Cliff Roth
212-777-1180
cliff@techeditorial.net
3000 words
10/25/22

Designing With High Speed SerDes Interconnects

Summary/Abstract:
Very high data transmission bit rates of 25-100 billion bits per second are now possible with the use of serialized data and differential pair connections. SerDes interconnects enable efficient transmission of data within devices, across printed circuit boards, and across systems. When properly designed, these interconnects can, with a very small number of point-to-point channels, transmit and receive billion of bits per second per channel. While zero-error-rates are impossible to achieve in practical systems, proper design minimizes transmission error rates due to physical PCB artifacts, signal path geometries and discontinuities. Failure to achieve low error rates can decrease overall system performance in the best case, and can create intermittent, hard to isolate and hard to locate failures in the worse case.

Data commonly moves at billions of bits per second, and rates keep increasing. Broadly speaking, this data, which may be carrying text, images, video and audio—can be manipulated in two forms: serial and parallel. A computer with a 64-bit CPU processor receives and sends and process data to and from its DRAM memory in parallel. Whereas a computer's connection to a camera is typically accomplished via the USB port, and data from the camera is sent in serial form. Only one bit at a time passes through the universal serial bus (USB). SerDes (Serializer / Deserializer) operations convert between these two forms of data, and have broad application across the full range of digital devices. Data may be serialized for transmission, deserialized for processing, and then serialized again for transmission into storage devices. Consider a USB camera's data after it goes through the CPU, for example, and then gets stored on a thumb drive.

But PCs are just the tip of the iceberg. Some of the many places where high speed data and SerDes operations are common include:

Communication Systems
New 5G cellular radio systems operate in the tens of gigahertz. Cable-TV can deliver hundreds of simultaneous channels of 5-40 megabits/second per channel, amounting to many gigabits per second (Gbps). Fiber optics recently demonstrated the ability to transmit at 178 terabits per second (that's 178,000 Gbps.)

Data Centers

While smartphones, gaming PCs and social media represent the exciting side of tech to consumers, it is the underlying data centers that make it all possible. Many terabytes of data per second may be served from a single data center, with individual server communications running at many gigabits per second.

Vehicle Systems
Cars with automatic braking and other advanced self-driving technologies are making crucial high-speed decisions. There's also entertaining back-seat passengers with video and gaming. It's an electromagnetically noisy environment where long cable lengths necessitating well-isolated serial communications are common.

Devices
CPUs still use parallel data for processing and for interfacing with system memory. Computer systems uses serial interfaces for most other interfaces, such as USB, SATA, PCIe, Ethernet and HDMI connections. FPGAs (field programmable gate arrays) and ASICs (Application Specific Integrated Circuits) and DSPs (Digital Signal Processors) commonly have serialized inputs and outputs, primarily to save device pin count, and to enable easier and simpler interconnect topologies.

NVMe (Non-Volatile Memory Express) is a high-speed serialized interface for modern SSDs (solid state drives) that takes advantage of the increased bandwidth and widespread use of PCIe, an interface that leverages 4 SerDes channels (see below). HBM (High Bandwidth Memory) stacks are commonly used for graphics processor memory. The HBM3 specs operate at 5.2Gbps. NVMe runs at 16 Gbps per differential pair.

Standard Interfaces
The standard PCIe (Peripheral Component Interconnect Express) interface for computer graphics cards (GPUs) and other devices such as solid-state memory is based on 1 to 16 channels (lanes) of serial data. Each lane consists of two differential pairs, one pair to send data and the other pair to receive. A Wi-Fi card might use just one lane (PCIe x1), while a graphics card uses 16 (PCIe x16).

PCIe has evolved over a number of generations, starting from its early days purely as a parallel interface (PCI), to today's implementation over serialized lanes (PCIe). The third generation (Gen 3), for example, runs up to 1 GB/s (that's 8 Gbps) per lane, while the more modern Gen 4 can achieve up to 2 GB/s (16 Gbps) per lane. Other commonly used serial interfaces include Serial ATA (SATA), USB, Serial Attached SCSI (SAS), FireWire (IEEE 1394), and RapidIO. In digital video, HDMI, DVI and DisplayPort all use serialized data.

Why Use SerDes?
All these applications and interfaces are driving the use of SerDes technology. If one device has a parallel interface, another has a serial interface, a discrete SerDes device is needed to interconnect them. But more commonly, SerDes processing takes place internally, such as when an SSD receives data and converts it to parallel form to prepare it for storage.

Serializing offers advantages in the real world of designing products to be both durable and cost effective. A laptop's screen hinge is a case in point: Moving data across the hinge requires use of a flexible cable, and the more data lines in this cable there are, the more potential for failure. Serializing the data reduces the number of copper connections that are required.

Serial data usually costs less, especially where cables and or circuit board connections are concerned. Besides BOM savings, because the majority of high-speed serial interfaces are typically implemented over differential pairs or optical fiber, electromagnetic emissions and susceptibility is greatly reduced.

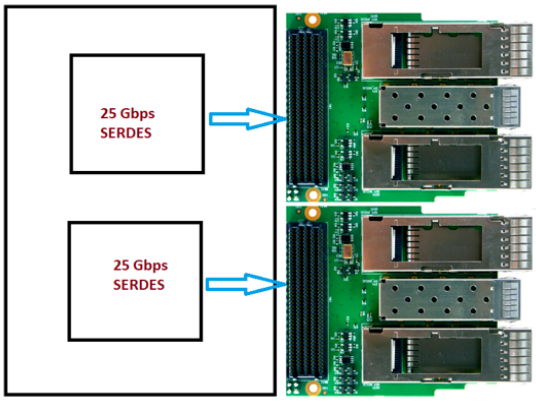
All these reasons speak to serialized communications' advantages. But the main disadvantage of serialized data is that, by definition, it must move at much faster data rates than the equivalent amount of parallel data.

Data Rates And Performance

A digital device's maximum data rate is rarely a brick wall, but more likely as you get closer to the rated maximum the signal becomes noisier, less reliable and more error prone. So one would expect a gradual roll off near the max.

But data running at 1 Gbps and more has special characteristics resulting from its millimeter-wavelength electromagnetic effects, with circuit paths and connectors acting as unintentional RF transmitters, and resonances occurring from a variety of sources relating to how the circuit board or stack of circuit boards is built. Potential problems may arise with the printed circuit board material itself, the thickness of traces and grounding pads on the board, connectors between boards and between boards and devices. Since resonances occur at specific frequencies, data at those frequencies may be more prone to errors.

A typical high speed data scenario is depicted below. Two 25 Gbps Ethernet connections to QSFP (quad small form-factor pluggable) utilize the 25GAUI interface. This topology exhibited a theoretical loss of 10dB, which is the allowable maximum loss based on IEEE standards.



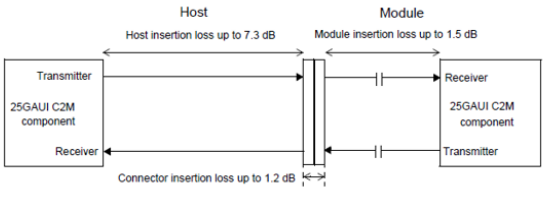


Figure 109B–2—25GAUI C2M insertion loss budget at 12.89 GHz

Caption: In this example, a signal loss of 10dB is expected across the 25Gbps signal path.

However, in actual testing at various frequencies, the boards were found to exceed the 10dB calculated loss at certain frequencies, measuring almost -14dB at 12.5 GHz. (Note that for modulated digital data, the Nyquist theorem essentially states that the maximum rate for binary data is double the frequency of the carrier signal, so a 25Gbps signal runs at 12.5 GHz.)

A more detailed research paper on this subject, "Debugging High Speed SERDES Issues in Multi-board Interconnect Systems," describes a testing setup to simulate signal transmission between circuit boards and interconnects and devices, and to investigate problems that arise with high speed data. This research revealed several sources of resonance in various aspects of circuit board design, and from cascading boards, resulting in performance drops (increased errors) at data rates well below the system's maximum data rate.

The purpose of this research was to investigate common complaints and observations with high-speed serial interconnect circuit board designs, including:

- "The system works at 10 Gbps, but fails to operate in the extended 11.25 Gbps mode"
- "The system runs error-free at the PCI Express Gen3 Data rate but fails to run at Gen4"
- "All lanes except two run error-free at 25 Gbps"
- "Eight out of ten boards appear to run error-free"
- "It appears to work fine when I press hard on the connector"
- "We added ground planes and a lot of ground vias and the problem went away"

The research focused primarily on data travelling over differential pair connections. Differential pairs are widely used in applications where rejection of noise is paramount. If noise is picked up equally by both wires or data paths in the pair, it will cancel itself out, since the signal is based on the difference between the two wires. (The noise is said to be a common mode signal, and eliminating it is called common mode rejection.) Ethernet networking and many other communications systems rely on differential pairs to reject noise over long cable runs.

Test Setup

The multi-board research test system consisted of a PCI Express circuit board and an NVM Express SSD module. The SSD had 4 Transmit and 4 Receive differential connections. A physical loopback board was used to connect the Tx and Rx lanes together. Tx amplitude was set to maximum with no pre- or post-emphasis and receiver DFE (Decision Feedback Equalizer) turned off. The interface was run at 16Gbps.

Caption: The testing setup consisted of a PCI-e board and an NVMe SSD

Caption: Close-up of the NVMe SSD interconnect

Ten boards with this setup were tested, and one lane failed repeatedly on most of these boards running at 16Gbps. It was different from the other lanes because it had asymmetric routing at the input to the connector (the two paths of the differential pair were not identical), which was its main distinguishing characteristic compared to the other, more reliable lanes.

A complex simulation model was created to compute what was going on with this failure-prone lane, based on analysis of S-Parameters.

S-Parameters and Signal Loss

S-parameters (scatter-parameters) describe the behavior of electrical networks and allow us to predict the impact that these networks will have on signals as they propagate through these networks.

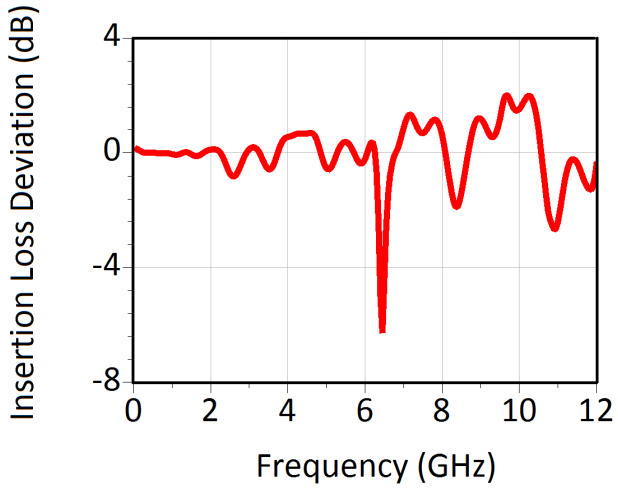
Using these parameters, designers can arrive at three fundamental mechanisms that result in signal degradation; Insertion Loss, Return Loss, and Crosstalk.

Insertion Loss is the signal energy that is lost (typically dissipated as heat) through a given link.

Return Loss is the loss in reflected signal energy at the interface boundary.

Crosstalk is the result signal energy leakage (by an aggressor signal) and the unintentional coupling of this leaked energy onto another signal (victim signal). The coupling mechanism can be inductive or capacitive, or via electromagnetic radiation, and can impact signals on printed circuit boards, cable bundles or any other multi-conductor topology.

For the multi-board research report, the Insertion Loss was computed for the problem lane, as shown here:



Caption: Insertion Loss for data going into the SSD module

Signals at 6.4 GHz, a resonant frequency of the test setup, experienced more than a 6dB drop in signal strength.

The Return Loss for the problem lane was calculated for each end of the differential pair (blue and green), as well as the combined signal (red):

Caption: Return Loss for Differential Signal

The red differential graph reflects the rejection of common mode interference, but still suffers from the resonance problem. Also, the P and N signals are slightly out of phase and not identical in amplitudes.

Since the system was designed to work at data rates up to 16 Gbps (corresponding with a frequency of 8 GHz, per Nyquist), the big resonance at about 6.4 GHz was very problematic, with over 10dB of return loss, and 6dB of insertion loss.

Circuit Boards and Signal Paths

The research report investigated how qualities of a circuit board itself affect performance with high-speed data, including both the board materials and circuit board design.

Traces on a printed circuit board have a big impact on the S-parameters at high speeds. The physical characteristics of traces include the type (microstrip, stripline, edge coupled or broadside coupled), the trace impedance, trace coupling, the trace thickness and surface roughness, trace coatings, and trace spacing.

All circuit boards suffer from signal loss along their traces. For high quality glass epoxy boards, loss of about 1db/inch is typical. Inexpensive boards typically lose about 2dB/inch, or more. While that doesn't sound like much, for a 5-inch trace it's significant.

Reference Planes and Ground Pins provide grounding for multi board systems. On each board itself the distribution of ground pads and pins at various points on the board affect S-parameters.

