PCB Only Solution for 112Gbps VSR Channels for Ethernet Switches

Riaz Naseer, Amazon
riazrao@amazon.com

Yuri Tretiakov, Amazon
ytretiak@amazon.com

Eric Tran, Amazon
erictran@lab126.com

Abstract

Ethernet speeds of 112Gbps/lane pose significant challenges for designing switch systems to meet increased throughput demands of today's data centers with high traffic volumes from artificial intelligence (AI) servers. Due to increased PCB losses at such high frequencies, fly-over cables are being promoted for the longest length channels. Fly-over cables pose manufacturing reliability and mechanical challenges for the system design; along with increased system cost. System designers prefer to have PCB only solution. Development in PCB materials and current data rates have brought us to a stage where copper losses dominate than the dielectric losses. In this scenario, copper-roughness modeling becomes even more important than before. When it comes to Cu-roughness modeling, there are a lot of parameters that are used in datasheets. Furthermore, there are a lot of Cu-roughness models available in literature compounding the issues system designers have to take into account. It is important to identify operating regime of the design parameters and adopt an approach that results in practical outcomes. This study captures state-of-the-art developments in materials, processes and modeling methodologies to meet today's design challenges. This work shares our design approach to the upcoming switch systems to meet high throughput demands.

Author(s) Biography

Riaz Naseer completed his Masters and PhD degrees in Electrical Engineering from University of Sothern California, Los Angeles. He has been working in Signal and Power Integrity domain for the past 15 years. He is currently helping with design of data center switches at Amazon Web Services in a Senior Signal Integrity Engineer role. He has previously worked at Intel Corporation, Rockwell Automation and Ciena Corporation.

Yuri Tretiakov received his PhD degree in Electrical Engineering from Arizona State University in 2001 and MSc in Electrical Engineering from Moscow Institute of Physics and Technology in 1997. Yuri has been working with Amazon since 2017 in various roles covering signal and power integrity analysis. Yuri has previously served as Principal Engineer at Rambus and as Staff Engineer both at Samsung and Altera.

Eric Tran has been with Amazon since 2012 and have served in various roles for signal and power integrity and ESD engineering. More recently, Eric has been leading the new product development for data center switches from SI/PI perspective and working across domains and organizations in developing state-of-the-art cost-effective solutions for AWS. Prior to Amazon, Eric worked as Signal Integrity engineer at Cisco Systems. Eric completed his BS and MS degrees in Electrical Engineering from University of Washington.

1.0 Introduction

Cloud infrastructure is at the forefront of technology enabling advances in artificial intelligence (AI), communications and computation. Ethernet switch devices play a critical role for the flow of information across the network. While the industry is working towards the development of 224Gbps/lane Ethernet specifications, the first-generation switch devices using 112Gbps/lane speeds (and aggregate bandwidth of 51.2Tbps) are about to hit the market. Enabling 112Gbps speeds for the VSR (very short reach) chip-to-module (C2M) channels is challenging given end-to-end channel insertion loss (IL) budget of 16dB; of which only 11.9dB is dedicated to the switch host PCB. Figure 1 shows the C2M channel loss budgeting across different components of the channel as specified in IEEE 802.ck-2022 standard – Annex 120G [1]. OIF-CEI-5.1 implementation agreement (IA) also specifies a similar total channel loss budget of 16dB for CEI-112G-VSR-PAM4 PCB channels with some variations across the constituent components [2]. According to OIF-CEI-112G-VSR-PAM4 loss budgeting, host PCB loss is capped at 12dB and connector and module losses are equally set at 2dB at the Nyquist frequency.
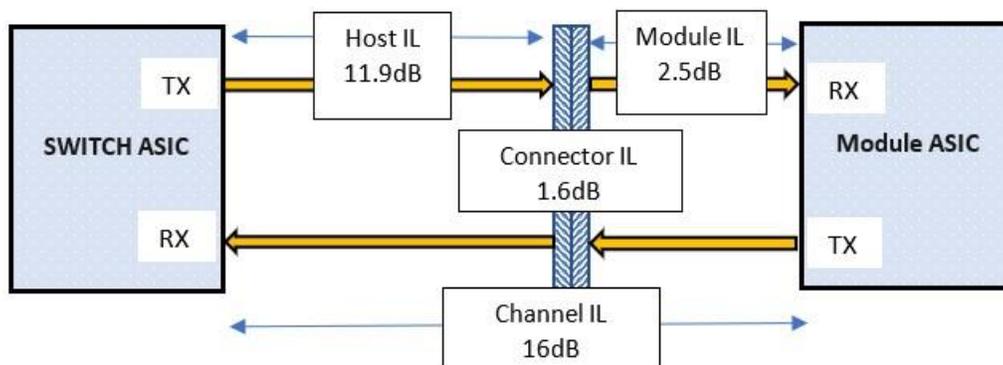

Figure 1. C2M Channel Loss Budgets – IEEE802.ck-2022 Annex 120G

Solutions have been suggested to use fly-over type cables from the switch ASIC to the front-panel ports especially for the longest channel ports on the edges of the pizza-box (generally referred as Top-Of-the-Rack TOR) switch form factors that go in standard server racks. This is due to the longest channel traces could be as large as 12-14 inches long and with the PCB copper losses it was deemed impractical to meet the 11.9dB PCB channel loss at the Nyquist frequency of 26.56GHz/28GHz. Fly-over type cables have been promoted to replace PCB copper since those offer very low loss/inch metric compared to PCB traces since the introduction of 25Gbps/lane speeds. In addition to the increased channel reach, cables offer better thermal signal integrity performance. Figure 2 below show fly-over cable performance from Samtec with their cabling system called "flyover" [3].

