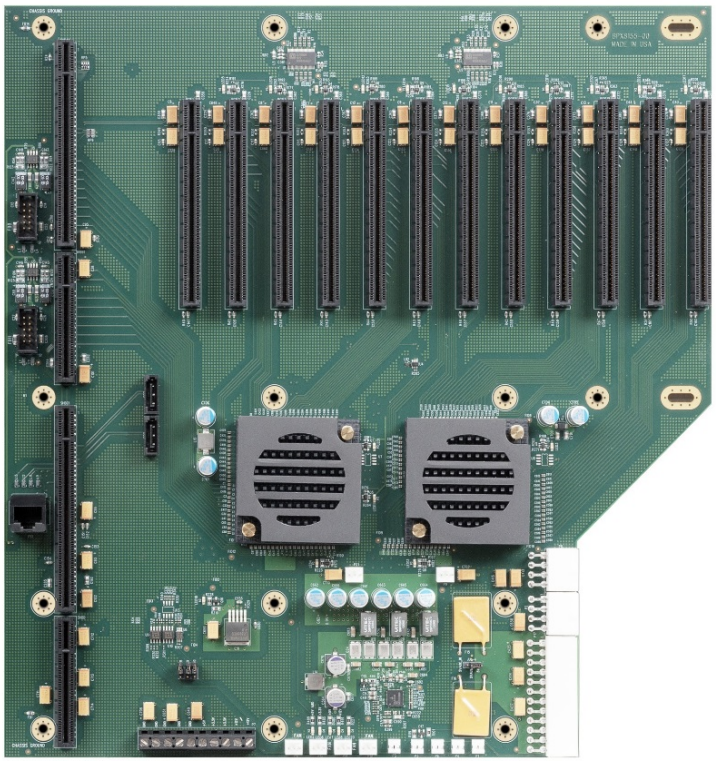
Optimizing PCIe 3.0 Backplane Designs for Performance and Reliability

By

Steve Riley, Sr. Design Engineer

Contents



Introduction

PCIe 3.0, also known as Gen3, is the newest release of the ubiquitous PCI Express high-speed peripheral interconnect standard.

The interconnect bandwidth for PCIe 3.0 is a 8GT/s bit rate, which effectively doubles the PCIe 2.0 bit rate, while still preserving full compatibility with all existing software and mechanical interfaces.

This version-to-version continuity is one of the key advantages of using the PCI Express standard, which supports performance scaling and interoperability for a wide range of board level functionality, along with low cost, low power and minimal changes at the system level.

PCIe 3.0 also provides enhanced signaling and data integrity optimization features, including transmitter and receiver equalization, clock data recovery, PLL improvements, and channel enhancements.

By using a new 128b/130b encoding scheme, PCIe 3.0 reduces the data overhead as compared with the previous 8b/10b encoding used in PCIe 2.0 and helps achieve the higher interconnect bandwidth. However, the more complex encoding scheme and higher bandwidth require excellent signal integrity.

At the backplane design level, the doubling of data transfer rates in PCIe 3.0 means much higher signal speeds and more signal integrity challenges. Some key issues include the placement of components and the length of the traces that connect them together. Other issues involve the use of active vs. passive signal management methodologies.

This paper provides an overview of the critical issues that are important for optimal PCIe 3.0 backplane design as well as the performance advantages and system configurability benefits that can be achieved through a proper PCIe 3.0 backplane implementation.

Enhanced Performance Opportunities with PCIe 3.0

The on-going stream of enhancements to PCI Express speed and performance is being fueled by increases in both system demands and board level capabilities. For example, new graphics, video, GPU, FPGA, and telemetry cards as well as other high-performance sub-systems along with advanced SBCs and mission-critical applications have been constantly raising the bar for delivering higher system performance.

As was the case with PCIe 2.0, the push for PCIe Express 3.0 usage has come from the graphics arena, where there is always an insatiable demand for bandwidth. As graphical displays continue to increase in resolution, color depth and real-time motion, the need to move more data across the backplane is a critical element for meeting application performance requirements. In systems where inter-card data movement has been the limiting factor, moving from Gen2 to Gen3 bandwidth means the time required to display images at a given resolution can be cut in half, thereby doubling graphics performance.

In addition, PCIe multi-cast capabilities can boost performance even more by enabling simultaneous channels between the CPU and multiple GPUs. PCIe 3.0 is playing a key role in enabling advanced audio, video, telemetry and signals intelligence analysis by providing high-bandwidth, multi-channel interconnect options that support better aggregation and scaling while minimizing power consumption.

