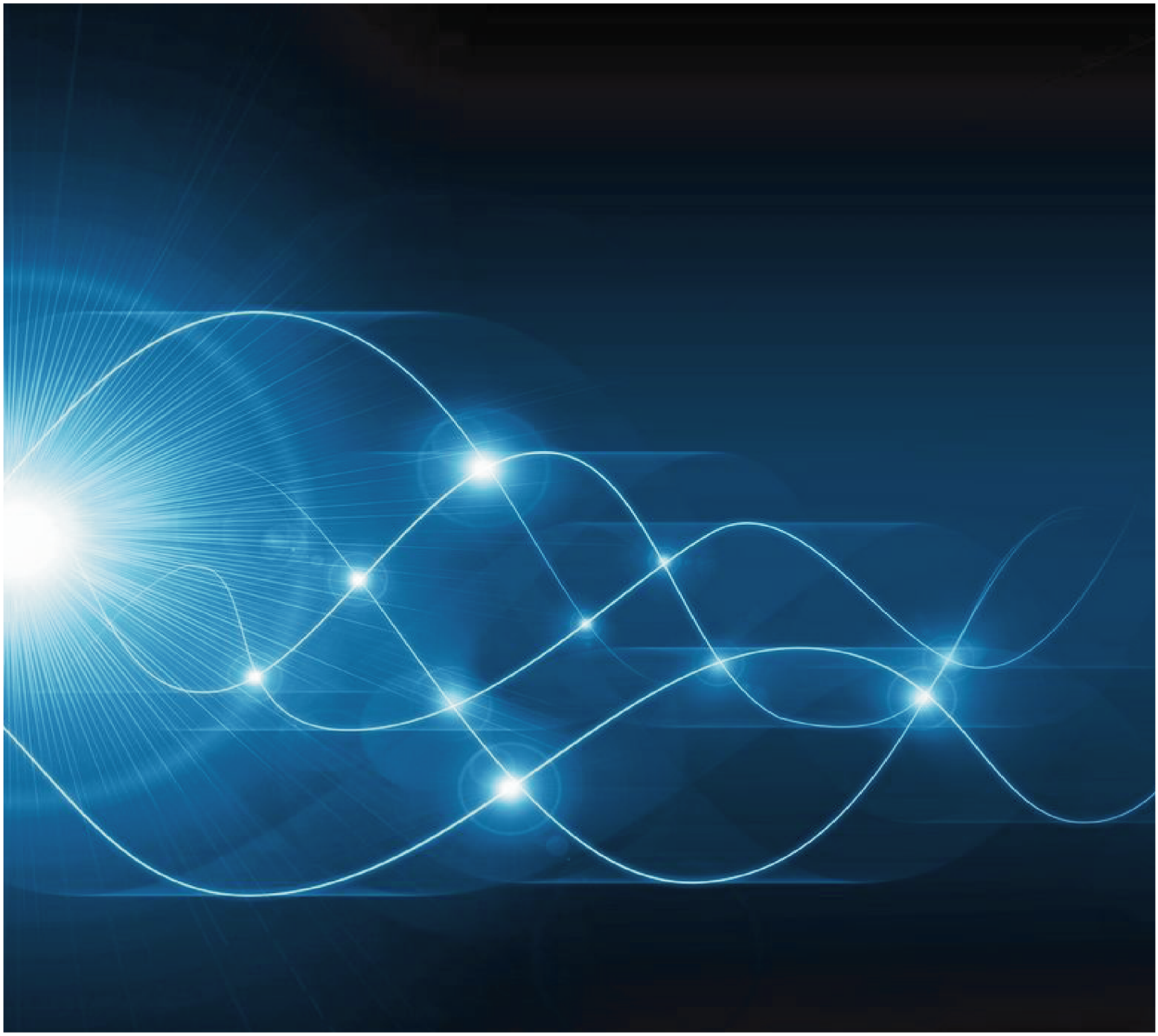
multiLane

Innovation for the next generation

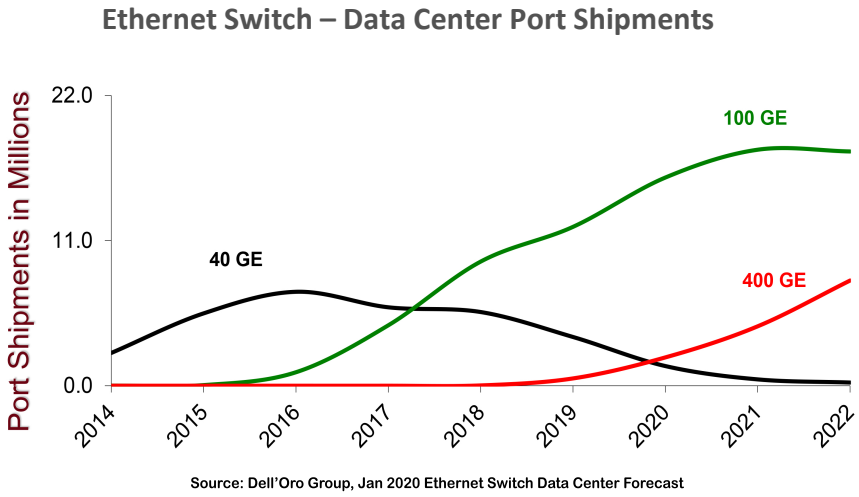
FEC Solutions
for the 400G Frontier

White Paper



Forward Error Correction: No Longer Optional

The transition to 400G ports has prompted the high-speed IO industry to embrace a paradigm shift. Demand for improved bandwidth from cloud customers and service providers shows no sign of slowing down. The projected demand growth of 400GE ports is made evident in the graph to the right as 100GE shipments begin to slow down and 40GE continues to be phased out. To align with this new frontier, semiconductor and optic vendors have explored new methods to enable 400G interconnection while overcoming the physical limitations of today's optoelectronic components.



This evolution has been driven forward by Forward Error Correction (FEC) encoding schemes. Once seen as an optional implementation to stretch out link budgets in the 100G era, FEC has emerged as a mandatory mechanism to successfully close 400G links. Enabled by a fresh wave of standards publications, solution vendors can now go to market with chip architectures that operate at high error rates by design.

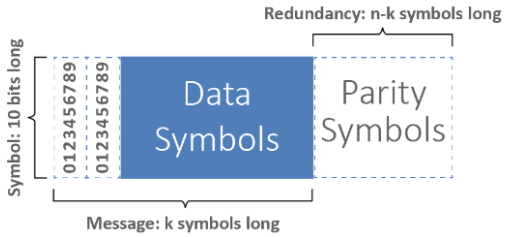
This new requirement has driven Digital Signal Processor (DSP) vendors to produce module-based solutions that maintain a delicate tradeoff between hardware complexity, packet latency and coding gain across a link while abiding by a critical power consumption envelope. This evolution necessitated a departure from traditional Non-Return to Zero (NRZ) signalizing in which a signal can occupy either a '1' or a '0' state. Pulse Amplitude Modulation (PAM4) has arrived as the optimal transmission scheme resulting in 100G capacity per unique wavelength over the line side.

The upgrade from traditional 25G wavelengths to 100G is facilitated by PAM4 signals transitioning in intervals half the size of NRZ with twice the number of possible information states (Four levels in the signal as opposed to a traditional two levels). To address these novel complexities, vendors require comprehensive analysis tools to capture both the deterministic and random error behavior of various noise sources.

As interconnect vendors freeze their designs and begin ramping up for production volume in 2020, system integrators have voiced an urgent need to measure Real FEC statistics. This white paper presents an overview of MultiLane's capability to deliver high-value FEC solutions for every stage of the development cycle.

Technical Overview of FEC

In its essence, Forward Error Correction is a process that appends a string of data with additional bits that allows a certain rate of bit errors to be detected and subsequently corrected. An optimized approach to this method adds sufficient performance margin to a link while increasing link latency and system power consumption to a limited extent. This section will provide an overview of FEC implementation in 400G transmission today.

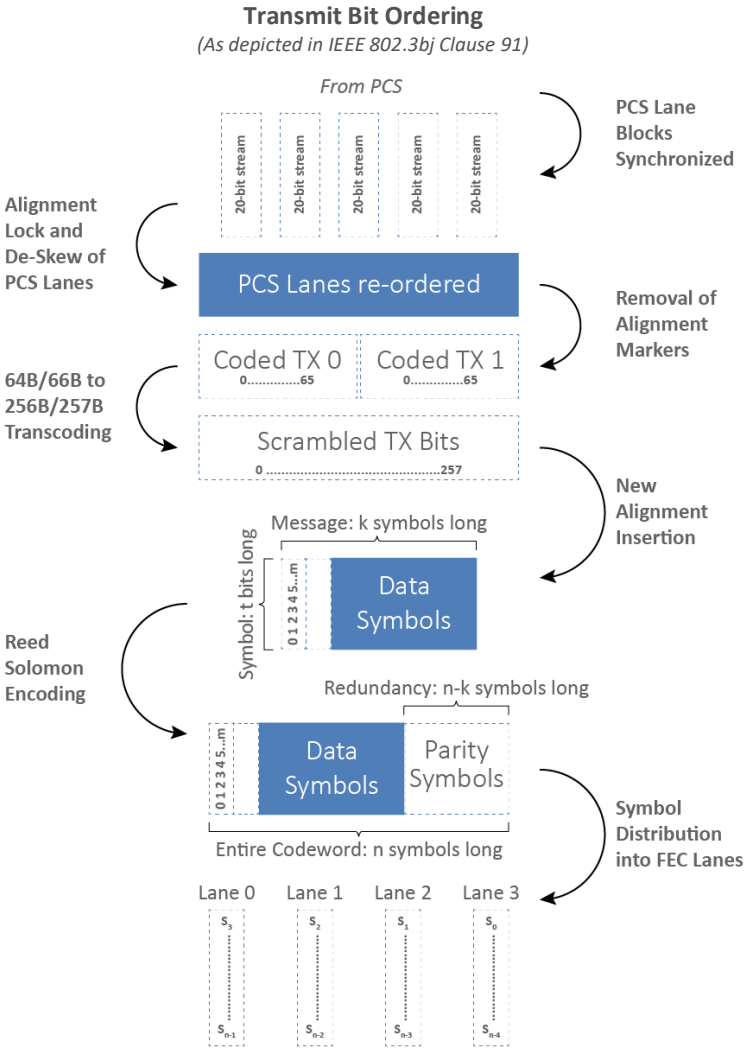


FEC Statistics vs TX Performance

The optics industry has invested significant attention towards characterizing 100G-per-lambda transmitters with the Transmission and Dispersion Eye Closure Quaternary measurement (TDECQ). While TDECQ does give an accurate measurement of TX performance relative to an ideal transmitter, it is still the subject of some debate. Optimal TDECQ performance does not directly correlate to optimal Bit Error Rate (BER). To that effect, BER measurements associated with Forward Error Correction encoding schemes have emerged as a key figure of merit for module link performance and system characterization.

Construction of a FEC Block

Within the RS-FEC Sublayer of the PHY environment, a bitstream flows in from the Physical Coding Sublayer (PCS). This data is formed into 20-bit streams which are subsequently arranged into 66-bit blocks by use of synchronization headers, enabling block lock. The next step is obtaining a first alignment lock, which enables PCS streams to be de-skewed and have their alignment reordered. After this bitstream is transcoded from 64B/66B to 256B/257B encoding, scrambled bits are inserted with alignment markers to arrange the message into a format that groups bits into symbols.



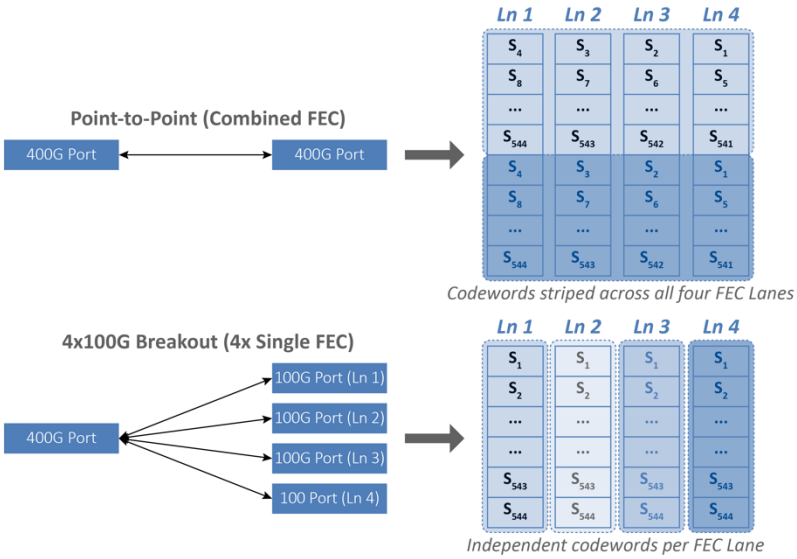
These aligned symbols are then passed through a Reed-Solomon encoder.[1] Symbols from the assembled codewords are then arranged into FEC lanes in round robin format. For the scope of this document, a simple block diagram will suffice.

After the encoded bitstream traverses through the signal path via the physical medium attachment (PMA), the receiver circuit essentially inverses this process. Blocks are deconstructed in the Reed-Solomon decoder, where codewords are corrected where possible. If and when the maximum number of correctable symbols in a codeword are exceeded, the decoder indicates that the specific codeword is uncorrectable. The conditions that deem a block correctable or uncorrectable will be highlighted later in this document.

[1]Details pertaining to the RS encoding approach can be found in Clause 91 of the 802.3bj publication.

FEC Striping

The nature of FEC striping is defined by the interface implementation approach. Network operators can choose to break out a 400G module (KP4 FEC) into four 100G ports (4 x KP1 FEC) or simply stick to a 400G point-to-point link. Depending on the preferred switch implementation, encoded symbols will be striped into FEC lanes differently. While a traditional point-to-point 400G link would employ round robin symbol arrangement across the four FEC lanes, a breakout link requires each FEC lane to operate as an independent interface. The right figure elaborates with a simplified KP FEC example.



KP4 and KR4 FEC Definitions

We will address two fundamental Reed-Solomon encoding types, KP for PAM4 signals and KR for NRZ signals. These encoding schemes have been widely adopted for 100G-per-wavelength optics due to their limited latency hit and ability to correct both random and burst errors while adding considerable performance margin to module link budgets. In terms of applicable use cases, KR FEC is applied to NRZ signals for traditional 100G ports. This translates to grouping 4 lanes of 25Gb/s into a 100GBASE-KR tunnel (DR1, FR1[2], etc.). On the other hand, KP FEC is applied upon PAM4 signals to group electrical lanes of 50Gbit PAM4 into one to four 100GBASE-KP tunnels (DR4, FR4, etc.). While KR1 FEC is applied exclusively to 100G ports, KP can be applied both to 100G (KP1) or 400G (KP4) ports.

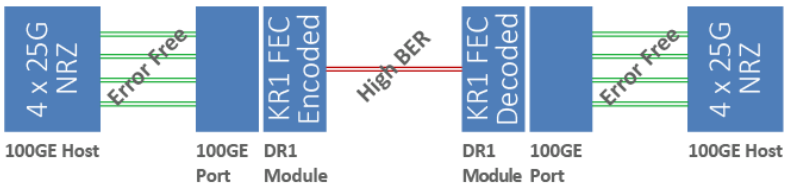
General Reed-Solomon FEC Format: RS (n, k, t, m)		KP PAM4 FEC: RS (544, 514, 15, 10)	KR NRZ FEC: RS (528, 514, 7, 10)
n	Total size of FEC codeword (CW) in symbols	544 Symbols	528 Symbols
k	Total size of message in the CW, in symbols	514 Symbols	514 Symbols
t	Max number of correctable Symbols per codeword t = (n-k)/2	15 Symbols (max)	7 Symbols (max)
m	Symbol size – Number of bits per symbol	10 Bits	10 Bits
n-k	Redundant length, total size of parity bits in the CW	30 Symbols	14 Symbols
NCG	Net Coding Gain – Link margin improvement	~6.9 dB *(IEEE Publication)*	~5.7 dB *(IEEE Publication)*
R	Code rate, fraction of CW that is non-redundant – R = k/n	514/544 ≈ 0.945	514/528 ≈ 0.973
d_{min}	Min. distance between two valid codewords – $d_{min} = (2*t) + 1$	31 Symbols (max)	15 Symbols (max)

A FEC scheme is primarily identified by the highest number of correctable symbols per codeword block and the resulting net coding gain (NCG). NCG is an estimation of the performance improvement in dB added by the encoding process. Another note here is that the 't' value in these FEC codes can be derived from the minimum distance, d_{min}. This metric is an indicator of the minimum number of symbol transitions required to alter one valid codeword into another. Minimum distance is a key determinant towards the error correction capability of a FEC code.

FEC Implementation at the Port

Modern DSP chips have been designed to enable versatile encoding and decoding of FEC traffic within the architecture of a transceiver. While this is a crucial feature in some implementations, it adds unnecessary latency in others. Two primary use cases make this evident.

[2] DR Optics are specified for 500m reach across single-mode fiber, while FR optics target 2km for that same medium.



Case 1: 100G DR1 Link in Legacy Switch

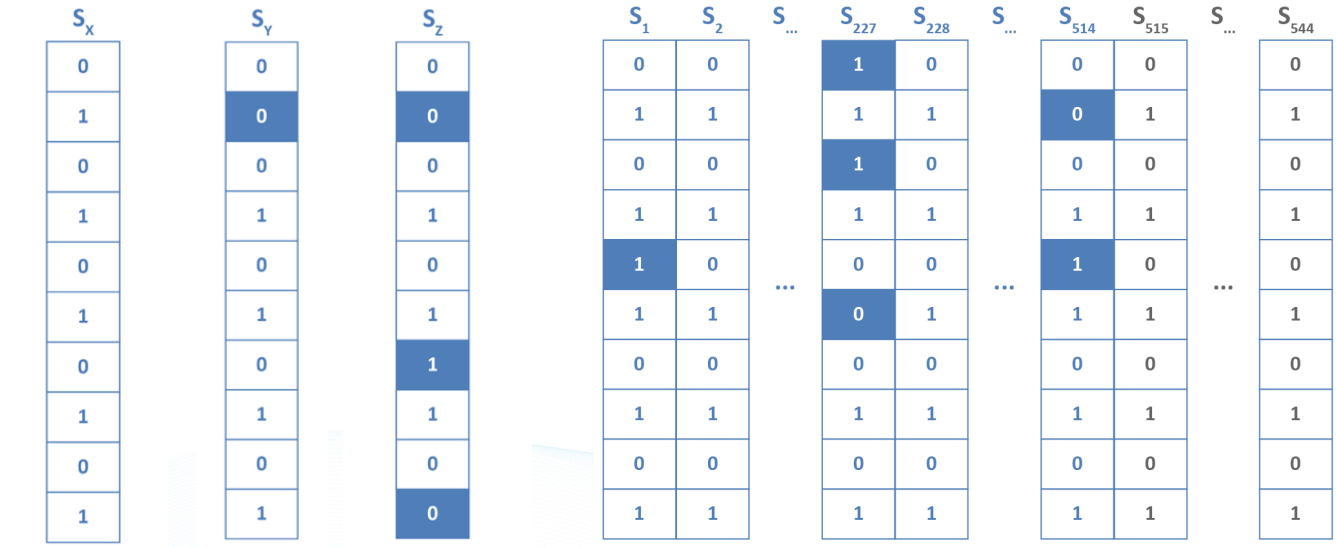
Case 1 depicts a direct DR1 link across legacy 100GE host ports. DR1 optics can leverage existing 4x25G NRZ ports for single wavelength fiber transmission. As this optic has the same electrical path as traditional 100G pluggables, traffic typically passes error free from host to module port. However, due to the high-BER nature of DSP designs, FEC must be implemented in the module chip itself. In this case, KR1 FEC is applied. This process is subsequently reversed on the remote end. The module DSP decodes the FEC frames, thus removing the parity bits and delivering the original 4x25G traffic stream back to the remote host.

Case 2, on the other hand, is more straightforward. This figure covers a point to point 400GE link between two DR4 modules. While the 8x50G PAM4 electrical signal driven from the host to the module port does incur a low rate of errors, it is negligible compared to the high BER that occurs within the module components. To that effect, Ethernet traffic is encoded within the host ICs on either side of the link. Due to latency and power dissipation concerns, it is considered impractical to decode and re-encode FEC within the 400GE transceiver itself.

Case 2: 400G DR4 Link in Modern Switch

Codeword Errors

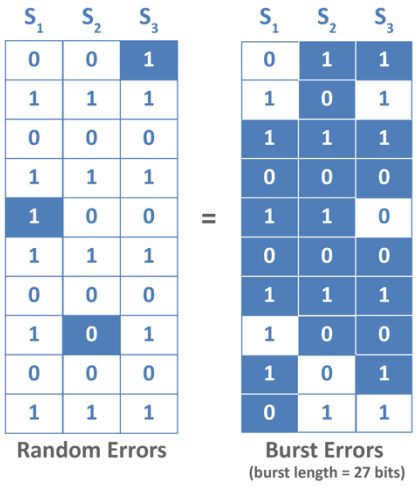
As has been touched upon earlier, a FEC codeword is primarily defined by the number of symbols that can be corrected within it before the block is lost. The following figures differentiate between traditional bit errors and symbols errors, along with the conditions that make a word uncorrectable:



As depicted previously, a symbol can only be correct if all bits that comprise it are correct. Whether a single symbol has ten errored bits, or just a single errored bit, the result is a single symbol error. The *t* value in a FEC encoding scheme defines how many of those errored symbols can be corrected per word. Sticking to our KP example, a codeword can only be recovered if *t*=15 or less of its symbols are errored. Those 15 symbols can each contain single errors or complete loss; it makes no difference to the FEC decoder, the codeword will be corrected and preserved. The ability of this encoding mechanism to recover lost symbols is critical for applications within 100G-per-wavelength interconnects due to the various types of signal noise that can cause both random and burst errors. This will be discussed in the following section.

Error Types and Noise Sources

The crucial nature of Reed-Solomon error correction is made evident by the high error rates that occur in 100G-per-lane circuits by design. Errors can be sorted into two primary categories; random and bursts. The IEEE classifies the highest correctable rate of these errors in PAM4 optical links as 2.4E-4[3]. The occurrence of random (uncorrelated) errors satisfies the random distribution curve; they arise independently of each other and are typically caused by random noise within the channel. Burst errors, on the other hand, resemble a consecutive series of bits where the first and last bits in the burst are incorrect, caused by undesired yet deterministic behavior within a transmission circuit.



Primary noise sources that can attribute to random errors include fluctuations in the DC power level delivered to active components along with the cross talk and insertion loss that occur due to marginal signal integrity from wire bonding. While the limitations of lossy signal traces are indeed compensated by a variety of equalization techniques, the techniques themselves can be responsible for the large strings of errored bits we identify as bursts.

Equalization Techniques and Error Bursts

EQ (Equalization) approaches are implemented both at the transmit and receive sides of an active circuit to compensate for loss within the channel. While a TX-side equalizer can apply signal emphasis to an outgoing signal, an RX side equalizer typically applies CTLE (Continuous Time Linear Equalization) or FFE (Feed Forward Equalization) for linear gain or DFE (Decision-Feedback Equalization) for adaptive gain.

Directing focus to the DFE approach, we must discuss a dominant source of signal noise known as Intersymbol Interference (ISI). ISI, caused by channel dispersion, is a form of signal distortion in which one symbol state will interfere or corrupt the states of adjacent symbols. A DFE is more effective than linear equalizers to compensate for this phenomenon, since it uses past symbol decisions to eliminate the ISI affecting the decision of the current symbol.

Often, the DFE circuit applies this feedback compensation upon the incorrect symbol due to effects from jitter (signal uncertainty due to fluctuation across the time axis). This results in a propagation of errors (an error burst) in the RX circuit since a string of bits improperly equalized. As the frequency of burst errors is high due to the dependency on DFE circuits, the ability of RS codes to correct said bursts is highly valued.

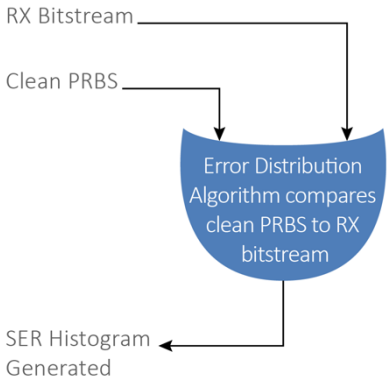
[3] A maximum of 24 errored bits per 100,000 overall bits; see IEEE 802.3ck.

FEC Testing with MultiLane

The critical role filled by Reed-Solomon Forward Error Correction in closing 100G-per-lane links has been plainly outlined. As both random and burst errors are typical in 400G component architecture, the test and measurement sector must provide vendors with solutions to certify their designs. This performance visibility can be realized with the full FEC measurement suite supported by MultiLane Bit Error Rate Testers. With a comprehensive family of 400G and 800G instruments enabled by the brand new ThunderBERT software, users are equipped with Real and Emulated FEC performance data to validate their designs.

FEC Emulation – Quantifying Performance Margin

The FEC Emulation suite offered by select MultiLane BERTs offers an accurate estimation of DUT performance via a comprehensive error distribution algorithm embedded in the design architecture. As opposed to traditional FEC mechanisms, this feature drives an unframed Pseudo Random Bit Sequence (PRBS) through the DUT circuit. The emulator then receives this input bit stream and identifies the location of each bit error by comparing them to the original correct PRBS. The algorithm then compiles an assessment of the post-FEC Symbol Error profile by analyzing the error distribution based on the selected FEC striping method.

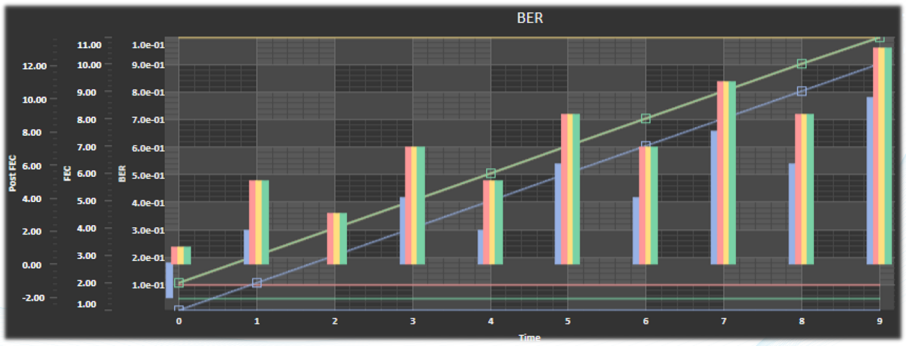


Real FEC – Replicating Host Environments

Real FEC measurements employ the same bit stream realignment, encoding and decoding approach executed in actual 400GE switch ASICs. DUT performance can be characterized in a process that replicates the FEC block construction approach outlined in the IEEE 802.3bj standard literature. Users can leverage this capability by acquiring an exact assessment of their DUT's ability to overcome the turbulent error behavior prevalent in modern optics. Equipped with a hardware-based Reed-Solomon decoder, post-FEC link performance is made available in real time.

ThunderBERT – Precise Performance Tracking

The MultiLane BERT graphical user interface has been comprehensively overhauled. Already recognized as a prevalent software tool in the test and measurement industry, ThunderBERT is equipped with a brand-new minimalist look and feel, intuitive navigation and enhanced measurement capabilities. Measurements like signal-to-noise ratio (SNR) and post-FEC symbol error rate (SER) can now be captured in instantaneous and cumulative modes simultaneously across all BERT lanes. The right graph shows multiple BER plots over time in both acquisition modes.



FEC Measurement Overview

This section will highlight the FEC-related measurements supported by MultiLane's BERT lineup. ThunderBERT's fast data acquisition capability enables each of the following measurements to be depicted both in instantaneous[4] and cumulative modes:

Real FEC Suite

Pre-FEC Bit Error Rate
Raw and unframed ratio of incorrect bits (Bit Errors/Total Bits) on a channel-by-channel basis.

Corrected Ones Count
Number of "1" bits that were corrected to be "0" bits after decoding.

Corrected Zeros Count
Number of "0" bits that were corrected to be "1" bits after decoding.

Corrected Bit Count
Sum of corrected "0" and "1" bits after decoding.

Processed Codeword Count
Total number of codewords, correctable and uncorrectable, processed by the decoder.

Corrected Codeword Count
Number of codewords (FEC blocks) that were corrected by the decoder. In other words, this is the number of codewords that were determined to have less than or exactly t[5] symbol errors.

Uncorrected Codeword Count
Number of codewords that were deemed uncorrectable by the decoder. In other words, this is the number of codewords that were determined to have more than t symbol errors.

Uncorrected Codeword Rate (Frame Loss Rate)
Number of uncorrected codewords divided by the number of processed codewords.

Symbol Error Count
Total number of symbol errors processed by the decoder.

Corrected Bit Rate
Number of corrected bits divided by total number of received bits.

Symbol Error Rate
Number of symbol errors divided by the total number of processed symbols.

Symbol Error Rate Histogram
Breakdown of processed codeword symbol error distribution. Blocks are sorted into different buckets depending on how many symbols errors they incurred.

Post-FEC Bit Error Rate
Total number of bit errors remaining *after* real FEC decoding, divided by the total number of received bits.

Emulated FEC Suite

Pre-FEC Bit Error Rate
Raw and unframed ratio of incorrect bits (Bit Errors/Total Bits) on a channel-by-channel basis.

Post-FEC Bit Error Rate
Total number of bit errors remaining *after* emulated bit correction, divided by the total number of received bits.

Error Count
Total number of PRBS errors captured by the error detector for a specific test duration.

Corrected Errors
Number of bits that the error distribution algorithm in the FEC emulator deems to be correctible.

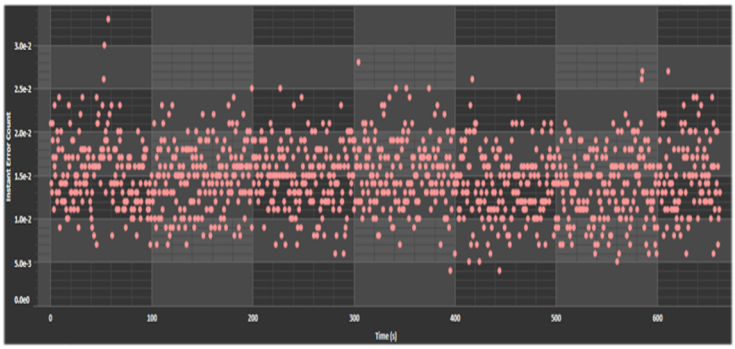
Block Count
Total number of bit stream blocks (emulated codewords) processed by the emulation algorithm.

Saturated Block Count
Total number of emulated codewords that were determined to be uncorrectable based on the distribution algorithm implemented by the FEC emulator.

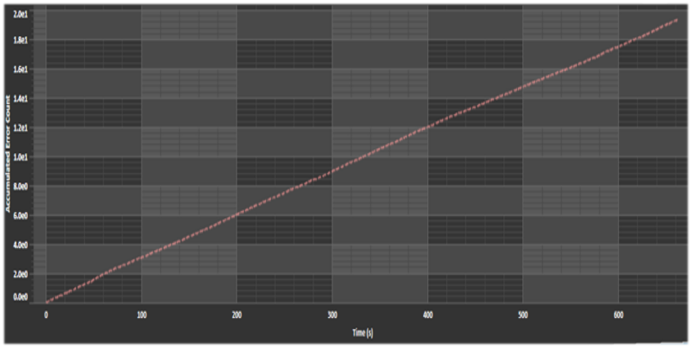
Symbol Error Rate Histogram
Calculated distribution of bit stream blocks sorted by number of symbol errors per emulated block.

[4] Instantaneous measurements are captured every 200 milliseconds
[5] See *KP4 and KR4 FEC Definitions* section

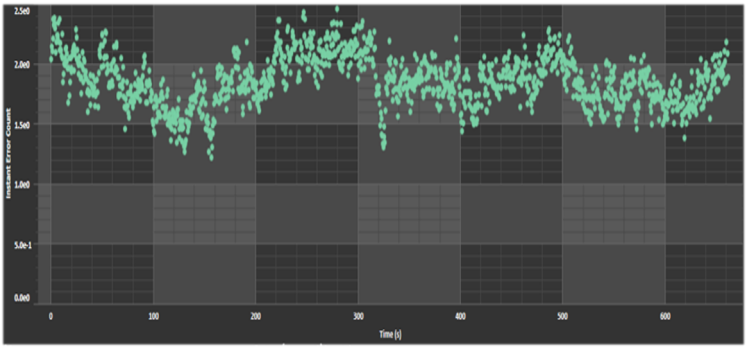
Sample ThunderBERT Measurements



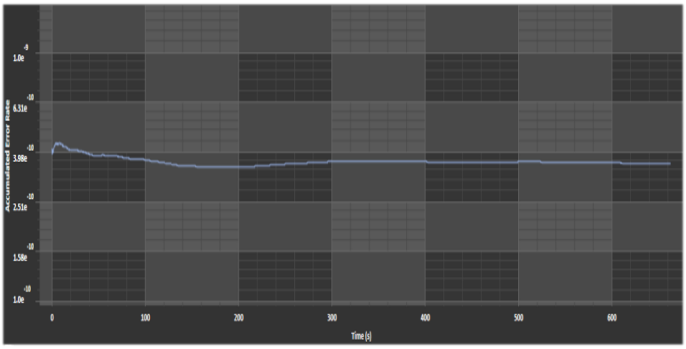
Instant Corrected Zeros Count Vs. Time



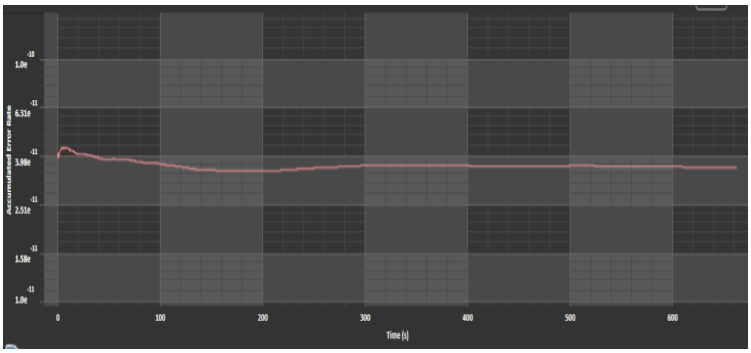
Accumulated Zeros Count Vs. Time



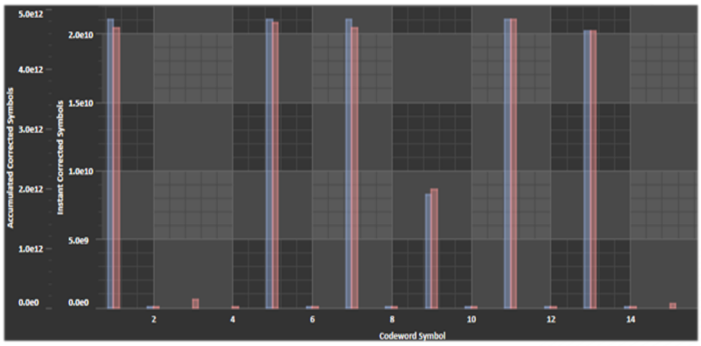
Instant Symbol Error Rate Vs. Time



Accumulated Symbol Error Rate Vs. Time



Accumulated Post-FEC Bit Error Rate Vs. Time



SER Histogram vs. Number of Symbol Errors per Codeword

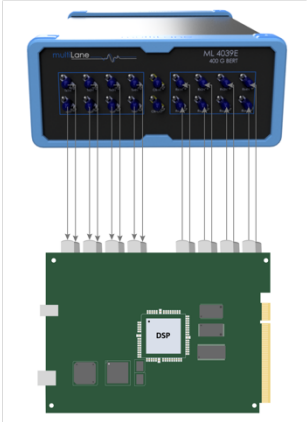
FEC Options with MultiLane BERTs

BERT Platform	Implementation	Supported FEC Striping	Supported FEC Types	Max Supported Rates
ML4039D	FEC Emulator	50G, 100G, 200G Stripes	KP (50G, 100G, 200G, 400G) KR (50G, 100G)	4x28 Gbaud NRZ/PAM4
ML4079D		50G, 100G, 200G, 400G Stripes		8x28 Gbaud NRZ/PAM4
ML4039B-BTP	Real FEC	50G, 100G, 200G Stripes	KP (50G, 100G, 200G) KR (50G, 100G)	4x28 Gbaud NRZ/PAM4
ML4054-400				8x28 Gbaud NRZ/PAM4
ML4039E/EN ML4035		100G, 200G, 400G Stripes	KP (100G, 200G, 400G) KR (100G)	4x56 Gbaud NRZ/PAM4
ML4079E/EN				8x56 Gbaud NRZ/PAM4

Application Examples

This section features a brief application note to enable users to characterize their designs with MultiLane FEC BERTs. While the scope of uses for this product family is comprehensive, we will narrow the focus onto two primary cases.

1- Digital Signal Processor (DSP Chip) Testing with ML4039E



This application enables characterization of host-side DSP performance in a benchtop setting. A user can set a desired line rate, PRBS pattern and equalization taps in the ThunderBERT configuration window and execute the following sample measurements while cycling the DUT through various operational modes:

-Quantify SNR value, histogram and more with REAL FEC disabled

-Capture in-depth FEC measurements like corrected 1's, post-FEC BER, codeword statistics and more with REAL FEC enabled

-Assess DUT resilience to crosstalk noise injection (ML4039EN option required)

2- 400G Transceiver System Testing with ML4054-400

As this BERT is also equipped with a 400G host interface, a user can force a DUT (OSFP or QSFP-DD) into a preferred operating condition, which would in this example be FEC Bypass Mode. After setting transmit characteristics within ThunderBERT and connecting a fiber loopback to the module, a sample of supported test operations follows:

-Track module pre-FEC BER performance with REAL FEC disabled

-Monitor and log SER histogram, codeword statistics and more over time in either instantaneous or continuous captures with REAL FEC enabled

-Test module performance across supply voltage and temperature sweeps

Conclusion

In this white paper, we deliver insight into the role FEC plays towards modern transceiver development in the 400G era. Following an overview of the state of the optoelectronic industry today, we delve into FEC encoding within the PHY environment. This is complemented by a breakdown of FEC types and the noise/error behaviors that they are able to correct. Finally, the latter section provides a comprehensive dive into the Real FEC and Emulated FEC test capabilities offered by MultiLane's instrumentation lineup. Driven by the novel ThunderBERT test suite, MultiLane is the ideal test and measurement provider to equip industry players with solutions to validate and produce their FEC-based products.

Worldwide
Houmal Technology Park
Askarieh Main Road
Houmal, Lebanon
+961 5 941 668

Asia
14F-5/ Rm. 5, 14F., No. 295
Sec.2, Guangfu Rd. East Dist.,
Hsinchu City 300, Taiwan
(R.O.C)+886 3 5744 591

North America
48521 Warm Springs Boulevard
Suite 310
Fremont, CA 94539 USA
+1 510 573 6388