Although fly-over cables offer advantages from signaling performance point of view, overall system design constraints make those unattractive choice. Fly-over cables increase system cost significantly; each 800Gbps port would need 8+8 fly-over cables as
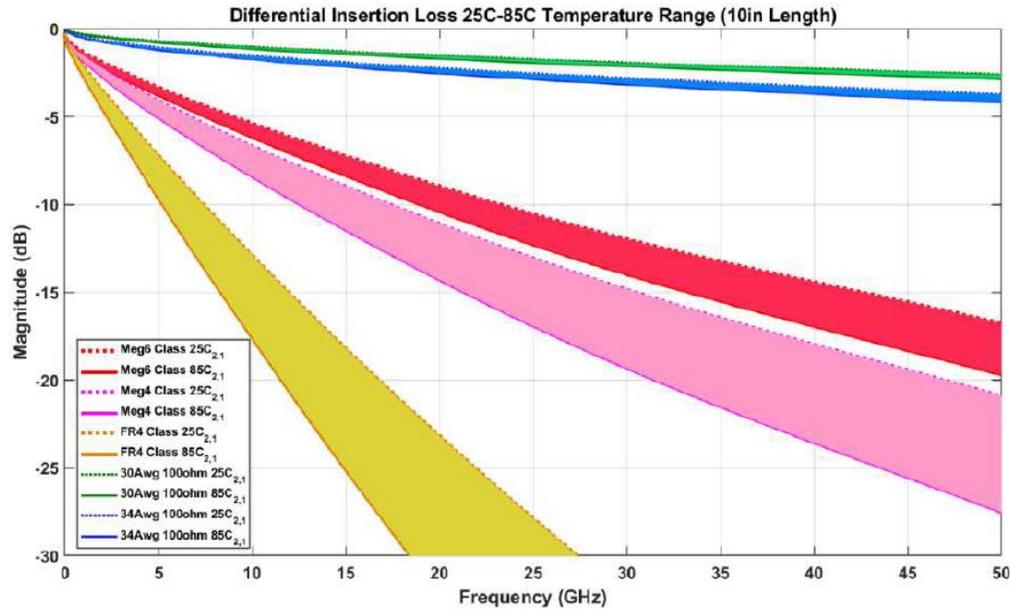
Figure 2. Differential Insertion Loss for a Samtec Flyover Cables (10-inch) versus PCB materials
Courtesy: [3] Anthony Fellbaum; Samtec

well as a connector on the PCB somewhere near the ASIC BGA area. These high-performance cables cost much more than the state-of-the-art ultra-low-loss PCB materials. The additional connectors placed on the PCB to tap the signals for cables increase PCB area and hence the cost of the system. In addition, fly-over cables pose mechanical, thermal & reliability challenges for insertion/removal & vibrations and hence result in hard-to-debug errors during the operational lifespan of the switch devices. These factors make the PCB only solutions highly desirable.

When the Ethernet speeds increased to 25Gbps/lane to support 100Gbps ports in 2012-16, the PCB laminate suppliers came up with the 6th generation PCB materials such as Megtron-6 with dissipation factor Df in the range of 0.003 to 0.005. During the 200G-400G port speeds (56Gbps/lane PAM4), the PCB laminate suppliers transitioned to 7th generation materials such as Megtron-7 with dissipation factor Df in the range of 0.002 to 0.003. For supporting today's 800G Ethernet ports, the laminate suppliers have introduced 8th generation materials with Df < 0.002. At the same time, the copper-finish on the PCB has improved significantly; going from roughness profile of very low profile (VLP) copper to HVLP (hyper very low-profile copper), HVLP2, HVLP3, HVLP4 etc. All of these advances in PCB manufacturing technologies have helped the industry in extending the cost-effective PCB solutions to meet the ever-increasing performance demands of the industry. Eric Bogatin has written a very interesting article in Signal Integrity Journal with an apt title, "The Reports of Copper's Death May be Exaggerated" [4].

The rest of the paper is organized as follows. Section 2 discusses PCB manufacturing process briefly and then discusses channel loss estimation techniques including various copper roughness modeling approaches. Section 3 presents our simulation and modeling

results for the 112Gbps/lane case and compares our simulation results with the coupon measurements. Section 3 also shares our channel simulation results. Section 4 presents the summary.

2.0 PCB Manufacturing Process and Cu-loss Estimation

This section provides a brief background on PCB manufacturing process and PCB channel losses. There are many references available in literature with more details regarding the PCB manufacturing e.g. [5][6].

2.1 PCB Manufacturing Process

The PCB fabrication process starts with copper clad laminates (CCL) which consists of sheet of copper on both sides of a dielectric core. The copper cladding is created in two different ways; rolled annealed copper and electro-deposited (ED) copper.
Rolled annealed copper is prepared by passing the copper billet/ingots under the pressure of gradually smaller rollers; as depicted in Figure 3 [5]. This process results in the smoothest copper foil but is quite expensive. Another method is electrodeposition process (shown in Figure 4) which employs large rotating drum partially submerged in a copper sulphate solution that deposit the copper on the rotating drum using electric charge through the anode and cathode nodes [6]. The drum-side (aka shiny side) of this process is smoother than the solution side (aka matte side). The core material is bonded to this matte side to have good adhesion properties. For making high layer count PCBs, the PCB fab-house combines etched/processed core layers with uncured pre-preg dielectric material between multiple core layers. As a result, the shiny side of the copper bonds with the pre-preg material. In order to have good adhesion for this side, the PCB fab-house roughens the shiny side with oxide (or oxide alternative) treatment before pressing it with multiple cores. Another category laminates are also available which are called reverse-treated foils (RTF). For the RTF cores, the CCL supplier roughens the shiny side and attach that side to core material. For the RTF foils, the core side may be rougher than the pre-preg side. The main point to remember is that the two sides of the copper may have different roughness profiles depending upon the processes it goes through and the signal integrity (SI) engineers needs to take this information into account while estimating channel losses.
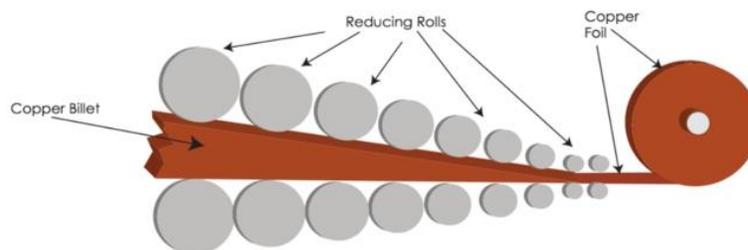