Real-time applications such as airborne telemetry and other mission-critical military systems utilizing the latest high-speed data acquisition I/O cards are also leveraging the higher bandwidth and performance capabilities of PCIe 3.0 based platforms.

Looking ahead, with PCIe 4.0 (Gen4) already scheduled to provide another doubling of performance when the specification is available in 2014-2015, all of the lessons learned with Gen3 board design and system implementation will provide a key foundational baseline for the next leap forward in serial interconnect technology.

Overview of New Technical Changes in PCIe 3.0

High-efficiency Encoding Scheme

As mentioned above, the PCIe 3.0 encoding scheme enables a doubling of the effective bandwidth without having to double the encoded bit rate. This provides a significantly higher level of data transfer efficiency and reduces overhead.

PCIe 2.0 uses an 8b/10b encoding scheme that delivers an interconnect bandwidth of 4 Gbit/s per lane from encoding the raw bit rate of 5 Giga Transfers per second(GT/s). In comparison, PCIe 3.0 uses a scrambling methodology with a 128b/130b encoding scheme that delivers a useful bandwidth of nearly 8 Gbit/s from a raw bit rate of 8 GT/s. This effectively provides a 98.5% efficiency as compared to 80% efficiency with the 8b/10b encoding.

Therefore the PHY is able to operate at almost a 20% lower frequency than the 10 GT/s that would be needed to deliver the same bandwidth with the previous encoding scheme.

The 128b/130b scrambling methodology offers the advantages of higher bandwidth at lower bit rate which conserves power consumption and avoids over complicating the PHY design.

However, as will be further explored in the next section, the more complex encoding methodology creates additional pressure for board-level hardware designers to actively manage signal integrity to assure proper performance

PCIe architecture	**Raw bit rate**	**Interconnect bandwidth**	**Bandwidth per lane per direction**	**Total bandwidth for x16 link (bidirectional)**
PCIe 1.x	2.5GT/s	2Gbps	~250MB/s	~8GB/s
PCIe 2.x	5.0GT/s	4Gbps	~500MB/s	~16GB/s
PCIe 3.0	8.0GT/s	8Gbps (approx.)	~1GB/s	~32GB/s

Enhanced Transaction Layer Features

In addition to doubling the bandwidth, Gen3 includes a number of new enhancements that address:

- Host-intelligent device interactions
- Power management features

The PCIe protocol extensions are primarily intended to improve interconnect latency, power and platform efficiency. These protocol extensions pave the way for better access to platform resources by various compute- and I/O-intensive applications as they interact with and through the PCIe interconnect hierarchy.

The protocol extensions and enhancements range in scope from data reuse hints, atomic operations, ID-based ordering, dynamic power adjustment mechanisms, loose transaction ordering, I/O page faults, BAR resizing and so on. Together, these protocol extensions will increase PCIe deployment leadership in emerging and future platform I/O usage models by enabling significant platform efficiencies and performance advantages.

Active Equalization Capabilities

Equalization is a method of distorting the data signal with a transform representing an approximate inverse of the channel response. It may be applied either at the Tx, the Rx, or both. A simple form of equalization is Tx de-emphasis as specified in PCIe 1.x and PCIe 2.x, where data is sent at full swing after each polarity transition and is sent at reduced swing for all bits of the same polarity thereafter. Since channels exhibit greater loss at high frequencies, the effect of equalization is to reduce these impacts. Equalization may also be used to compensate for ripples in the channel that occur due to reflections from impedance discontinuities such as vias or connectors.

The two most common algorithms are linear (LE) and decision feedback (DFE). Linear equalization may be implemented at the Tx or the Rx, while DFE is implemented at the Rx. Trainable equalization refers to the ability to adjust the tap coefficients. Each combination of Tx, channel, and Rx will have a unique set of coefficients yielding an optimum signal-to-noise ratio. The training sequence consists of adjustments to the tap coefficients while applying a quality metric to minimize the error.

Key Issues for PCIe 3.0 Backplane Design

Design Goals

Any backplane design must take into account these key factors for success:

- Achieving target performance levels
- Maintaining signal integrity and interoperability
- Minimizing cost and optimizing manufacturability
- Assuring lifecycle reliability and TCO targets

PCIe Gen3 interface performance increases require PCIe 3.0 backplane designers to still meet the above objectives while dealing with an ever-expanding set of board design challenges.

Board Design Challenges

At PCIe 3.0 speeds, maintaining signal integrity and achieving target performance levels present a number of challenges. At 8GT/s bit rates the signal loss increases, both capacitive and signal/noise ratio loss. Clock-jitter specifications also have been tightened.

