High-Speed Backplane Connectors

Howard W. Johnson, PhD.

Signal Consulting, Inc.
Twisp, WA U.S.A.
howie03@sigcon.com

***Abstract*—Printed circuit board (PCB) backplanes and backplane connectors form the cornerstone of many large system architectures, particularly in the fields of communication and high-performance computing. This paper investigates the limits to the electrical performance of backplane connectors. It discusses the main factors affecting performance, and predicts the future of backplane connector development.**

I. Introduction

When working with objects on the scale of a printed circuit board, or a backplane connector, if something else does not preclude you from doing so, the simplest and best means of increasing the circuit bandwidth and speed is to physically shrink the circuit. That much seems obvious.

If that were all there were to making good, high-frequency connectors this would be a very short paper indeed. What complicates the story, and makes it interesting, are the multitude of practical factors that prevent shrinking.

This paper begins with a review of the technical criteria associated with the theory of three-dimensional scaling. The review leads to a discussion of physical scaling as it affects three types of systems: integrated circuits, HDMI connectors, and RF connectors. The last section treats the ongoing development of high-performance backplane connectors, and the question of what future changes the author anticipates in connector technology.

A. 3-D Rule of Scaling

Any passive, lossless electrical structure shrunk to half its normal size responds electrically twice as fast as its predecessor [1]. This is the key mechanism by which we make passive circuit go faster: we shrink them. The *response* of the circuit may be interpreted as either the impulse or step response of the circuit, measured either for signals passing through the circuit or for signals reflecting from it.

In mathematical terms, a system **A** having response $h(t)$, when scaled to a new size by multiplying every dimension of **A** by a constant scale factor k, produces a new time-domain response $g(t)=h(t/k)$, and a new frequency-domain response the same as the old, only scaled to a new operating frequency $G(f)=H(fk)$. This is the guiding principle by which we design antennas for operation at different frequencies, or capacitors of varying sizes, or inductors. The rule of scaling is an elementary consequence of Maxwell's electromagnetic equations.

The rule of scaling applies only to circuits that are *passive*, meaning that they contain no sources of power. The rule is also limited to *lossless* structures, meaning that the circuit dissipates no power (or very little power) internally.

The rule of physical scaling applies well to low-loss conducting structures like metal plates, conducting wires, connector pins, and semiconductor packages. Vias, connector pins and BGA balls all scale beautifully. Any metallic interconnection short enough that losses from skin effect resistance, dielectric absorption and radiation do not significantly affect its operation will follow the rule of scaling.

The 3-D rule of scaling is not an approximation, it is an exact solution to Maxwell's electromagnetic equations for passive, lossless systems. In the event that the initial system **A** is not lossless, you may still use the rule of scaling provided that you follow these guidelines [2]:

1) Scale every part of the circuit, in every dimension, by the same factor, k.
2) Ensure that the new dielectric permittivity, dielectric loss tangent, and magnetic permeability have the same values at the new operating frequency as they had in the original circuit at the old operating frequency.
3) Scale the resistivity of the conducting members by the factor, k.

The last condition fixes a problem with the way resistance scales. The resistance of a solid conducting body scales in proportion to its length and inversely with its cross-sectional area. If you divide all the dimensions by two, the length diminishes by a factor of two, but the cross-sectional area, which varies quadratically with size, diminishes by a factor of four. Unless you change in the resistivity of the conductor, the ratio of factors described doubles the overall circuit resistance. Scaling the resistivity in proportion to the physical scale factor k holds the circuit resistance constant, invariant to scale. The resistivity-scaling condition is rarely satisfied, as discussed below in the section on the scaling of integrated circuits.

The rule of scaling is valid over any range of physical scales for which useful conducting objects may be constructed. It breaks down for microscopic structures near the atomic level, for which the conducting surfaces cannot be scaled due to the inherent quantization of atomic matter. As far physicists know, the rule applies to structures of galactic dimensions, although such structures have not been tested to verify conformance with the rule.

B. Integrated Circuits

Imagine a circuit that satisfies the *lossless* condition of the rule of scaling, meaning that the losses are small enough to ignore. The lossless property may tempt you to assume that any new circuit formed by scaling the original will also have losses small enough to ignore, but that is not alway the case.

When scaling a circuit down to the extremely small dimensions common in the IC industry, losses in the scaled circuit may become quite significant, distorting the circuit behavior. Remember that the rule of scaling requires that to maintain an exactly constant level of circuit performance at any scale factor you must scale the resistivity of the circuit conductors in proportion to the overall change in physical scale. If you fail to change the resistivity, because, of course, you can't actually *change* the resistivity of copper or aluminum, then the resistances in the circuit all scale inversely with the scale factor, *k*.

As an example, begin with a connector having 1 mOhm of contact resistance. Scale it down by a factor of ten (k=0.1), keeping the resistivity constant. The new resistivity now ranks ten times more significant than before, at a new level of 10 mOhm. Compared to a 50-Ohm circuit, 10 mOhm may not seem very large, but if you scale by a factor of 1/1000, the resistance jumps to 10 Ohms, a much more noticeable figure.

Circuits on the physical scale of an integrated circuit are so small that resistive losses internal to the circuit conductors may become extremely significant, even if those conductors are formed from highly-conductive metals. When you take into account the resistive losses due to scaling, combined with the effects of lossy poly lines and lossy FET gates as found inside typical integrated circuits, the theory of scaling becomes quite complicated, but its basic character still shines through: physically shrinking a circuit pays great benefits in terms of speed. Figure 1 illustrates the astonishing accomplishments that have already been made possible by shrinking [3].

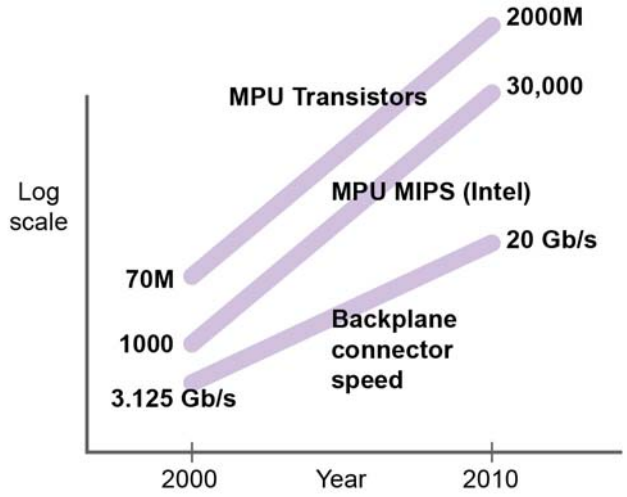

FIGURE 1—MICROPROCESSOR UNIT (MPU) PERFORMANCE HAS accelerated dramatically during the past decade.

Over a period of ten years the number of transistors in a large microprocessor unit (MPU) has grown from 70 million to 2000 million, a factor of 30 times. That was accomplished entirely by shrinking the relative size of each circuit element with no appreciable increase in the overall die size.

Over the same period the number of instructions per second increased from 1000 million instructions per second (MIPS) to 30,000 MIPS, also a factor of 30. That increase in speed was made possible mostly, but not entirely, by shrinking the individual circuit elements, thus permitting use of a faster clock. Although an expert in computer architecture might argue that advances in processor architecture made during the latter portion of the decade contributed substantially to performance, my point here is not to debate the fine points of Moore's Law, but simply to observe that MPU performance has increased quite rapidly, and that physical miniaturization played a great role in that increase.

The attainable minimum feature size in integrated circuits is set by lithography. We can only make things so small. As time passes, we seem able to improve our lithographic processes, but ultimately, new limits will arise due to the atomic structure of matter [4].

Lithographic limitations do not apply to macroscopic objects like backplane connectors, which could, if one wished, be easily shrunk by a factor of ten or a hundred without encountering any lithographic or resistive difficulties.

Yet, Figure 1 indicates that backplane connectors, over a ten-year period, realized an increase in advertised speed capability from 3.125 Gbps to only 20 Gb/s, a factor of only six. Why haven't backplane connectors kept up with the increase in speed of MPU's? Perhaps it is because they have *failed to shrink.*

C. HDMI Connectors

It is the general nature of electronics that, over time, things shrink. As an example, Figure 2 depicts the popular series of High-Definition Multimedia Interface (HDMI) connectors used in consumer products for transmitting video data streams [5-7]. This series of drawings, all rendered to the same scale, indicate a progressive shrinking of the mechanical dimensions with each successive version of the standard.