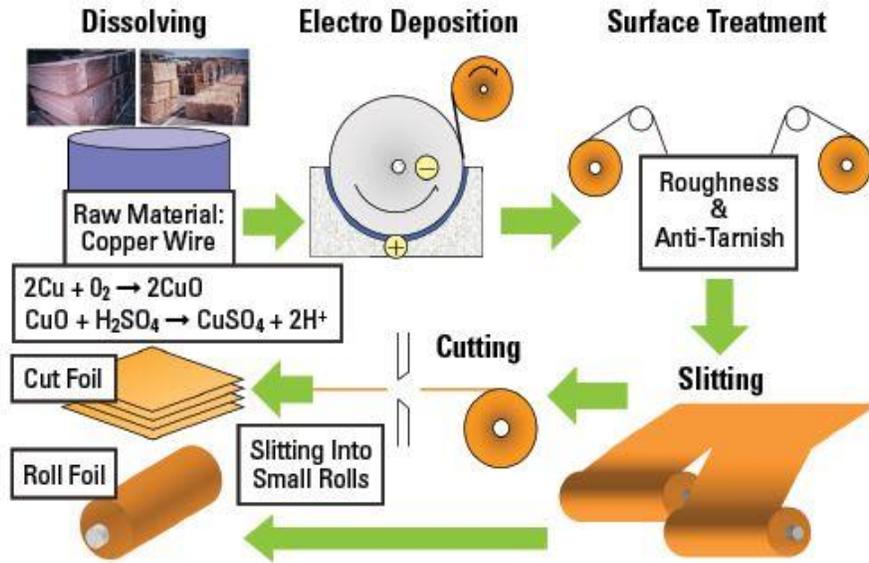Figure 3. Rolled Annealed copper process; Courtesy: Rogers Corp [5]

Figure 4. Electro-Deposited Copper Foil process
Courtesy: [6] Eric Bogatin, SI Journal

Copper roughness profile is generally specified using three parameters; Rz, Rq and Ra. Rz is a 10-point mean roughness measurement using five highest peaks and five lowest valleys usually measured with a profilometer. Rq is the root mean square (RMS) value of the roughness profile while Ra represent the absolute average value of the roughness profile and is usually more suited for defining roughness profiles of the much smoother copper laminates. Since latest copper laminates are much smoother hence instead of mechanical contact-based profilometer, imaging or light profilometer are being used more commonly. With older technologies, the Rz values used to be in the 5-10um ranges. For meeting signal integrity performance needs for increased data rates, the industry has continuously improved the roughness profile of the copper laminates such that today's copper foils have Rz range between 1-2um or even less. There are many good references in the literature which define and present measurements for the copper roughness [7][8][9].

2.2 PCB Channel Loss

The PCB channel loss is dominated by copper and dielectric materials used. Table 1 shows Dk/Df values for various generations of PCB materials along with a total PCB channel loss for a 10-inch line at 10GHz. With the technology advances in the PCB materials, the roles of loss contribution between the dielectric and copper have changed. The Figure 5 shows percentage contribution of the dielectric and ideal/smooth copper losses towards the total PCB loss for 10-inch line segments at 10GHz for various generation materials as characterized with the dielectric parameters given in the Table 1. In order to keep comparison simple, the trace width, separation and target impedance are maintained while characterizing below numbers for copper with no roughness.

Table 1. PCB Material Dk/Df over various generations

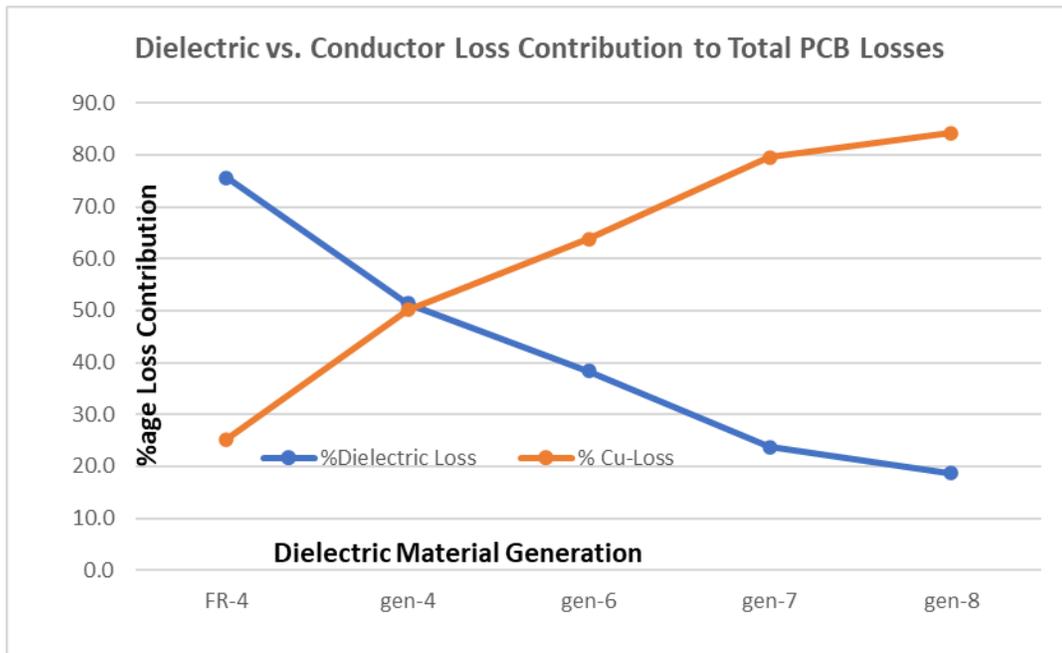| PCB Material	Dk @ 1GHz	Df @ 1GHz	Total PCB Loss/10 inch (dB) @10GHz
Standard FR-4	4.2	0.02	12.93
Gen-4	3.7	0.007	6.3
Gen-6	3.7	0.004	4.95
Gen-7	3.3	0.002	4
Gen-8	3	0.0015	3.85