In addition, the change to 128b/130b encoding means DC wander becomes a more significant issue that must be tightly controlled in the hardware rather than compensated for by the encoding scheme (as was the case with 8b/10b encoding).

In the past, links could use “dumb” redrivers but now the Gen3 links need more end-to-end integrity and can’t simply be broken up with passive redrivers.

Other major issues include placement of components and the length of the traces that connect them together. The max overall length of a PCIe trace depends on several factors including the dielectric loss – Cu roughness – weave of board material used, trace width, rotation of the panel weave during PCB fabrication, and the successful application of the automatic 8Gb/s equalization training.

The longer the signal has to travel and the more layers it has to pass through, the more prone it is to data loss and link failures. With previous generations of PCI Express it was best practice to keep traces well below 16 inches to insure optimum operability. The PCIe Gen3 specification makes the length requirement even more restrictive.

Additionally, the Gen3 specification also requires a pre-validation of the link before data transmission. If the automatic equalization training cannot establish a reliable link, it will not allow the transmission of data at 8Gb/s speeds, which significantly compromises the target performance goals.

Active vs. Passive Approaches

There are many ways to address these restrictions, each with their own benefits and disadvantages. The use of high-speed materials and back-drilled via stubs can be used to improve the expected signal integrity, but it comes at a high manufacturing cost. However, these passive solutions do not account for PCB material variance from batch to batch, vendor to vendor, or between assembly houses.

Using a passive approach for PCIe 3.0 not only drives up the cost of PCB and the risks of performance loss; it also limits the opportunity to take advantage of key new features such as options for fine-tuning of links between devices. For Gen3, an active design using retimers, as described below, is the optimal approach.

The use of PCI Express 3.0 retimers ensures signal quality over extended distances while offering simplified design by alleviating board layout constraints at a relatively low cost. The inclusion of retimers can extend the PCI Express 8Gb/s traces beyond 30 inches without the added cost or complexity of the other options.

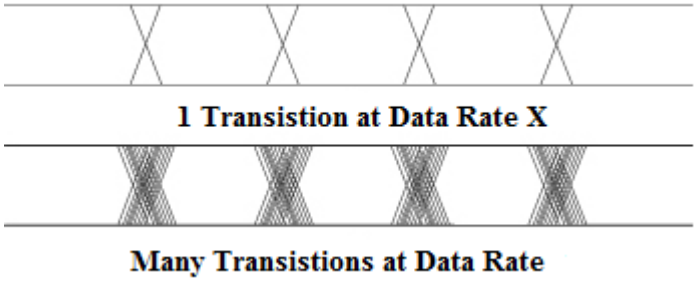
Additionally, retimers are intelligent devices that allow designers to fine tune the PCI Express signal with CTLC (Continuous Time, Linear Equalizer), DFE (Decision-Feedback Equalization), and FIR (Finite Impulse Response) capabilities.

These capabilities reduce jitter and clock fluctuations to guarantee the best operability along the entire trace. The ability to actively monitor and modify the signal channel is not possible with any of the passive methods.

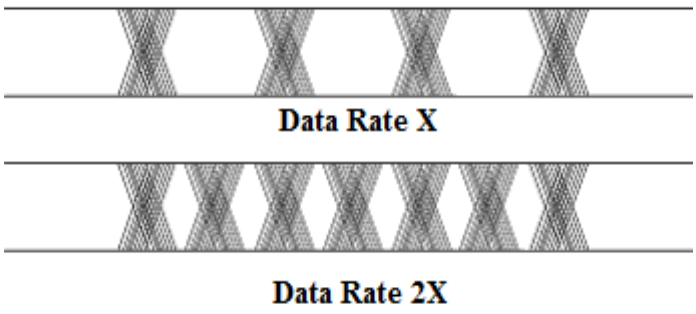
Jitter Control, Impedance and Signal Loss

The goal of optimizing loss/signal integrity is to achieve a consistently open signal transmission "eye". Key challenges are impedance discontinuity, signal loss, crosstalk and jitter, all of which work to close the eye. This is especially a problem for backplanes because trace lengths are longer, resulting in loss and crosstalk, and connectors in the signal path add impedance discontinuities and jitter.

Jitter: Jitter arises from variations in the timing of the transition edge. The total amount of jitter that can be allowed is called the 'jitter budget'. As the data rates increase, the jitter budget is reduced.



As illustrated below, doubling the data rate, while keeping the jitter constant, results in a much more closed signal eye.