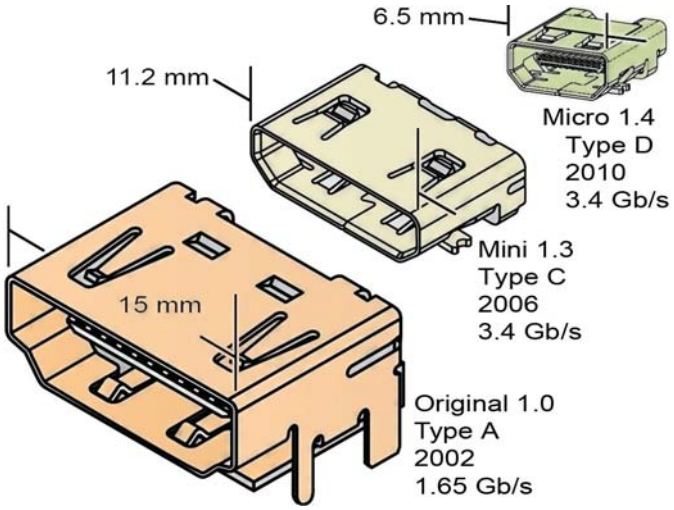

Figure 2—HDMI connectors have shrunk on a regular, progressive basis.

These HDMI connectors are macroscopic, meaning that resistive losses internal to the conductors play a very insignificant role in connector performance. In that case, exact application of the rule of scaling predicts an increase in operational speed in direct inverse proportion to the shrinkage in dimensions. Indeed, the rated operation speed per

differential channel approximately doubled over this period from 1.65 Gb/s to 3.4 Gb/s.

A close examination of the HDMI specifications, however, reveals that connector type C operates twice as fast as type A despite not having been scaled down by a full factor of two. That suggests that the type A connector may harbor some excess capacity.

Looking back over the history of computer development, a generous degree of excess capacity is the usual case for connectors. At each stage of technology, from relay logic to tubes, to transistors, to through-hole DIP-packs, to surface-mounted packages, each technology begins in a package fundamentally smaller than the one before. The packaging is, initially, not a limiting factor for speed. As each technology matures, and the operating speed of the internal workings of each packaged device increases, the technology eventually presses up against the bandwidth limits of its physical packaging. At that point, engineers must devote significant design resources towards the problem of improving the performance of the packaging, and the rate of increase in performance slows.

In every case, the industry breaks out of that slowing trend with the development of a new, smaller package. The introduction of a new packaging standard wreaks havoc within the electronics industry as manufacturers scramble to accommodate new methods of assembly. Billions of dollars worth of old assembly equipment are scrapped in the conversion. The ultimate result is a new point on the packaging performance curve that represents smaller packages with much better high-frequency performance—at which point engineers stop worrying so much about package performance and go back to thinking about system architecture and product features.

The HDMI connectors shrank not for reasons technological, but ergonomic and economic. Consumers like tiny products, and smaller products are cheaper to produce. These factors drive the continual shrinking of consumer-type connectors like the HDMI.

I should point out that it was not *trivial* to produce a connector the size of the HDMI Several innovations were required, including a novel form of torsion-spring contacts.

Limits to the size of the HDMI connector will eventually emerge due to ergonomic and mechanical-ruggedness factors. An HDMI connector must be a certain minimum size in order for humans to see it and gently plug it into its receptacle without breaking the mating half. Mechanical strength, offset alignment, the bending of pins, and the stress of dropping the product onto its mated connector, will all factor into the determination of the ultimate minimum size of an HDMI connector.

Ergonomic factors do not apply to the design of high-performance backplane connectors that are guided into their seats by card guides and alignment pins.

D. RF Connectors

The world of high-precision RF connectors has matured over a much longer period than HDMI. In this world, the requirements of ultimate fidelity and mechanical perfection trump almost all considerations of cost. A single high-end 1-mm connector can cost upwards of US $1000.

One might expect the designers of such connectors to have achieved a degree of perfection close to some maximum theoretical limit. That is indeed the case.

One primary factor limiting the performance of any coaxial transmission medium is the non-transverse-electric-and-magnetic (non-TEM) mode boundary, a frequency above which the coaxial medium develops spurious modes of transmission in addition to the ordinary TEM mode [9-11]. These additional modes propagate at slightly different velocities, with the inconvenient result that signal power input to the coaxial system at one time may emerge later over a smear of different times from the far end of the system.

For a coaxial cable, the non-TEM mode boundary changes inversely with the diameter of the outer conductor (the shield). A smaller diameter works to a higher frequency without experiencing non-TEM mode distortion. That principle motivates the use of small coaxial cables.

The resistive losses internal to a coaxial cable work in the opposite direction. Larger cables experience less resistive loss. That principle motivates the use of large coaxial cables.

An optimal (meaning lowest-loss) design specifies cable with a non-TEM mode boundary just barely above the highest frequency of interest. That will be the largest possible cable that still avoids the non-TEM mode problem.

Available connectors for high-precision RF applications like metrology are categorized according to the size of the coaxial cable shield diameter with which they are expected to work. The geometry of the connector itself is highly constrained as it must emulate, with as little disturbance as possible, the smooth flow of signals along the coaxial structure.

Figure 3 shows end-view diagrams for five types of high-precision RF connectors. The performance of each connector is related to the size of the air gap (dark ring) surrounding the shiny center conductor. The outer conductor diameters indicated on the figure range from 3.5 mm to 1.0 mm. As you can see, the frequency rating for each connector varies in almost exact inverse proportion to its outer conductor diameter.

The pictures are all drawn to approximately the same scale. Note that the external dimensions of the connectors do not vary strongly. Only the internal dimensions of the critical bits vary. The external dimensions of these connector are established not by considerations of cost or connection density, but almost entirely according to ergonomic considerations involving the convenience of test technician who must thread and un-thread the connections with their fingers, and then apply a standard torque wrench to tighten the connection.

Figure 3—The performance of each connector is limited by the diameter of the air gap (dark ring) surrounding the shiny center conductor.

The internal dimensions are fixed by the geometric relationship of the connector to a specific, exact coaxial cable shield diameter determined by the tradeoff between non-TEM mode propagation and resistive losses.

Backplane connectors need not comply with such stringent requirements for the simple reason that we do not have such a high expectation of performance from a backplane connector. A typical backplane connector includes multiple right-angle transitions and abrupt changes in its internal geometry. These transitions and changes create reflections at a level that, while not acceptable for metrology, perfectly suits the needs of digital signal transmission.

Backplane connectors are not subject to exact geometric constraints.

II. Backplane Connectors

The backplane connectors in Figure 4 all belong to a series of parts originally designed by Teradyne in 2002, called the GbX series. The original GbX connector released in 2002 was advertised for use at a data rate of 3.125 Gb/s. The GbX Series-E extended that range to 6.125 Gb/s. In 2007, Molex and Amphenol released another upgrade called the GbX I-Trac, rated for use up to 12.5 Gb/s.

The pictures are all rendered at the same scale. As you can see, each successive generation in this series retains the same physical pin spacing as its predecessor. Successive generations of connectors are getting better in terms of impedance control, or crosstalk, or other differential-mode balance, but not size. The size remains locked in place, preventing shrinkage, by the limitations of PCB construction, described next.

A. Pin Spacing

PCB manufacturing technology, especially regarding high-performance backplanes, has progressed very slowly during the past decade. By the term *progress*, this author refers to two factors, (1) connector pin spacing, and (2) overall board thickness.

Pin spacing affects the overall size and, through the rule of scaling, the overall high-frequency performance of a backplane connector. All other factors being equal, a connector with 10x smaller pin spacing, and 10x smaller features elsewhere, should realize a 10x increase in its data transmission capability. If the performance gains are that good, why are such tiny connectors not on the market? The reason is that such tiny connectors would be incompatible with existing standards for printed circuit board (PCB) construction.

A typical PCB backplane comprises multiple signal routing layers separated by solid reference plane layers and connected by electroplated holes drilled vertically through the board. The plated-through holes are called vias. To support the electroplating operation, the holes must be a certain minimum size. After accounting for all the various tolerances and offsets, the complete via structure including its anchor pads on top and bottom layers attains a minimum diameter on the order of 18-36 mils, depending on board thickness. Vias that harbor press-fit connector pins, or span very thick backplanes, lie at the top end of this range.

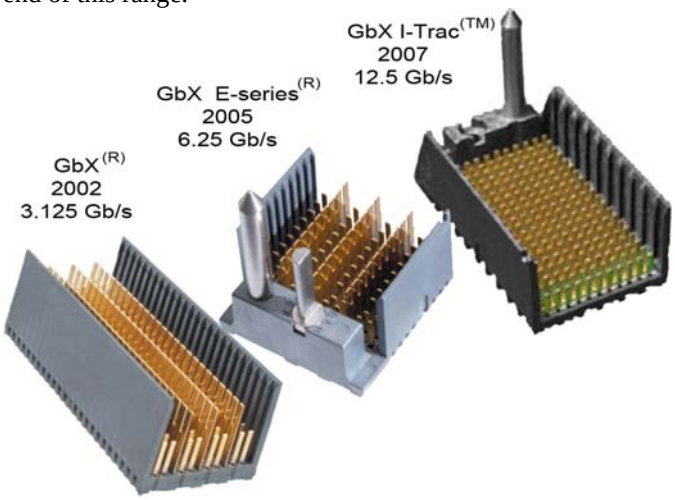

Figure 4—These connectors from Molex achieve improved performance without the benefit of shrinking.

Between vias, the connector designer must leave sufficient room to route pcb traces, otherwise there would be no way to reach pins located deep in the interior of a large field of connector pins. The pattern of traces interfacing to the connector pin field is called a breakout pattern. Figure 5 illustrates two possible breakout patterns. On the left, the via spacing permits double-track routing. Double-track routing, refers to the ability to sneak two side-by-side traces in the space between two adjacent connector pins. On the right the more closely-spaced vias permit only single-track breakout.

Double-track routing cuts almost in half the number of layers required to breakout traces from the connector pins to the surrounding pcb circuits. Large connectors with a deep row structure benefit considerably from double-track routing, therefore high-speed backplane connectors universally adopt a pin spacing sufficient to permit double-track routing.

Connector layouts that support only single-track routing may still be used, but require more layers, and therefore thicker boards, than their double-track-routed counterparts.

Connector layouts so dense that they cannot support single-track routing are not useful beyond a single row, except in some specialized configuration where the connector is

approached from two directions, in which case two rows are reachable.

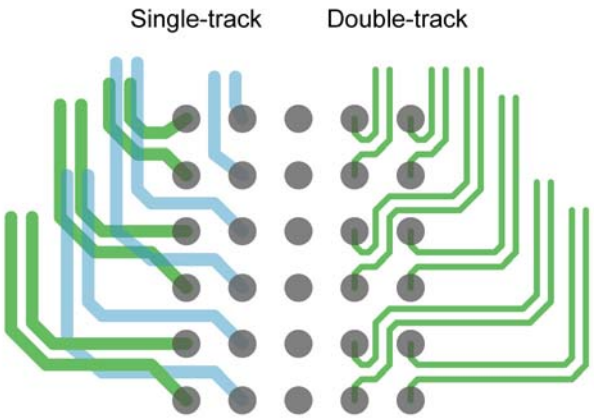

Figure 5—Double-track routing sneaks two traces between each connector pin.

Data in Table 1 from FCI, Amphenol, Ernie and Molex indicate that pin spacings between 1.85 and 2.5 mm are required to provide adequate space for double-track breakout routing. The connectors listed in the table take different approaches to the breakout problem.

The FCI Airmax(R) uses a fairly regular pin grid with a column spacing of 2.0mm, intending that double-track routing be used in the 2.0-mm corridors between columns.

The Amphenol Excede uses an irregularly-spaced array of pins that provides specialized double-track routing corridors in particular locations. The pin spacing on either side of these corridors is 1.85 mm generally, and 1.56 mm at its narrowest.

The Ernie Zd-plus connector uses a regular array of vias on a 2.5x1.5-mm grid, with double-track routing taking place within the 2.5-mm wide corridors.

The Molex connector shifts every odd-numbered column to the left, making every other corridor very wide and the others quite narrow. Molex intends that quad-track routing be used in the wide corridors, which have a spacing of 2.7 mm, and no routing in the narrow corridors, which have a spacing of 1.0 mm. The table lists the effective corridor spacing calculated per note [2] below the table.

Even going all the way back to 1980 (VMEbus), the connector pin spacing has not much changed. The backplane connector industry has undergone a 30-year period with little to no improvement in signal density.

All spacings represent the center-to-center separation of vias on opposites of the routing corridor. The actual routing space available depends on the size of pads surrounding the signal pins, which in turn depends somewhat on the board etching and drilling technology available and the degree of board reliability desired.

TABLE 1

THESE HIGH-SPEED CONNECTOR PRODUCTS USE SIMILAR PIN SPACING.

Mfr.	Product (Year)	Corridor spacing[1]	Density (pair/mm^2)
FCI	Airmax VSe(TM) (2011)	2.0	.119
Amphenol	Xcede(R) (2008)	1.85	.101
Ernie	Ermet Zd plus (2008)	2.5	.089
Molex	I-Trac(TM)[2] (2007)	1.89[2]	.073
FCI	Airmax VS(R) (2005)	2.0	.119
Ernie	Ermet-Zd (2001)	2.5	.089
VMEbus	DIN 41612 (1980)	2.54	.155[3]

NOTE [1]—Effective via-to-via spacing, measured from the center of one via to the next, for vias on opposite sides of the routing corridor.

NOTE [2]—The I-Trac supports two pairs (four wires) in a corridor 2.7 mm wide (center-to-center). Using .88-mm vias that leaves a total of 1.82 mm to hold the four signal wires and five inter-wire spaces. Assuming equal space-and-trace dimensions, that works out to 0.202 mm spaces and 0.202 mm traces, or roughly 8-mil spaces and 8-mil traces. The equivalent via-to-via spacing for a single-pair routing would a corridor 1.89 mm wide (center-to-center) with 0.88-mm vias, leaving a total of 1.01 mm for two signal wires and three spaces.

NOTE [3]—The VMEbus system was a single-ended system with a signal-to-ground ratio of approximately 8:1 for its bus signals. The single-ended pin density is 0.155 signals per square millimeter.

The connectors in table 1 are differentiated by the internal construction of their signal routing pathways and the position and construction of (or lack of) solid ground shield plates separating columns of pins, but are not significantly differentiated by their pin spacing. Pin spacing has remained almost constant over a ten-year period. Backplane connectors aren't shrinking.

Most of the electrical difficulties we have with backplane connectors stem from that fundamental shortcoming.

In addition to the excessive size of the backplane connector itself, the vias underneath a modern high-speed connector exert an astonishing influence on connector performance. One large via with excessive parasitic capacitance can literally destroy the performance of an entire high-speed digital link. Successful high-speed digital designers recognize the importance of good via design.

B. Board Thickness

Thicker boards require large vias, and large vias perform poorly at high frequencies.

Figure 6 plots the insertion loss performance for one differential via pair. This pair is implemented on a board 0.284-in. thick. Backplanes in large communication or high-performance computing system are often that thick. The via size and spacing are designed to create a 100-ohm characteristic impedance *if the via is used throughout its length*. Unfortunately, and this is what makes Figure 6 an interesting test case, the differential signal enters the via at its top layer and exits one layer down, leaving the bottom portion of the via dangling, unused. In that configuration the bottom

portion of the via acts as an un-terminated stub connected to an otherwise perfectly good 100-ohm differential transmission line. At an operating speed of 10 Gb/s, the effects of the stub interfere substantially with data transmission.

Lambert Simonovich [12] analyzed the performance of the via in Figure 6 and provided data for the plot. The "Plain via" plot shows the response of a 30-inch differential channel equipped with good-quality differential vias at each end. The good-quality vias are cut off short, with no dangling stubs. The "With stub" plot shows precisely what happens when the extra-long dangling stubs are included. The stubs create a first-order zero in the transmission response at a frequency just above 4 GHz. The exact frequency of resonance is a function of the length and size of the vias.

The last plot in the figure, labeled, "With stub and termination," shows the effect of a technique suggested by Dr. Nicholas Biunno at Sanmina-SCI Corporation. His technique, called MTSvia™, shunts the end of the dangling stub with a tiny metal thin film or polymer thick film resistor embedded within the pcb. The shunt damps the stub anti-resonance. The damping tends to flatten the overall insertion loss curve at the expense of a flat fixed loss, 10 dB in this example, that reduces the signal level evenly across the whole frequency spectrum. If the fixed flat loss can be made up at the receiver, the stub termination resistor could be a good idea.

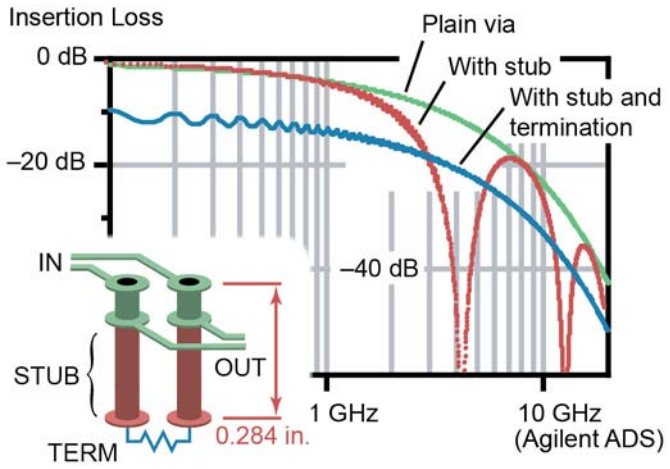

Figure 6—The unused stub dangling below this differential via interferes significantly with data transmission.

I mention stub termination as an example of the extreme design practices to which we are sometimes forced by the poor parasitic performance of overly-large signal vias.

The best way to reduce the impact of signal vias is to shrink them. That is best done in today's technology by either using blind vias or vias back-drilled to eliminate the dangling stubs.

Another way to reduce the impact of vias is to adjust the board's dielectric constant. Lowering the dielectric constant, while holding trace width and trace impedance constant, enables the use of somewhat thinner layers, which shortens the vias. In addition, the dielectric material surrounding the via creates less capacitance. The overall reduction in via capacitance with this technique is roughly commensurate with the degree of reduction in dielectric constant.

Unfortunately, the dielectric constant already stands at about 4. Reducing the dielectric constant to perhaps 3 (an achievable goal) will help, but it's not going to close the connector performance gap in Figure 1.

III. Conclusion

The design of any electrical system becomes substantially more complex as you push up against the ultimate theoretical limits of its performance. The key to making inexpensive, reliable systems is to identify the primary limits to performance and find a way around them, thus removing the obstacles, rather than attempting to scale the face of what may turn out to be some very steep cliffs.

In the case of backplane connectors, the dominating factor that influences connector performance is size. The smaller the connector, the better it works in a high-speed application. Yet, over the last decade, the size of backplane connectors has changed little, almost none in comparison to the enormous improvements made in on-chip geometry.

Why are backplane connectors not shrinking? Backplane connectors do not suffer limits to performance from lithographic, or ergonomic, or geometric considerations as do other interconnection systems. The size of a backplane connector is constrained primarily by historic considerations involving the traditional means of fabricating the vias and breakout patterns on PCB underneath the connector. If we overcome that obstacle, the whole connector can shrink, and performance will skyrocket.

The two papers that follow this introduction each address issues related to the subject of backplane connector performance:

1) Brett Grossman

Careful shaping of the via geometry underneath the connector, back-drilling, wine-glass vias, resistive damping, and other techniques can mitigate the via issues, but all are subject to fundamental limits on the accuracy of manufacturing process control. Brett Grossman's paper in this session covers manufacturing and measurement tolerances and will go into some of the issues limiting our ability to control via geometry.

2) Joe Fjelstad

Shrinking the via geometry may be possible only after a radical re-thinking of the whole PCB assembly process. Joe Fjelstad's paper in this session will address new means of PCB fabrication.

References

[1] H. Johnson and M. Graham, *High-Speed Signal Propagation*, Prentice-Hall, 2003, section 1.3
[2] H. Johnson, "Trace Scaling," *EDN Magazine*, 3/4/2010, http://www.sigcon.com/Pubs/edn/tracescaling.htm
[3] Mark Bohr, "*A 30 Year Retrospective on Dennard's MOSFET Scaling Paper*," IEEE Solid-State Circuits Newsletter, Winter 2007
[4] *International Technology Roadmap for Semiconductors*, www.ITRS.net , updated annually

[5] *HDMI Ver.1.0 Specification*, HDMI Licensing LLC, 2002, www.HDMI.org
[6] *HDMI Ver.1.3a Specification*, HDMI Licensing LLC, 2006, www.HDMI.org
[7] *HDMI Ver.1.4a Specification*, HDMI Licensing LLC, 2010, www.HDMI.org
[9] Constantine A. Balanis, *Advanced Engineering Elecgromagnetics*, John Wiley & Sons, 1989
[10] Dimitrios, James, "Exact Cutoff Frequencies of Precision Coax," *Microwaves*, June 1965, pp. 28-31
[11] Green, Harry E., "Determination of the Cutoff of the First Higher Order Mode in a Coaxial Line by the Transverse Resonance Technique," *IEEE Trans. Microwave Theory and Techniques*, Vol. 37, No. 10, October 1989, pp. 1652-1653
[12] Lambert Simonovich, "Stub Termination," *EDN Magazine*, May 13, 2010, http://www.sigcon.com/Pubs/edn/StubTermination.htm