Figure 5. Dielectric vs. Conductor Loss Contribution to total PCB Loss for a 10-inch line

As Figure 5 shows, dielectric losses were the dominant factor for overall PCB loss in older technologies. Somewhere along the technology development; e.g. in the Megtron-6 technology generation (gen-6), the copper losses dominated the overall PCB losses. With further improvement in dielectric losses in latest technologies, copper losses account for more than 80 percent of PCB losses. It also points that going forward the improvement in losses will be small since dielectric loss is a smaller portion of overall PCB losses. So even though the dielectric material suppliers work towards development of highly low-loss materials with dissipation factors <0.002, the copper losses severely limit the PCB loss budgets.

In the above paragraphs we talked about the copper losses without any regards to copper roughness profile. When we account for copper roughness losses, the overall contribution of the copper losses increases. In the early days, the reverse treated copper foils (RTF) used to be very rough. To meet SI performance, industry developed increasingly smoother copper laminates. Although there is no agreed upon definition, but designer

hear the words such as VLP (very low profile), HVLP (hyper very low profile), HVLP2, HVLP3, HVLP4 etc. For SI engineers, the key parameter to consider is the roughness measure such as Rz, Rq or Ra. Recently laminate suppliers are able to produce copper-clad laminates with such smooth profiles where Rz<1um but the etching/oxidation processes to enhance bonding of layers by the PCB fab-houses result in increasing the copper roughness. To mitigate this, chemical bonding solutions are being proposed such as Glicap [10].

At the PCB design stage, modeling the channel losses is a critical task for the SI design engineers. There are a lot of models available in the literature to account for copper surface roughness such as: Hammerstad-Jensen Model, Hall-Huray Model, and Simonovich Cannonball-Huray Model [11][12][13][14][15]. In the following paragraphs, we capture the essence of the 3 most commonly used copper roughness models.

The oldest used copper roughness model is Hammerstad-Jensen model (generally called Hammerstad model) [11]. It assumes the roughness profile to be a 2D triangular corrugated surface and is based on a curve fitting to an old work about copper power losses by S.P. Morgan published in 1949. The beauty of this model is in its simplicity i.e., it requires a single parameter i.e. RMS tooth height (or Rq) which is readily available from the datasheets or can be easily calculated from Rz (Rq = Rz/2√3). On the down side, it has been reported in literature that it saturates the frequency dependent losses at a factor of 2.

Hall-Huray "snowball" model [12] was developed to overcome the frequency dependent saturation limit of Hammerstad model. Hall-Hurray model tries to capture increased copper losses due to roughness using a physics-based 3D model approach; but it requires nodule radius, number of nodules to represent roughness, and the rough to smooth copper surface ratio. These parameters are not easy to estimate at the design stage unless one builds a test structure and measures it.

The Simonovich Cannonball-Huray model (commonly known as Cannonball-Huray model and labelled as such henceforth) [15] is the culmination of successive attempts by Bert Simonovich ([13][14][15]) to get the Hall-Huray model accuracy using only the parameters which are available to SI engineers from datasheets. This model simplifies the estimation of Hall-Huray model parameters by geometrically stacking of 14 equal sized spheres on to a square tile base resembling a stack of cannonballs. Another model from Bert Simonovich is called HCPES (hexagonal close packing of equal squares) roughness model [14] is similar to the Cannonball model, except it stacks 11 equal sized spheres on a hexagonal tile base. Both of these models provide the same results but the equations are different.

The industry preferred the use of Cannonball-Huray model due to its simplicity of being able to get parameters from the datasheets and due to the fact that it allows the scaling of losses with frequency. This model is available in industry standard tools such as HyperLynx, Polar, Z-Zero etc.