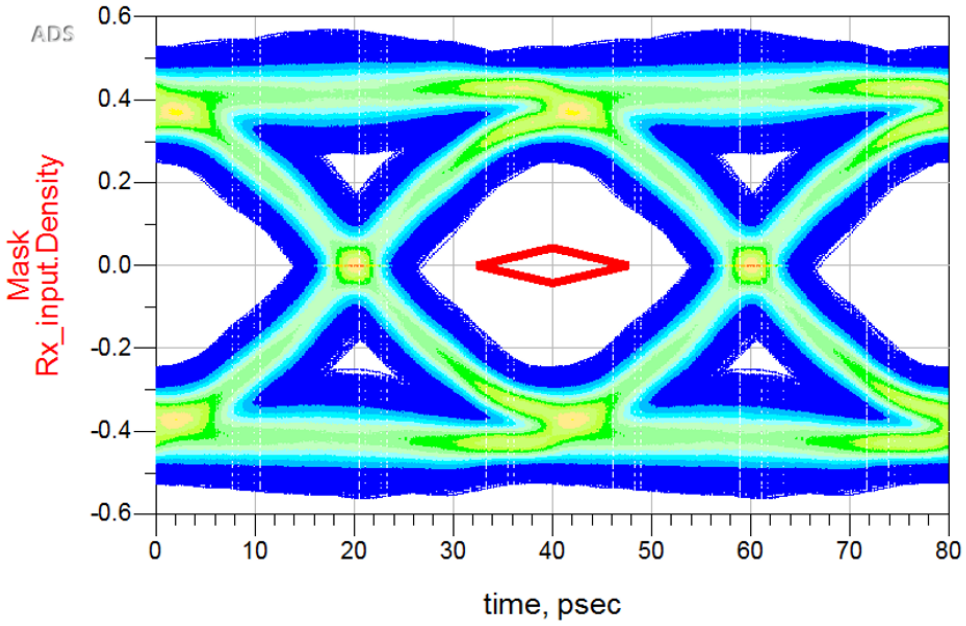
PCB Material & Fiber Weave: The PC boards manufactured for the research used standard single ply laminates with 1078 weave style. Fiber Weave Skew is the amount of time-delay that may occur randomly (only in some boards) along signal paths due to physical characteristics of the board. Values of fiber weave skew in the range of 7 pS/inch can be expected for this board type. Boards with tighter weaves (higher weave density) exhibit less skew.

Other Sources of Skew: In addition to fiber weave, many other factors contribute to skew of signals on a board. These include the PCB Traces, the PCB Vias, AC Coupling Capacitors, BGA (ball grid array surface-mount packaging) Breakout, Connectors, the PCB Stackup, and the PCB Material. *Differential Pair Skew* occurs when one part of a differential pair signal arrives at the receiving end slightly later than the other part, creating distortion of the original signal. Parallel connections may similarly suffer from skew issues, when one data path arrives slightly early or late compared with the others.

Eye Diagrams Display Signal Integrity

The eye diagram provides a convenient visual display of digital signal integrity. It is an overlay showing the transitions from 0 to 1 and 1 to 0 of millions of bits.

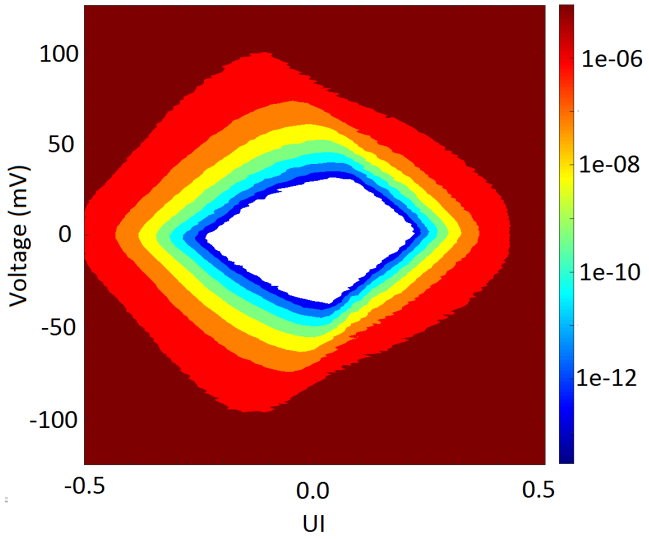
Theoretically if the transition from a lower to higher voltage, or higher to lower, took place instantaneously, in zero time, the eye diagram would look like a box with straight up and down lines representing the transitions (a square wave). But real-world high-speed transitions take a small amount of time, and so the edges are slanted in opposite angles going up and down, leading to the eye-like shape.



Caption: The eye diagram overlays millions of bits of data transitioning from 1 to 0 and 0 to 1.

When the middle area of the eye diagram has a big, wide opening, signal integrity is considered good. When the opening is small, the signal is more prone to errors.

The eye diagram can use colors to represent various aspects of the data. Here, for example, is an actual measured eye pattern from this research report showing a lane, or data path, that was working properly at eight different bit rates:

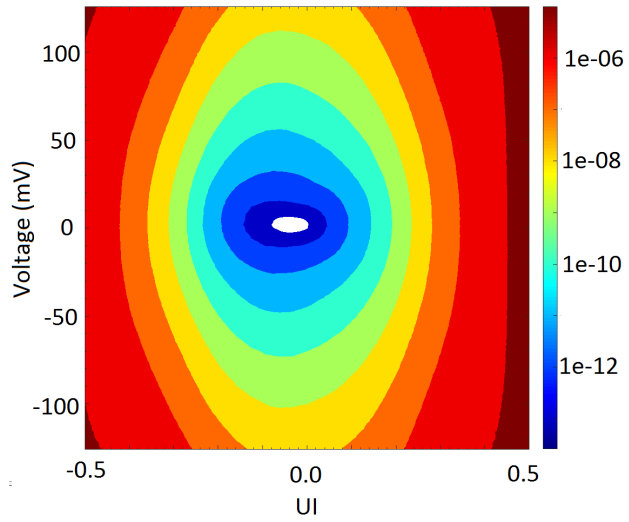


Caption: Measured Eye diagram of an error free lane

The color code goes from red to blue as the Bit Error Rate goes down, and the Unit Interval (the amount of time for a single bit of data) gets shorter. (The amount of time represented by the UI varies, depending on the data rate.)

The eye width and height at the lowest bit error rate determine the margin available. The larger the white area, larger the margin.

Compare that with a different measured eye pattern from the data lane with high bit error rate:



Caption: Measured Eye diagram of the failing lane

This eye diagram shows only a tiny area of white in the middle, reflecting a very high likelihood of errors.

Solutions and Techniques For Minimizing Errors and Maximizing Data Rates
Based on this research, there are a number of concrete steps that designers can implement to prevent these problems and exercise best practices for high-speed board and interconnect design:

Symmetric Path Solutions
For differential signals, maintaining absolute perfect symmetry in the circuit board traces of the differential pair is crucial to signal integrity. The traces should be the exact same lengths, same widths, and ideally should run right next to each other to maximize common mode rejection.

Grounding Solutions
Grounding on both the circuit boards and the interconnects between the boards greatly minimizes resonance issues. Ideally the boards should have large ground pads located in multiple places, and the interconnects should incorporate multiple ground pins. Reference ground planes should be located close to the boards, with multiple grounding connections to the boards.

Adaptive Solutions
Available corrections include use of an adaptive receiver and DFE (Decision Feedback Equalizer), both of which were shown in the research to provide good solutions to the problem data paths. However, the research showed these remedies should not be applied across the board, as using them on data paths that had no problems actually worsened performance.

Materials Solutions
Fiber weave skew from the circuit board itself can be reduced by using glass epoxy circuit boards with denser fiber weaves. The higher the weave number, the less skew. Also the quality of pin connections, quality of interconnects, and thickness of connector pins and traces may all impact high speed performance.

Conclusion:
Careful Design Enables Reliable High-Speed Operation
Designing circuit boards and interconnects for reliable high-speed data requires consideration of many factors that might not be of concern at lower data rates. It can be a long road from a lab-based prototype to a finished product that operates reliably. Among the issues designers can encounter on this journey, ground plane resonance and fiber weave skew and other issues discussed in this whitepaper can severely impact the product development roadmap. Differential signal paths must be identical, and grounding must be continuous to maintain a uniform impedance across the signal path. Failure to adhere to these basic principles may result in overall system degradation. Consulting with designers who have surmounted these hurdles and know the subtleties of high-speed circuit board design may ultimately provide significant manufacturing savings and improved time to market. Once designed, be sure to make and test many sample boards—at least ten—as some problems will appear in only a fraction of the boards.

###