BOARD DESIGN GUIDELINES FOR PCI EXPRESS™ INTERCONNECT

Intel Corporation

The industry is on the threshold of a parallel-to-serial interconnect transition. This paper provides guidelines for that transition. From a board design viewpoint, the interconnect guidelines are different from prior experience and, generally, easier to implement.

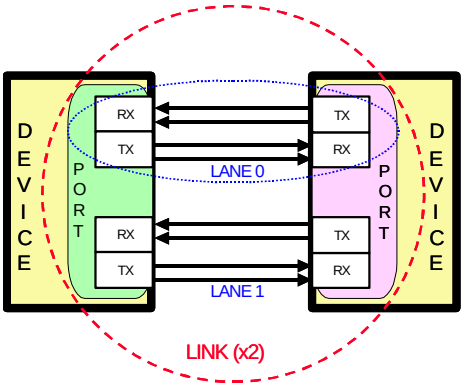
Introduction

The input/output technology for connecting digital ASICs is unable to keep pace with Moore's law, which predicts the doubling of silicon device performance every 18 months. The parallel bus I/O technology based on threshold voltage crossings, such PCI or PCI-X, popular in Intel Architecture platforms of the past decade, is now faced with several severe physical limitations. This bus technology is based on a multi drop, time-shared concept, where several devices share the bus and only two devices can communicate at a given instant. Some of the limitations with this technology are the power and ground noise, stubs due to multiple devices on the bus, trace skews, loss due to skin effect and dielectric loss, which in turn increases the settling time and large power consumption by the bus drivers. These limitations introduce various design challenges for high-speed and performance-intensive applications to use parallel, shared bus technology in a cost-effective manner. Point-to-point interconnect technology based on serial differential links is now being widely recognized as a suitable alternative for high-speed data transfer applications. This technology is based on zero-crossings of differential signals and differential signaling, which has the potential to be highly scaleable in performance with silicon devices of the next decade. Many of the limitations inherent in a parallel multidrop bus architectures are not present in point-to-point technology. This paper discusses design trade-offs required to implement the PCI Express interconnect—a point-to-point differential interconnect technology based Printed Circuit Board (PCB) trace layout.

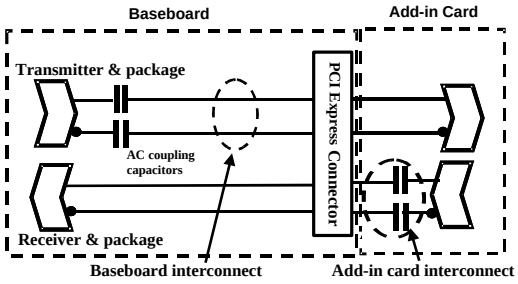
PCI Express Background

PCI Express is a dual simplex point-to-point serial differential low-voltage interconnect. The bit rate is 2.5 Gbit/sec/lane/direction at introduction. The signal is 8b/10b encoded with an embedded clock. Each lane consists of two pairs of differential signals. A link between the ports of 2 devices is a collection of lanes (x1, x2, x4, x8, x12, x16, x32 width). For example, a x16 link has 16 lanes which will provide an aggregate bandwidth of 10 Gbytes/sec. The PCI Express interconnect is cost-optimized for a 4-layer FR4 board design. The requirement is to support 20 inches between components using standard (no premium) FR4 and high volume/low cost connectors. Higher quality interconnects will support even longer distances. The Point-to-point, serial PCI Express interconnect offers layout advantages over traditional multi-drop parallel/stubbed interconnect. Trace routing occupies less space for simple point-to-point connections. The embedded clock simplifies routing rules by removing the length matching requirements between signal pairs. All sideband signals are eliminated to reduce pins and routing constraints in the base technology. The increased bit rate for PCI Express requires some design considerations for board designers that are discussed below. Interconnect losses, jitter, crosstalk, and mode conversions are the key parameters to meet instead of clock-to-data timing skews. Detailed simulation and validation are necessary to guarantee a successful design. We will now highlight some of the key points for motherboard and add-in card designers when implementing PCI Express interconnect on a PCB.

Topology and Interconnect



The PCI Express topology consists of a transmitter (Tx) located on one device connected through a differential pair (D+ and D- signals) connected to the receiver (Rx) on a second device. One of the



devices may be located on an add-in card in which one or more connectors may be present on the topology.

Each lane is AC coupled between its corresponding transmitter and receiver. The AC coupling capacitor is located on the board where the transmitter of the device resides. Each end of the link is terminated on-die into nominal 100 Ω differential DC impedance. Board termination is not required.

The D+ and D- signals from the transmitter must be attached to the D+ and D- signals at the receiver.

Polarity Inversion is required to be supported by each receiver on a link to invert the received signal if necessary. This will alleviate layout that resulted in "bowtie" scenarios (D+ and D- crisscross).

Lane Reversal (Re-ordering) support is optional between two devices. Lane Reversal allows for the re-ordering of lanes that connect transmitting and receiving devices. For example, in a x8 link, lane 0 of the transmitting device may be connected to lane 7 of the receiving device, and vice versa.

Lane Width Negotiation will allow a device with a given size port to connect to a device with a different size port. For example, a 4-port device can connect to a 2-port device with only two of the lanes active.

It is important to note that Polarity Inversion and Lane Reversal do not imply direction reversal. The Tx pair from a device must still connect to the Rx pair on the other device.

Physical Layout Design Considerations

One of the key points to note for PCI Express board design is to obey a set of rules that minimize losses, jitter, crosstalk, and mode conversion for differential traces on FR4. Another consideration is to rely on simulation analysis instead of generic layout guidelines to ensure the design will meet the interconnect budget specification.

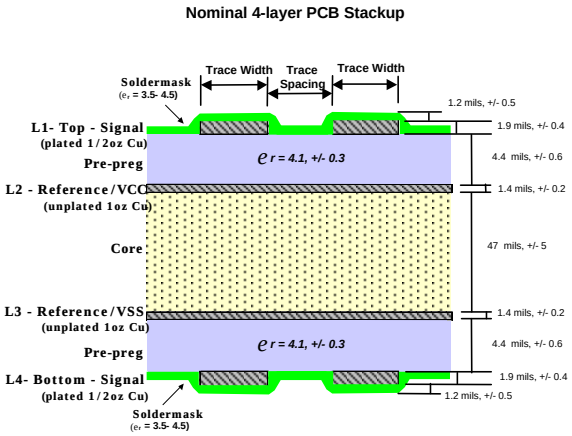
Loss for interconnect is the differential voltage signal swing attenuation from transmitter to receiver on the trace. The trace is subject to resistive, dielectric and skin effect losses. Loss increases as trace length and/or signal frequency increases. Loss also increases as trace width decreases. Vias and connectors also exhibit losses which are a part of the interconnect budget. Total loss allowed on the interconnect is 13.2 dB.

Jitter includes both data dependent and random jitter contributions on the interconnect. Total jitter allowed is 0.3 Unit Interval (1 Unit Interval = 400 ps).

Crosstalk is the coupling of energy from one signal trace to the other that results in a change in signal voltage and phase. Far-End Crosstalk (FEXT) appears at the receiver while Near-End Crosstalk (NEXT) appears at the transmitter. Crosstalk within the differential pair is not a concern. Cross talk can be minimized between differential pairs by keeping a large pair-to-pair spacing compared to spacing within a pair. Stripline traces show far less FEXT than microstrip traces.

Mode Conversions are due to imperfections on the interconnect which transform differential mode voltage to common mode voltage and vice versa.

Differential trace impedance target of 100Ω with a tolerance of 15% or better is desired. Tight coupling within the differential pair and increased spacing to other differential pairs helps to minimize crosstalk and EMI. If possible, Tx and Rx differential pairs should route alternately on the same layer (Tx pair next to a Rx pair rather than another Tx pair) to help minimize FEXT.



PCB stackup for a 4-layer system board uses microstrip trace routing. However, microstrip traces inherently display greater skin-effect loss and impedance variation than stripline traces. Thicker dielectrics and wider traces demonstrate less loss.

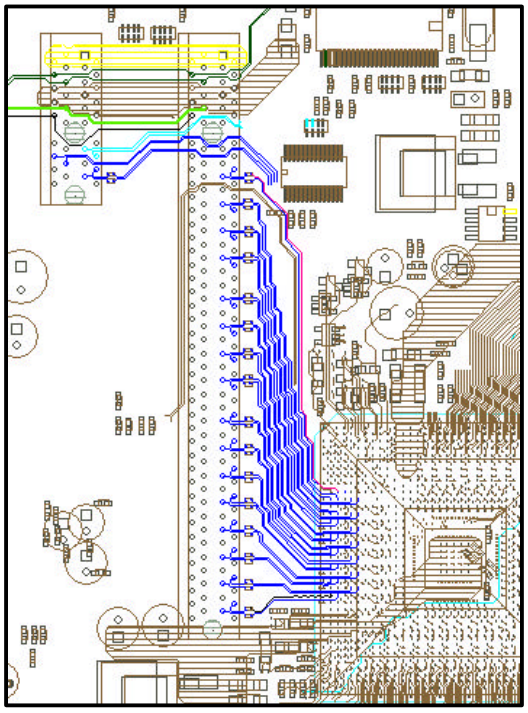
Trace lengths are determined from simulation analysis in order to meet the interconnect budget. Trace Shorter lengths may be used to tradeoff with more vias, tighter spacing, etc. on the interconnect. Inherent physical properties of the PCB trace can introduce as much as 1 to 5 ps of jitter and 0.35 to 0.50 dB of loss per inch per differential pair. For the add-in card form factor currently specified, trace length from the edge-finger pad to the device is limited to 3 inches. The exact trace lengths allowed for a given topology will be determined by analyzing the design tradeoffs for the interconnect.

Trace Length matching between pairs is not required due to the embedded clock architecture and generous pair-to-pair skew allowance. However, it is desirable to keep the length differences small to minimize latency.

Trace Segment Length matching within pair is required to ensure trace lengths are equal on a segment-by-segment basis. Length should be matched to less than a 5 mil delta overall. Each net within a differential pair should be length matched whenever possible on a segment-by-segment basis at the point of discontinuity. Examples of segments might include breakout areas, routes between two vias, routes between an AC coupling capacitor and a connector pin, etc.

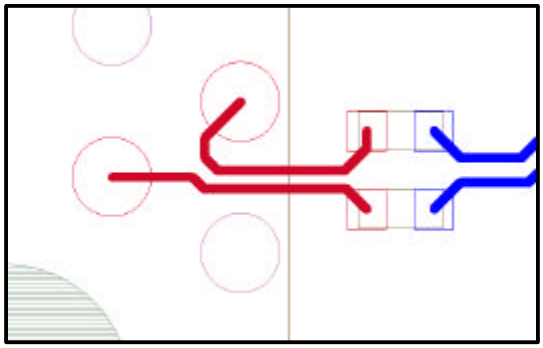
The points of discontinuity would be the via, the capacitor pad, or the connector pin.

Trace Symmetry is required between two traces of the same differential pair.



Vias are expected to contribute 0.5 to 1.0 dB per via to the loss budget. Limit the number to two vias on each trace for the differential pair on motherboard if possible. Vias on the add-in card should be limited to one near the breakout section and one at the edge finger. Vias should have minimum pad size (25 mils or less) and a small finished hole size (14 mils or less). Vias should not have pads on unused internal layers. Vias on the differential pair should not only match in number but in relative location on the differential pair to maintain trace symmetry.

Bends on traces should be kept to a minimum. If bends are used, they should be at a 45-degree angle or smaller. Attempt should be made to match the number of left and right bends as closely as possible to minimize skew due to length differences between each signal of the differential pair.



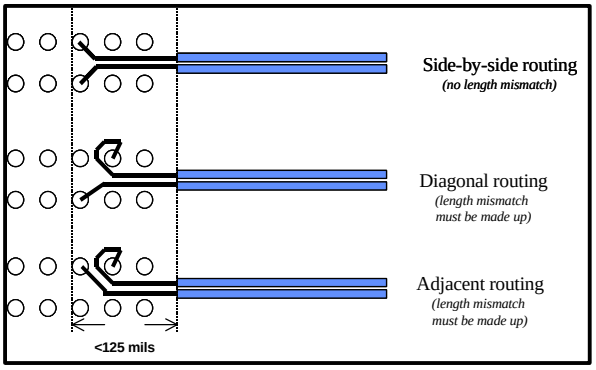
AC coupling capacitors of 75 nF to 500 nF should be placed at the same location (as close as possible) and should not be staggered from

one trace to the other within the pair. While size 603 capacitors are acceptable, size 402 capacitors are strongly encouraged. C-packs are not allowed for AC coupling capacitors. The exact same package size of capacitor should be used for each signal in a differential pair. Pad sizes for each of the capacitors should be minimized. The "breakout" into and out of the capacitors should be symmetrical for both signal traces in a differential pair.

Connector pins of a differential pair are offset from each other. This delta of mismatch between the pins should be directly accounted for by the PCB trace on the motherboard. Both traces of a differential pair should both route into a connector pin field from the same layer.

Ground Plane Referencing is required along the entire route of the differential pair. The signal pair should avoid discontinuities in the reference plane such as splits and voids. Traces routed near the edge of the reference plane should maintain at least a 40 mil air gap to the edge. Any layer transition must maintain a ground reference plane from one layer to another by stitching vias connecting the two ground planes together. Ground stitching vias should be placed close to signal vias. A minimum of 1 to 3 stitching vias per differential pair is recommended.

Breakout Areas near a device package that resulted in "neckdowns" and decreased spacing should be limited to no more than 500 mils in lengths. The necking down should be done symmetrically on both nets of the differential pair. Breakout sections require special attention to minimize crosstalk.



Edge fingers of the specified add-in card are designed to mate with the connector pins to produce the target impedance of 100 Ω. The reference planes under the edge finger pads are removed to meet the impedance.

Test points and probing structures may impact the loss and jitter budgets. If possible, test points and probe structures should not introduce stubs on the differential pairs.

Summary

PCI Express will serve as the general-purpose I/O interconnect for a wide variety of computing and communications platforms. Its advanced features and scalable performance will enable it to become a unifying I/O solution across a broad range of platforms—desktop, mobile, server, workstations, communications and embedded devices. A PCI Express link is implemented using multiple, point-to-point connections called lanes and multiple lanes can be used to create an I/O interconnect whose bandwidth scales linearly and provides ample headroom for performance-intensive applications for the next decade. PCI Express silicon building blocks and platform ingredients are being designed now for initial interoperability validation and testing beginning in late 2003.

The Point-to-point serial differential links, as adopted by the PCI Express, is now being widely recognized as a suitable alternative for high-speed data transfer applications. This technology is highly scaleable in performance with silicon devices of the next decade. Many of the limitations inherent in a parallel, multidrop bus architectures are avoided by the PCI Express technology. This paper has addressed design considerations for PCB engineers for successful implementation of the PCI Express serial interconnect. PCI Express opens up new system level opportunities to allow for creative form factors not possible with current parallel bus designs. Board designers will need to rely on simulation analysis and careful design tradeoffs to ensure a successful implementation.