3.0 Simulation and Modeling Results for 112Gbps

For industry's upcoming 51.2T Ethernet switches, each SerDes lane runs at 106.25Gbps using PAM4 signaling. All the leading design houses are recommending the use of fly-over type cables to meet the IEEE-802.3ck/OIF-CEI very short reach (VSR) channel loss budgets. But cables have their own challenges of increased system cost and operational reliability concerns in an increased airflow environment. To scope out if we can implement the 106Gbps/lane solution using only the PCB channels, we analyzed latest materials available to us with reasonable assumptions on copper roughness profiles with the two leading models; Hammerstad & Cannonball-Huray models, for which we can use the parameters from the datasheets. During this modeling, all other material and trace geometry parameters are kept the same except the two target Rz values;
Rz = 1um and Rz = 1.5um. Figure 6 shows our simulation results. The shades of green are Cannonball-Huray model results and the shades of red/brown are Hammerstad model results for the two Rz values.
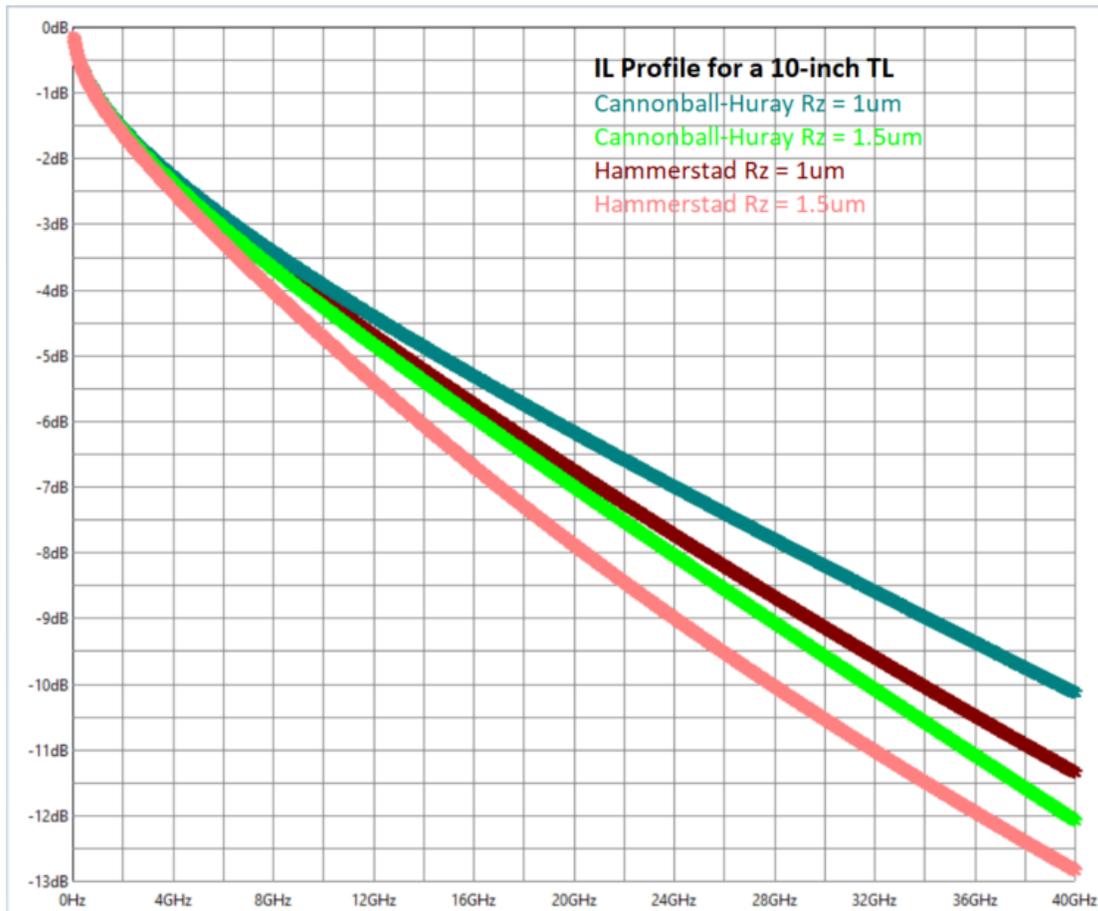

Figure 6. 10-inch Line Insertion Loss; Hammerstad vs. Cannonball-Huray

As seen in the figure, the Cannonball-Huray model results are on the lower side compared to Hammerstad model results for both cases. Although the difference is not

large; nonetheless there is a difference of about 1dB at 27GHz for the 10-inch line. As a design engineer, if one does not know which model is providing the correct answer, one would pick the pessimistic number at design time and be pleasantly surprised when the physical boards come up in the laboratory. Based on this modeling data, we knew that we would have to find the copper laminates with the best roughness profile if we were to meet the target trace lengths of 12-14 inches for ToR switch box.

With the aggressive technology scaling by the copper clad laminate providers, there is chatter in the literature about profile-free copper and chemical bonding solutions providing all the required mechanical adhesions [10]. Just to observe how the insertion loss trends will look like for the future generation of 224Gbps, we simulated profile-free copper losses and compared it to the Rz = 1um roughness profile copper using Cannonball-Huray model and Hammerstad model. Figure 7 shows the insertion loss profile for these cases for a 10-inch-long line up to 60GHz since the 224Gbps data rates will double the fundamental frequency if encoding schemes stays the same i.e. PAM4.
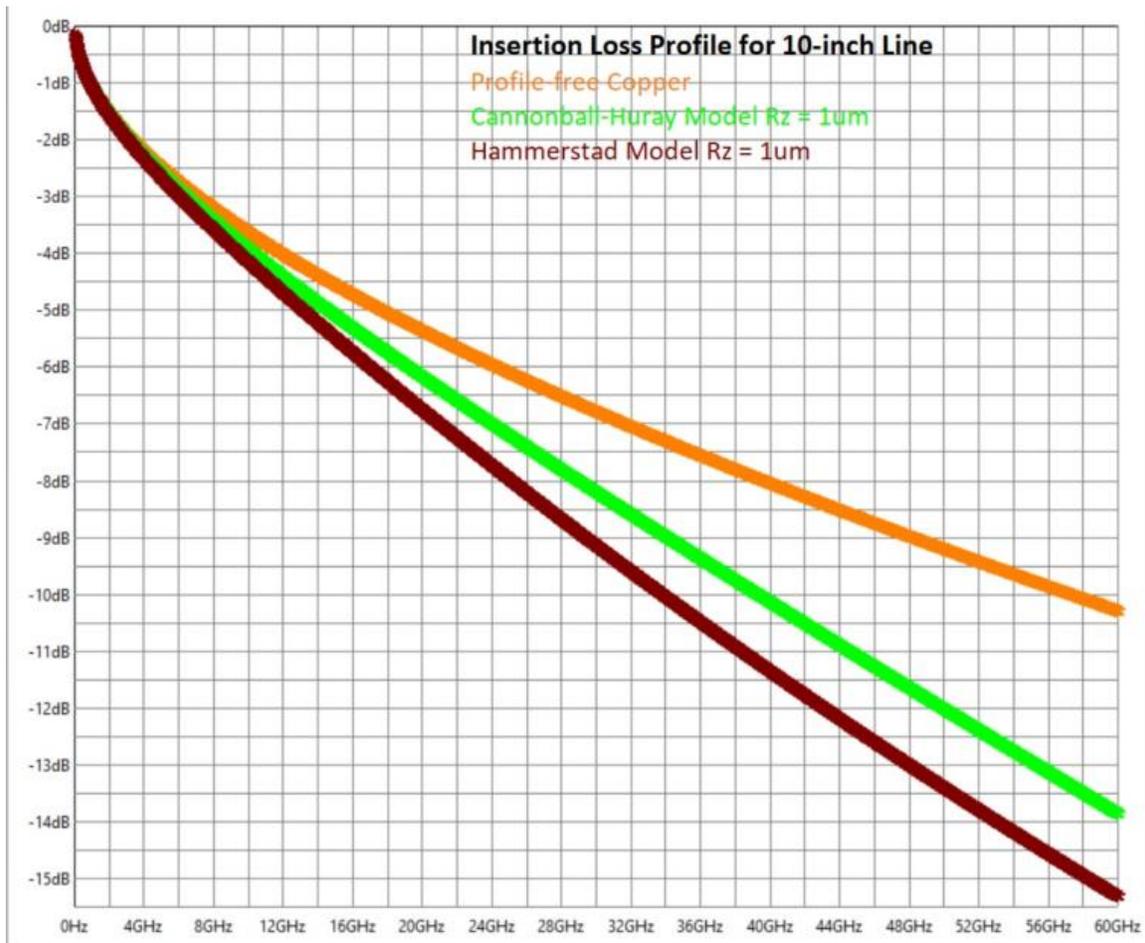

Figure 7. Profile-free Copper versus Roughness Loss Scaling to Higher Frequencies

We can see from Figure 7 that copper roughness can still increase the channel losses significantly compared to profile-free losses. But interesting point to note is that even for

the highest frequency of 60GHz, the roughness of the copper hasn't resulted in doubling the channel losses. It implies that Hammerstad model's saturation factor of 2 is not coming into picture up to the simulated frequency range with the assumed Rz=1um roughness with latest generation PCB materials where dissipation factors range below 0.002. This observation is somewhat in contrast to prevailing views in industry where Hammerstad model is presumed to under-predict roughness losses. Therefore, a modified version of Hammerstad model was developed to compensate for this mismatch in loss prediction [16][17]. But if we read the section 5.3.1 in reference [18] carefully, it states and even shows a figure (Figure 5-14(a)) that Hammerstad model predicted losses will match the measured losses up to very high frequencies if the RMS roughness profile is 1.2um which would approximate to Rz = 4.15um. Since the roughness profiles of the laminates under our deliberation are much lower than that; it would be expected that Hammerstad model results can be closer to measurements. So, this is practically the case for today's PCBs targeted for high data rates where copper roughness has typical values of Rz closer to 1um. Therefore, it is not surprising that Hammerstad model can provide reasonable estimates for the copper losses due to roughness.

As discussed previously, the copper clad laminate providers may provide much smoother copper but the oxidation/etching process/treatment at the PCB fab-house to bond cores with the pre-preg will introduce more roughness to increase mechanical bonding of the layers. This is where the new chemical bonding of Glicap [10] helps reduce the overall roughness of the fab-house and still provide the required mechanical strengths of withstanding multiple re-flow processes of PCB manufacturing, assembly and test.

We carefully screened our laminate providers and PCB fab-houses. We simulated our target trace geometries and built test coupons to perform loss measurements. The Figure 8 shows simulation and measurement results for a test transmission line with target trace geometries.

The measured results show that the insertion loss lies somewhere between the estimated loss by the two most commonly used copper roughness models. The measurements are reasonably matched to simulation models. The match can be further improved once we obtain our own optical profilometer measurements of the copper roughness instead of relying on the datasheet values provided by the CCL providers and PCB fab-houses. That is still a work in progress. Notice that Hammerstad model is not showing any loss saturation which is reported in prior literature and is the main reason for so many papers on roughness modeling. In contrast, the loss predicted by Hammerstad model is slightly more than the Cannonball-Huray model. Coupon measurement is also closer to the simulated losses predicted by Hammerstad model.
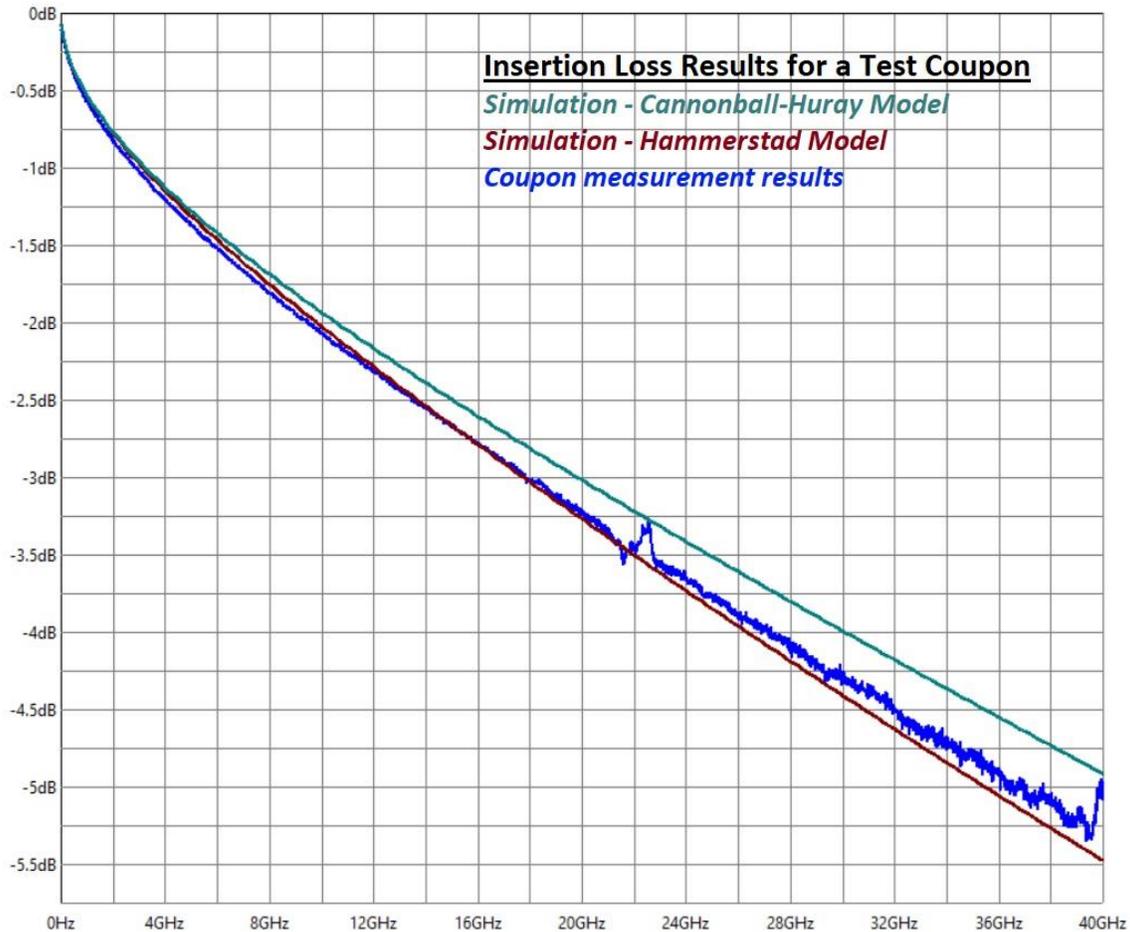
Figure 8. Simulated and Measured results of a test coupon

3.1 Channel Simulation Results

Based on previously mentioned simulations and test coupon result analysis, we determined that in order to meet VSR channel loss budget, we would have to restrict our maximum channel lengths to a particular target on our host PCB. We carefully planned the layout of our PCB to meet these constraints and built 3D models of the BGA breakout region and transition vias both in the BGA region and in the front panel port connector region.

Figure 9 shows the insertion loss and return loss profile of our longest channel including the connector and spec-compliant HCB (Host Compliance Board). As depicted in figure, our longest channel easily passes the insertion and return loss specifications. Although these results are for the nominal case and there are going to be process variations, but we feel confident that we have enough margins to mitigate process variations. We have performed channel loss analysis including the thermal impact on insertion loss and that also meets specifications.
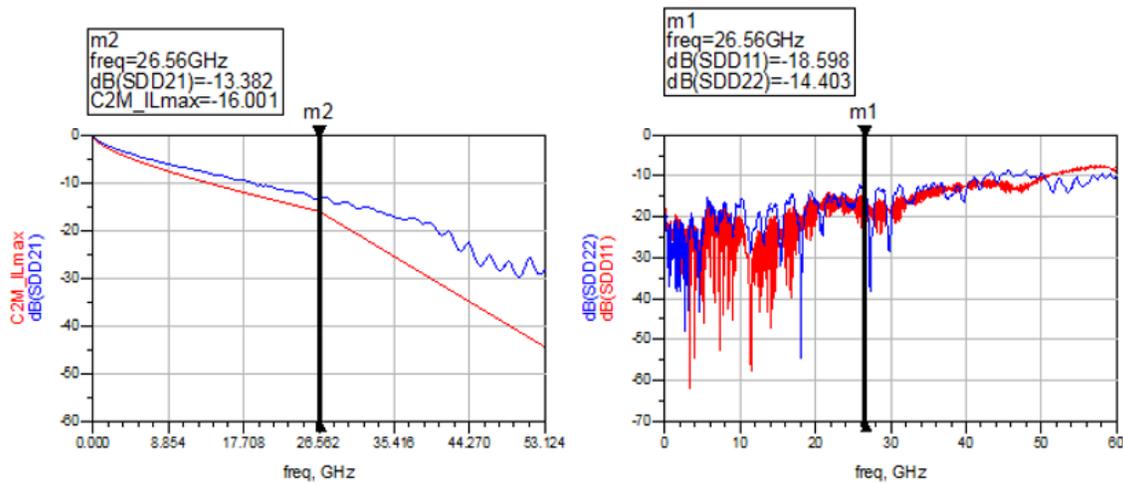
Figure 9. Insertion and Return Loss Profile for the Longest Channel

We also analyzed the channel using the IBIS-AMI simulations with ADS. The IBIS-AMI simulations passed for all the 3 corners namely; slow, typical and fast corners. Figure 10 shows an eye diagram for the typical corner case.

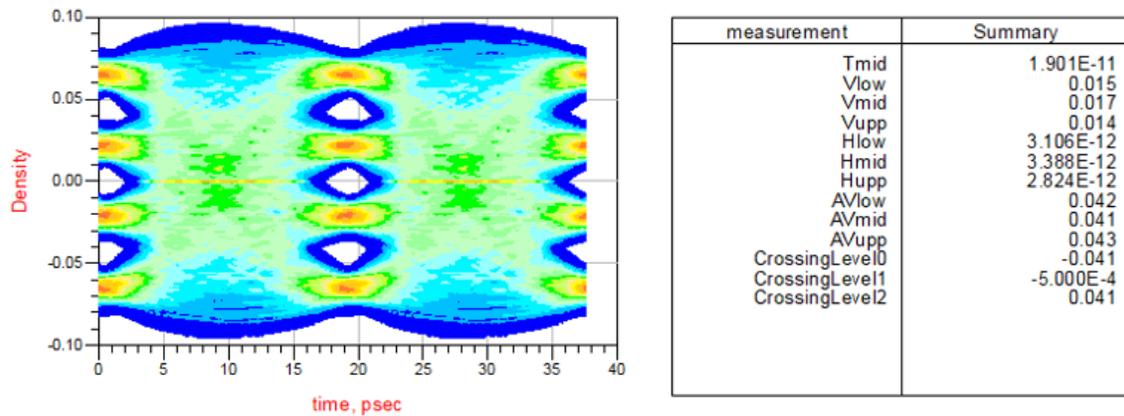| measurement	Summary
Tmid	1.901E-11
Vlow	0.015
Vmid	0.017
Vupp	0.014
Hlow	3.106E-12
Hmid	3.388E-12
Hupp	2.824E-12
AVlow	0.042
AVmid	0.041
AVupp	0.043
CrossingLevel0	-0.041
CrossingLevel1	-5.000E-4
CrossingLevel2	0.041