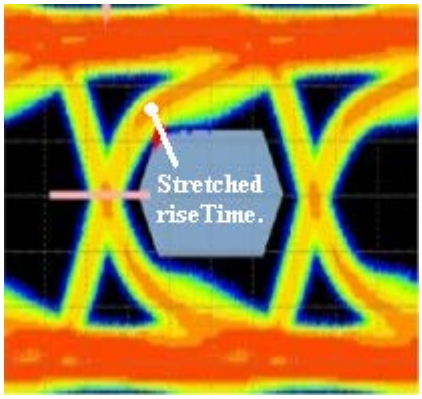
Common causes of jitter are poor pair length matching, InterSymbol Interference, and asymmetric clock cycles. A retimer can eliminate jitter from the incoming signal, resetting the jitter budget.

Impedance discontinuities: If you think of the signal as water through a pipe, impedance discontinuities are like ripples in the water that happen whenever the pipe is less than perfectly smooth. Some causes are trace or via stubs, variations in the characteristic impedance of the differential pairs, reference plane variations (gaps, etc.), close proximity of copper pours, and poor ground return paths. All of these have to be accounted for and minimized or prevented to avoid reflections, which can distort the signal and cause bit errors.

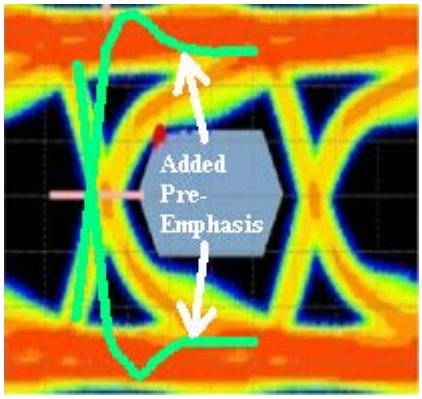
Signal loss: High frequency signals generate electromagnetic fields, which heat the dielectric material of a pcb, resulting in a loss in signal amplitude, proportional to the frequency. The longer traces means there's more signal loss.

If you think of jitter as reducing the horizontal width of the eye, then loss reduces the vertical height of the eye. An aspect of loss is the stretching of the rise or fall time of the transition, as the reactance of the PCB material, the trace length and connectors affect the signal.

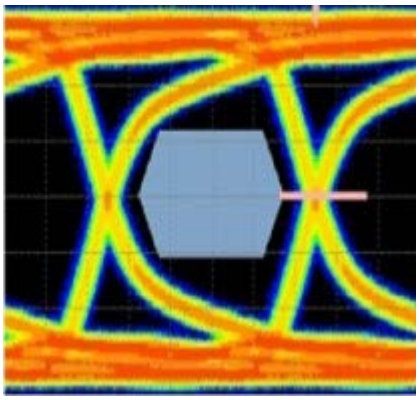
This is also considered InterSymbol Interference (ISI), as it is dependent on the data pattern. If there are a series of '1's or '0's, the voltage level of the signal has time to fully reach '1' or '0'. If this is followed by a transition to the opposite voltage level, the long transition, affected by the reactance may stretch into the next bit, like this:



To compensate for this, pre-emphasis is added to the original signal, whereby the signal is both amplified at those transitions, and de-amplified after them.



The sum of these two signals provides a more open eye, as shown below.



Unfortunately, pre-emphasis also increases cross-talk, so we can't just amplify the signal. There has to be a balance between pre-emphasis, de-emphasis and equalization.

Pre-emphasis is done at the transmitter. The other component to signal recovery is equalization, which is performed in the receiver. Essentially, equalization is similar to pre-emphasis, in that it amplifies the high frequency component of the signal. Equalization has limitations as well, since any high frequency noise in the signal will also be amplified.

Here again, retimers boost the amplitude of the signal in general, but also add equalization to their receivers and pre-emphasis to their transmitters, thereby assuring robust PCIe 3.0 links between all devices.

Real-World Deployment Issues

Previous sections have described key board design issues that must be considered when designing and optimizing a PCIe Gen3 backplane. This section offers a higher level look at real-world deployment benefits that result from adhering to stringent design disciplines.

One of the main reasons for the widespread adoption of PCI Express technology has always been assurance of robust interoperability between disparate devices across all levels of the specification. PCIe 3.0 continues that guarantee of interoperability while once again doubling performance.

In the real-world, this means that system designers can configure their optimal combination of standard and custom Gen3 cards and SBCs, with the confidence of getting the full 8Gbps performance between them.

Furthermore, in the high-performance, mix-and-match environment that is common to embedded computing and COTS products, suppliers of sub-systems often have no control over how their products are configured in the higher level system. In the case of many military applications, the suppliers may not even have access to information about how their products are being used.

