interface (AUI), giving hardware designers more freedom to experiment with error-correcting codes. The test community must be cognizant of how this could change future network emulation, however.

Spirent's 400/200/100/50GbE quad speed-test modules were first to market, and have been designed to support the mandatory FEC/PCS IEEE requirements.

FEC for Ethernet:

A New Path to Flexibility

FEC's Belated Relevance to Ethernet

Forward Error Correction (FEC) has a history exceeding 50 years in magnetic storage, optical storage, and optical transmission. Unlike simpler block codes which can correct for single errors, FEC provided a correction technique for burst and random errors, and provides a method for limiting errors in data communications. For more than 25 years, FEC was not considered a necessary element for Ethernet networks, because of the 32-bit **Cyclic Redundancy Check (CRC)** used within the frame structure of 10/100 Mbit Ethernet. DSP chips developed to support FEC in storage and long-haul optical transport applications made a fairly simple adjunct to networks, and were adopted by Ethernet task forces, beginning with 10Gbps, where FEC was listed as an option.

Gigabit Ethernet offered an implicit FEC in its 4D5PAM trellis code for copper transport, and the fiber version of the standard used an 8B/10B block code for timing, that required no FEC. The 10 Gigabit Ethernet standard did not specify a mandatory FEC, though the International Telecommunications Union adopted a Reed-Solomon FEC for the Optical Transport Network, a FEC also used for proprietary 10Gbps network extensions for Ethernet. Specialized 10G EPON and 10G backplane standards added Reed-Solomon FEC support through "clauses" in Ethernet standards. The FEC blocks typically had dimensions in the 200s of units, less than half the block size of Reed-Solomon codes today.

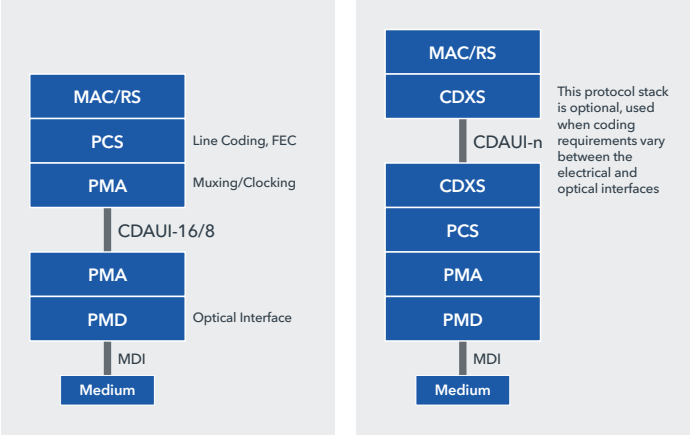


Figure 1. The Protocol stack with FEC.

The 802.3bj 100Gbps backplane and copper cable task force was the first to recognize the necessity of FEC, particularly for such high signaling rates carried over copper. This group specified "Clause 91" Reed-Solomon FEC layer, typically implemented between two Physical Medium Attachment (PMA) sublayers, or between a PMA and a Physical Coding Sublayer. Clause 91 was adopted as a baseline concept for all subsequent 802.3 standards requiring high speed and very powerful error correction capability. Early drafts of 802.3bj allowed 64B/66B block-encoded data to be sent without FEC, but the task force decided to make FEC use mandatory in the transmitter, even if receivers did not always correct all errors. The backplane PHY based on NRZ used a Reed-Solomon block size of (528,514), while the backplane using PAM4 featured a block size of (544,514). Given PAM4's new status as a modulation building block for faster Ethernet speeds, Reed-Solomon (544,514) is considered the FEC most likely to be standardized in DSP silicon across applications.

Clause 91 became part of the baseline for several more standards, including: 100GBase-CR4, specifying four lanes of 25GbE over direct attach copper up to 5 meters; 100GBase-KP4, applying four lanes of 25GbE with PAM4 modulation over a backplane with insertion loss of up to 33dB at 7GHz; and 100GBase-KR4, which uses four lanes of 25GbE with PAM2 modulation over a backplane for insertion loss of up to 35 dB at 12.9GHz. In all three cases, RS-FEC was a requirement, and additional 802.3 standards are expected to add the baseline FEC, even though it can add as much as 250ns latency in some systems. The 802.3bm task force for fiber optic 40- and 100-Gigabit Ethernet standards also has adopted Clause 91.

Additional clauses relevant to high-speed FEC include Clause 119, specified for the PCS for 200GBASE-R and 400GBASE-R using 64B/66B encoding. PHY/PCS data coded in 64B/66B is transcoded to 256B/257B to allow for the addition of FEC in 200/400GBase-R implementations. Clause 119 is typically considered a 64B/66B "spinoff" of Clause 91. In addition, Clause 108 defines Reed-Solomon codes for a 25Gbps MAC, while Clause 74 defined parameters for "fire code" FECs.

The 802.3by 25 Gigabit Ethernet task force wanted to give designers the option of implementing lower-latency links for special purposes, and allowed auto-negotiation between full Clause 91 FEC and an earlier Clause 74 BASE-R or Fire Code mode. The latter is designed specifically for recovery from burst errors in a backplane environment, where minimum latency is required.

Extenders for FEC Flexibility

The notion of allowing vendors to comply with baseline FEC while adding higher-performance proprietary FEC is nothing new in applications such as video transcoding or wireless LANs. In 802.11ac WLAN access points, for example, Broadcom has offered its own "TurboQAM" to support higher-order QAM constellations with proprietary FEC, while still meeting access point interoperability standards.

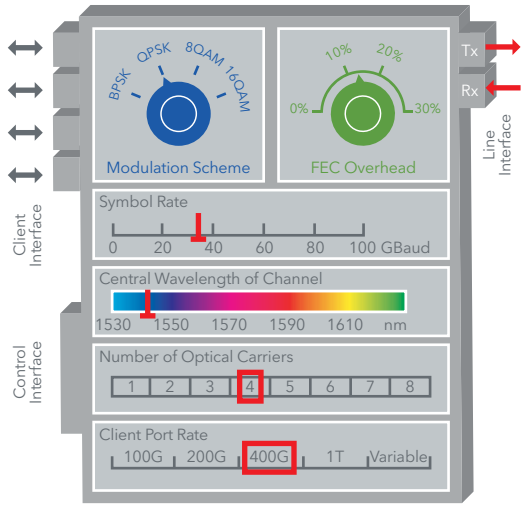


Figure 2. Configuring for FEC Overhead at 400G.

In order to accomplish this in Ethernet, a slight re-architecting of the MAC-to-PHY interface was required. In preparing the drafts for the 802.3bs 200/400 Gigabit Ethernet standard (expected to be approved by 4Q17), participants defined what they call a CDXS, or extender sublayer, referred to in some documentation as 400GXS (in faster Ethernet standards, Roman numerals are commonly used for higher speeds; thus "CD" means 400, while "XS" stands for the extender sublayer). This layer sits between the Media-Independent Interface, or MII, and the physical-layer Attachment Unit Interface, or AUI. IEEE has defined an eight-lane and 16-lane AUI for 400GbE, dubbed CDAUI-8 and CDAUI-16, respectively, but the CDXS (or 400GXS) sublayer allows vendors to experiment with new Physical Coding Sublayer (PCS) codes and new FEC methods. CDXS also would allow easier experimentation with line codes—NRZ, PAM4, PAM8—although PAM4 is anticipated to be the de facto standard for both 56Gbps serdes and possible 112Gbps serdes of the future.

Once the 200/400GbE standard is approved, developers could retain an existing PCS while adding new CDAUI options, or retain existing CDAUI-8 and CDAUI-16 options, while experimenting with new PCS that might offer new FEC types. In the past, the PCS and AUI of new Ethernet speeds always proceeded in lockstep. Now, they can be developed in different implementations using CDXS as the great equalizer.

In a new CDXS architecture, strong block FEC codes such as Bose–Chaudhuri–Hocquenghem (BCH) FEC could be used when an application demanded multiple errors being corrected simultaneously, even as decoding becomes a simpler algebraic process compared to other FEC methods. In the past, inserting of BCH FEC might require considerable changes in a PCS layer that would make an Ethernet interface incompatible with existing devices.

Test and Emulation in a CDXS and FEC Environment

In theory, an extender sublayer should ease both emulation and physical-layer validation of high-speed Ethernet, since a network node retains seven-layer OSI interoperability even as new PCS and AUI options are tried by specific vendors. But network managers should not underestimate the new normal of speeds in excess of 100Gbps, based on lane aggregation of 50Gbps lanes with new types of line code. IEEE has discussed single-lambda 100Gbps links that also would use PAM4 coding, so if optical transceivers become economically feasible, PAM4 encoding still would be relevant. PAM4 links will have greater error floors than ever before, with those error floors corrected through FEC—usually Reed-Solomon, though occasionally special block codes.

All testing in such an environment will be nonlinear, and will require a deeper understanding of error statistics and FEC characteristics. Framed emulated traffic provides a good baseline for performance, but tools must validate the margin of individual elements. Test regimes also must offer visibility in multiple domains. Physical layer testing must allow for the existence of FEC between AUI and MII. To understand skew, lane rate, and pattern sensitivity, tests must operate across multiple domains, with full post-FEC and pre-FEC monitoring of Ethernet-layer framed traffic.

About Spirent Communications

Spirent Communications (LSE: SPT) is a global leader with deep expertise and decades of experience in testing, assurance, analytics and security, serving developers, service providers, and enterprise networks.

We help bring clarity to increasingly complex technological and business challenges.

Spirent's customers have made a promise to their customers to deliver superior performance. Spirent assures that those promises are fulfilled.

For more information, visit: www.spirent.com

The chip, optical module, and line card elements present in this new flexible generation of high-speed Ethernet will further complicate testing. Advanced modules such as QSFP-DD, OSFP, and CFP8 will offer eight lanes of 50Gbps optical-electrical interfaces. While FEC will most commonly be implemented in ASICs or FPGAs on NICs or line cards prior to I/O, unique cases may call for implementation of FEC within an optical module. As the COBO Alliance works on on-board optics, the diversity of such implementations could increase exponentially.

In the new world, new PHYs, MACs, and sublayers are inserted into existing architectures, and optional FECs become one more modular option to take advantage of the new 400GbE standard. Operating full line rate 400GbE traffic translates into approximately 600 million packets per second of traffic. By utilizing Spirent's broad range of 400/200/100/50GbE quad speed test modules customers are able to test FEC error correction and statistics at line rate.

spirent™
Promise. Assured.

Contact Us

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

www.spirent.com

© 2018 Spirent Communications, Inc. All of the company names and/or brand names and/or product names and/or logos referred to in this document, in particular the name "Spirent" and its logo device, are either registered trademarks or trademarks pending registration in accordance with relevant national laws. All rights reserved. Specifications subject to change without notice.

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