Figure 10. Eye Diagram for the Longest Channel, Typical Corner

Eye simulations are a one way of predicting the performance of hardware channel and build confidence that the real hardware will work fine in lab but these simulations may not capture various other factors. Channel Operating Margin (COM) is another method that confirms the passing or failing of the specification of a target channel taking into account a lot of other factors such as fitted insertion loss, crosstalk, insertion loss deviation etc. Therefore, we also analyzed our channels using the COM metric and these results also demonstrated that our worst-case channel passes this metric. Figure 11 shows the screenshot of the COM results.
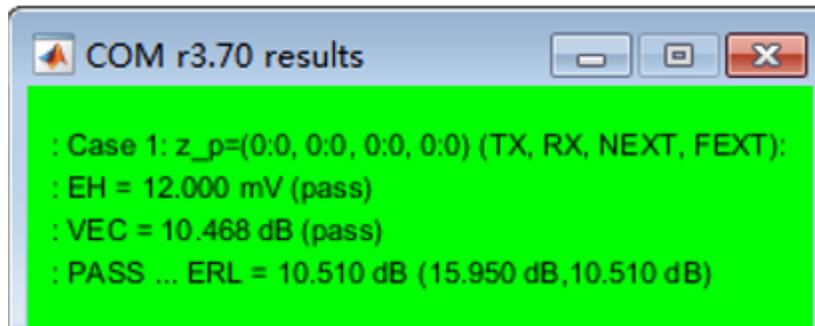
Figure 11. Channel Operating Margin Results for the Longest Channel

We have analyzed our host PCB carefully using different simulation and modeling approaches available to us. We have gerbered-out our board to the PCB fab-house for fabrication. We are looking forward to validating the board when it comes back in the lab. Although there is some uncertainty involved with any technology leading vehicle, we are confident that our board will meet signal integrity performance requirements.

4.0 Summary

This work shared our attempts to achieve PCB only solution for the 112Gbps VSR channels. The advances in PCB materials and fabrication have enabled the scaling of the copper solution to yet another generation of Ethernet data rates. The copper losses are dominating the dielectric losses in latest generation PCBs. It means that more benefits for future generation will come from smooth copper as opposed to low-loss dielectric material advances. Signal Integrity engineers should use both Cannonball-Huray model and Hammerstad model in the early stages of design to get an estimate of the expected channel losses based solely on the datasheet parameters. In order to design a successful product, SI engineers should carefully explore the design space with all the tools and models available to them.

Acknowledgements

The authors are greatly indebted to Bert Simonovich for his continuous coaching and mentoring.

References

[1] IEEE Standard 802.3ck-2022, Annex 120G
[2] OIF-CIF-5.1 Implementation Agreement, Section 25, Appendix 25.A, Dec 2022
[3] Anthony Fellbaum, "White Paper: Why use Flyover cable systems?"; Samtec; June 2023; (https://www.samtec.com/solutions/flyover/)

[4]   Eric Bogatin, "The Reports of Copper's Death May Be Exaggerated"; Signal Integrity Journal, June 2022

[5]   Rogers Corporation, "Copper Foils for High Frequency Materials"; Datasheet

[6]   Eric Bogatin, "The Quest for Copper May Have Reached Its Limit"; Signal Integrity Journal, February 25, 2020

[7]   John A. Marshall, "Measuring Copper Surface Roughness for High Speed Applications"; IPC Expo 2015

[8]   Bert Simonovich, "Practical Channel Modeling for High-Speed Design", EDICON 2018

[9]   Bert Simonovich, "PCB Interconnect Modeling Demystified"; DesignCon 2019

[10] Shikoku Glicap; https://insulectro.com/shikoku-glicap-gc-750/

[11] Hammerstad, E.; Jensen, O., "Accurate Models for Microstrip Computer-Aided Design"; Microwave symposium Digest, 1980 IEEE MTT-S International, p.407-409, 28-30 May 1980 doi: 10.1109/MWSYM.1980.1124303

[12] Huray, P.G.; Hall, S.; Pytel, S.; Oluwafemi, F.; Mellitz, R.; Hua, D.; Peng Ye, "Fundamentals of a 3-D "snowball" model for surface roughness power losses"; Signal Propagation on Interconnects, 2007. SPI 2007. IEEE Workshop on, p.121-124, 13-16 May 2007 doi: 10.1109/SPI.2007.4512227

[13] B. Simonovich, "Practical Model of Conductor Surface Roughness Using Cubic Close-packing of Equal Spheres"; EDICon 2016, Boston

[14] B. Simonovich, "Practical Method for Modeling Conductor Surface Roughness Using Close Packing of Equal Spheres"; DesignCon 2015 Proceedings, Santa Clara, CA, 2015

[15] Simonovich, B., "Practical Method for Modeling Conductor Surface Roughness Using the Cannonball Stack Principle"; White Paper, Issue 2.0, April 25, 2015

[16] Y. Shlepnev, C. Nwachukwu, "Roughness characterization for interconnect analysis"; - Proc. of the 2011 IEEE Int. Symp. On Electromagn. Compat., Long Beach, CA, USA, August, 2011, p. 518-523.

[17] Y. Shlepnev, "Unified approach to interconnect conductor surface roughness Modelling"; Proc. of 2017 IEEE 26st Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS'2017), October 15-18, 2017, San Jose, CA

[18] S. Hall, H. Heck, "Advanced Signal Integrity for High-Speed Digital Design", John Wiley & Sons, Inc., Hoboken, NJ, USA., 2009