In these cases, robust backplane design with active retiming and link management can make a critical difference when integrating systems with a variety of card types utilizing different PCI Express interface implementations.

For example, an airborne surveillance and telemetry application might use a combination of PCI Express Gen1, Gen2 and Gen3 data acquisition and NICs in conjunction with the latest graphics and video cards in a system with a standard COTS PCIe Gen3 backplane and single board computer.

In such a case, it is imperative that the backplane have sufficient robustness, performance, flexibility, and active link management capabilities to compensate for some level of variance in these specialized subsystems.

Summary

The introduction of PCIe 3.0 has once again opened new performance opportunities for system integrators, as well as COTS sub-systems and single board computer suppliers.

With a doubling of bandwidth, new enhanced features, robust link management capabilities, and backward compatibility, Gen3 represents a major step forward to enable next-generation system design. Gen3 is both increasing performance for a wide range of existing functions and is creating opportunities for innovative new capabilities.

Disciplined backplane design has always been the key linchpin in fulfilling the interoperability requirements for previous generations of PCI Express.

This is even more important with PCIe 3.0 where the backplane subsystem needs to perform a significantly more active role in signal propagation, equalization, link management and inter-device tuning.

System integrators and component suppliers alike are dependent on the performance, robustness and standards adherence of the backplane as a critical make-or-break factor in overall system success.

Trenton Systems – PCIe 3.0 Products

BPX8093 PCI Express Backplane

The BPX8093 PCI Express backplane when paired with a PCIe 3.0 capable single board computer like the BXT7059 delivers PCI Express Gen3 link speeds and expanded I/O card bandwidth. The backplane supports one PICMG® 1.3 single board computer and up to ten PCI Express I/O cards and nine of these option card slots are PCI Express 3.0 capable. Each option card slot uses a x16 PCIe mechanical connector to maximize system flexibility and to support a wide range of standard PCI Express I/O cards.

PCIe links on the BXT7059 SBC automatically match an option card's PCI Express link type and speed via the interface's built-in auto-training capability. This SBC and backplane design feature maximizes system flexibility by enabling seamless system support for standard PCI Express 3.0, 2.0 or 1.1 option cards having various PCIe electrical interfaces such as x16, x8, x4 and x1 PCIe links.

BXT7059 Single Board Computer

The BXT7059 SBC is a dual-processor PICMG 1.3 SBC featuring multi-core Intel® Xeon® processors and native support for PCI Express 3.0 I/O cards. This single board computer supports DDR3-1600 Mini-DIMMs with a maximum system memory capacity of up to 48GB and features SAS and SATA/600 ports, multiple I/O and communication interfaces including PCI Express 3.0.

The SBC's flexible I/O system design is driven by the PCIe 3.0 processor links and the Intel® C604 Platform Controller Hub (PCH). The external I/O card interfaces adapt to PCIe 3.0, PCIe 2.1/2.0 and PCIe 1.1 plug-in cards with either a x16, x8, x4 or x1 PCI Express edge connector. An additional PCI Express 3.0 x16 link expansion is available using the optional PEX10 module on the BXT7059 SBC.

About Trenton Systems & For Additional Information

Contact us for more information on the Trenton PCI Express Gen3 backplane offerings or any of our other embedded computing products or integrated computer systems. Trenton team members welcome the opportunity to work with you.

Trenton Systems is a designer and manufacturer of PCI Express expansion systems, rackmount computers, video capture and display wall controllers, GPU computing systems, custom systems, embedded motherboards, single board computers, system host boards, and backplanes for critical embedded computing applications.

All Trenton products are designed and built in the U.S. and are well suited to deliver long-life performance, precision and reliability in military, DoD, medical instrumentation, industrial control, and video capture and display wall systems.

Trenton industry partnerships with Intel, Matrox Graphics, Nvidia and other leading technology companies play an important role in the design, manufacture, integration and deployment of our high-performance system solutions. For example, Trenton is a member of the Intel® Intelligent Systems Alliance, a community of communications and embedded developers and solution providers. Trenton board level products and integrated computer systems are used to create a variety of mission critical solutions, from Government and Defense to Industrial Automation, Virtualization, Video Processing, Medical, Communications, Energy, GPU Computing, Test & Measurement and Video Display Walls.

Phone: 770.287.3100
Toll Free: 800.875.6031 (U.S.)
Fax: 770.287.3150
E-mail: Info@TrentonSystems.com
Web: www.trentonsystems.com

Follow Trenton on: