Highly-Resilient 25G Gen-Z PHY

Karl Bois, Chris Brueggen, Elene Chobanyan, Michael Krause, Pete Maroni, James Regan
{karl.bois, chris.brueggen, elene.chobanyan, mkrause, pete.maroni, jregan }@hpe.com

Application-specific use models require different specifications (e.g. memory, storage). A case for high resilience fabric performance while avoiding Forward Error Correction (FEC) is presented. The performance is achieved by limiting channel loss in contrast to full allowable interconnect reach.

What is "COM" and "BER"?

Chanel Operating Margin (COM) is an industry adopted method for validating channel compliance. COM is a single measure of the channel performance encompassing signal distortion analysis and methods to compensate for it. The COM measurement is obtained from simulated or measured channel S-parameters and addresses effects of channel ISI (inter-symbol interference), crosstalk, random jitter and noise. Once transmitter and receiver equalization schemes are applied and an eye diagram at a specified Bit Error Ratio (BER) is obtained, COM is computed as the ratio of the signal amplitude to the vertical eye closure. Using agreed upon transmitter and receiver parameters it is possible to define system performance expectations within the confines of a given channel.

BER is defined as a ratio of number of bits received in error (N_{err}), excluding additional bit errors resulting from DFE (decision feedback equalizer) error propagation, and the total number of bits transmitted (N_{bit}), i.e. BER = N_{err}/N_{bit}. Typical raw BER (BER prior to DFE burst error propagation and FEC application) targets are limited to 1E-4 or 1E-5. This paper will demonstrate that by limiting channel reach, platforms can deliver a FEC-free 1E-15 BER, which at 25G speed translates to approximately one error in 11 hours per lane.

Channel Topologies and Performance

To maximize system design flexibility, the *Gen-Z Physical Layer Specification* does not normatively specify channel frequency masks. Instead, to ensure interoperability and provide designers with a starting point, Gen-Z specifies informative insertion and return loss masks. *Figure 1* illustrates two frequency domain masks (see *Figure 2*) compliant channel examples with insertion losses at 12.89 GHz pushed to the high and low limits. The channel limits are representative of single-enclosure component-to-component topologies, e.g., processor to memory, I/O, storage, and switch components.

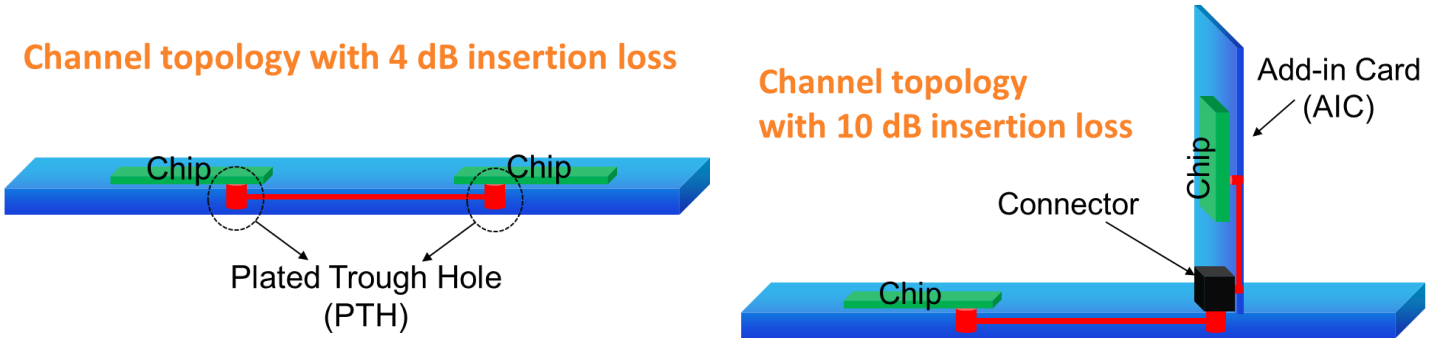

Figure 1: Channel Topologies

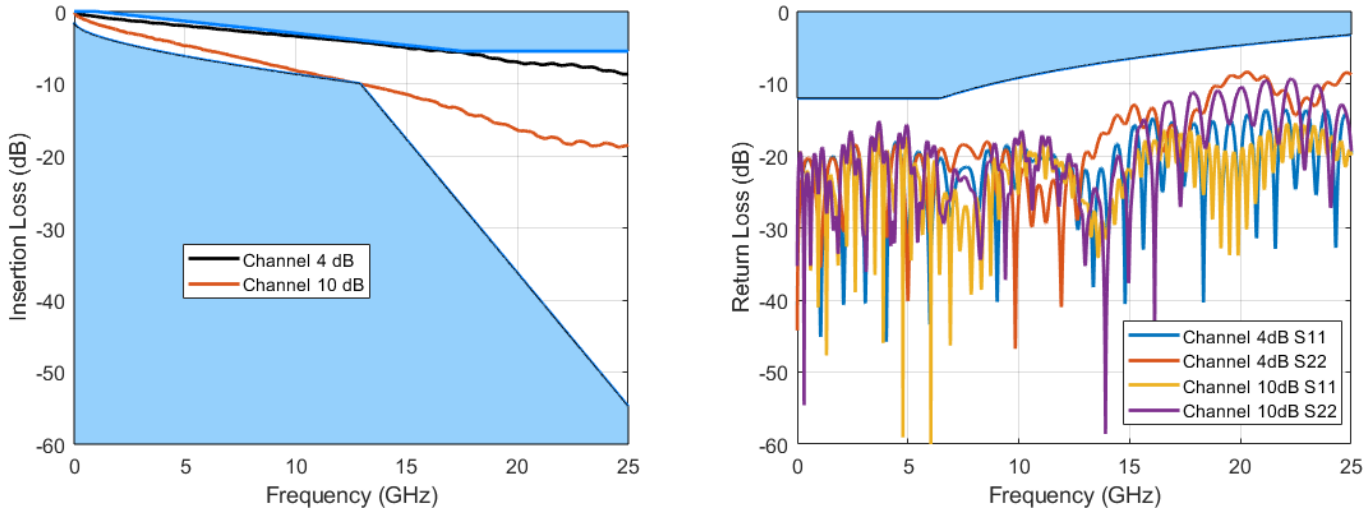

Figure 2: Informative Insertion/Return Losses

For each channel loss, the most complex interconnect topology was chosen. In particular, a channel with a 10 dB loss could have a connector (e.g., SFF-TA-1002 / Gen-Z Scalable Connector) in the path, while a 4 dB loss budget implies a single board local chip-to-chip topology. Time domain simulations to obtain COM values at BER = 1E-15 were performed with two far-end and three near-end aggressors (crosstalk contributors). COM values were computed for five different cases, where the number of DFE is varied from 0 to 4. The rest of the settings for COM computation are specified in the *IEEE Standard for Ethernet 802.3™-2015*, clause 83D.4. The results are shown in *Table 1*.

Table 1. Channel Operating Margin (COM) at BER = 1e-15

	10 dB channel		4 dB channel	
DFE Taps	**Short package**	**Long package**	**Short package**	**Long package**
0	**1.87**	2.563	**1.892**	3.628
1	2.45	3.242	2.134	3.824
2	2.663	3.583	2.168	4.076
3	2.481	3.771	2.748	4.101
4	4.042	3.925	3.475	4.137

The Gen-Z 25G NRZ short reach (aka Gen-Z-E-NRZ-25G-Local) electrical sublayer specification is leveraged from *IEEE Standard for Ethernet 802.3by™-2016* Annex 109A Chip-to-Chip 25 Gigabit Attachment Unit Interface (25GAUI C2C). Therefore, compliant channels can be built to this specification when COM ≥ 2 dB (failing channels with COM ≤ 2 dB are highlighted in bold red in *Table 1*). Further, these results demonstrate that if DFE-based designs are used, then Gen-Z 25G NRZ short reach channels can achieve a raw BER of 1E-15. The combination of a superior raw BER with Gen-Z's strong packet-based CRC and end-to-end and link-local packet retry mechanisms can yield solutions with reliability comparable to a highly-optimized DDR design (single-interface Mean Time to False Packet Acceptance > 1e20 years) with negligible solution-visible performance impacts.