Next-Generation Interconnect Research at Fujitsu Laboratories

● William W. Walker ● Yasuo Hidaka

This paper explains the need for continual bandwidth improvement in computer server interconnects and the work that is being done at Fujitsu Laboratories to provide the required bandwidth both electrically and, in the future, optically. It explains the concepts of frequency-dependent channel loss and equalization and the hurdles to be overcome to reach 25 Gb/s per lane using electrical high-speed input/output (HSIO) and 40 Gb/s per lane using optical HSIO while improving energy efficiency.

1. Introduction

The commercialization of the Internet in 1994 has led to exponential growth in the demand for communication bandwidth that continues unabated to this day.[1)] Video on demand, smart phones, and social networking are just three recent example applications that are driving the growth in Internet traffic. This bandwidth must ultimately be handled by a computer server in a data center. At Fujitsu Laboratories, we are designing high-speed input/output (HSIO) circuits that send data through computer backplanes.

This article explains some of the challenges we face in designing next-generation HSIO circuits. It explains in layman's terms why it is difficult to increase HSIO speed without affecting energy efficiency and why this difficulty will ultimately drive us from electrical to optical interconnects.

2. Problem of increasing bandwidth

Figure 1 shows our roadmap for HSIO. The x-axis is the year, and the left y-axis is the bit rate in Gb/s (billion bits per second) per lane. A lane is either one pair of wires used for differential electrical signaling or a single fiber or waveguide used for optical signaling. The right y-axis axis shows the target energy efficiency in pico-Joules (pJ) per bit. A pico-Joule is a very small amount of energy, a trillionth of a Joule. A liter of gasoline, for example, contains 32 million Joules of energy. As this roadmap shows, the energy per bit must be reduced while the bit rate per lane must be increased. Increasing the

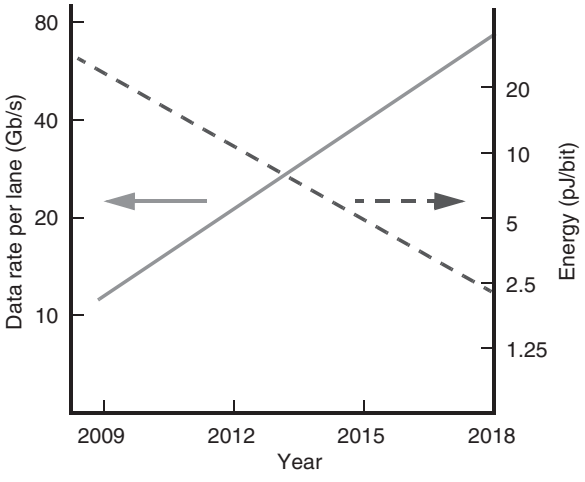

Figure 1
HSIO roadmap.

bit rate per lane is equivalent to increasing the HSIO bandwidth.

Let us continue by developing an analogy. A mountaineer atop a high ridge spots someone on the opposite side of a rocky canyon and calls to him, "Hello, I'm Joe!" (**Figure 2**). The other person hears a distinct but faint "Hello" followed by "I'm" with an echo of "Hello" superimposed and finally "Joe" with a secondary echo of "Hello" and a stronger echo of "I'm" superimposed. As a result, "Joe" is almost unintelligible. If Joe continues to call out in this fashion, his speech blurs into pure cacophony. To solve this communication problem, Joe could break up his utterance, allowing each echo to die out before calling out the next word. But then it would take him a long time to complete his utterance! We might say that his bandwidth, the rate at which information flows, must be reduced to make his speech intelligible.

This is precisely the challenge engineers face every day designing HSIO circuits although we are working with electricity, and (thankfully) our vocabulary consists of just two words, "1" and "0." By the way, communication theorists call vocabulary words *symbols*. In a more complex communication scheme, we might use a symbol set such as [0,1,2,3]. Joe's symbol set is every word in the English language! The medium through which we communicate is called a *channel*. In Joe's case the channel is air punctuated by giant boulders; in our case it is a copper conductor (a trace) on a printed circuit board (PCB) with possible inter-layer connections (vias) to other traces.

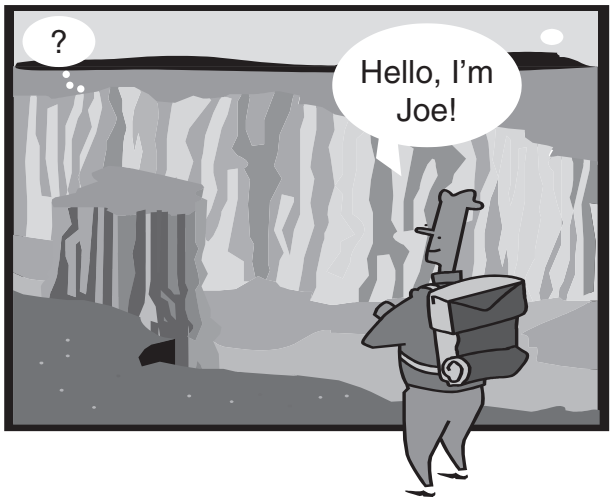

Figure 2
Problem of signal dispersion and echo.

3. Dispersion and echo

Joe's voice sounds faint to the other person because it is dispersed in all directions by the air; electrical signals are dispersed by electromagnetic radiation and resistance in the traces. At first glance, it might seem easy to solve the resistance problem—just make the traces wider. But that would not be effective because the available bandwidth depends on the number of output traces. If we doubled the trace width (pitch), we would have to cut the bandwidth in half unless we can somehow double the signaling speed per trace (symbol rate). At the frequencies we are using today (10 Gb/s) and as a consequence of Maxwell's equations governing electricity and magnetism, forces applied to the electrons cause electricity to flow mainly on the surface of a conductor, i.e., the skin effect. Doubling the frequency increases the resistance because of the skin effect, and we are back where we started.

The bottom line for HSIO circuit designers is that signal attenuation increases with the frequency, making it difficult to simply increase the symbol rate.

But it gets worse. A stream of "1" and "0" symbols at a frequency, f, actually contains energy content at a range of frequencies, and the higher frequencies are more highly attenuated than the lower ones, which causes the symbols to become rounded and bleed into each other. We call this effect *inter-symbol interference* (ISI). This is similar to but not quite the same effect as an echo.

Electrically, an echo is produced when a

signal passes through a discontinuity in a trace, such as a via. An echo causes part of the signal to bounce back until it hits another discontinuity, which causes an even smaller echo to bounce forward toward the receiver. This is analogous to the way Joe's voice bounces from boulder to boulder as it travels across the rocky canyon. The electrical effect is similar to the auditory effect: the message becomes garbled due to overlapping symbols (although an electrical echo is actually a derivative—the slope—of the original signal).

Figure 3 illustrates the combined effects of frequency-dependent attenuation and an echo on a digital signal. In the figure, we show a raw bit stream as transmitted and then the same bit stream after passing through an echo-free lossy channel. Finally, the figure shows that an echo is just a faint displaced derivative version of the bit stream that is added to the signal, producing the final distorted bit stream at the receiver.

As we mentioned before, we are thankful that our symbol set consists of only "0" and "1." This enables us to use some fairly simple electronic techniques, called *equalization*, to partially cancel the ISI and echo.

4. Equalization

Equalization is an electronic technique used to amplify or attenuate certain frequencies to compensate for an undesirable physical effect. The process is reversed, as necessary, at the receiver so that the output signal matches the original input signal. Equalization has a long history. For example, it led to the development of the long-play (LP) record—attenuating the bass frequencies enabled the grooves to be placed closer together. The attenuation was reversed at the playback amplifier to make the bass sound natural.

In our latest state-of-the-art backplane HSIO transceiver (transmitter plus receiver),[2] we use three types of equalization (**Figure 4**).

1) Pre-emphasis: Pre-emphasis adds or subtracts a fraction of the preceding or following bits to each transmitted bit to counteract the effect of ISI.
2) Continuous-time linear equalization (CTLE): An analog circuit in the receiver amplifies the higher frequencies more than the lower ones to reverse the effect of frequency-dependent loss in the channel. This helps to cancel ISI. As shown in Figure 4, the CTLE circuit partially reverses

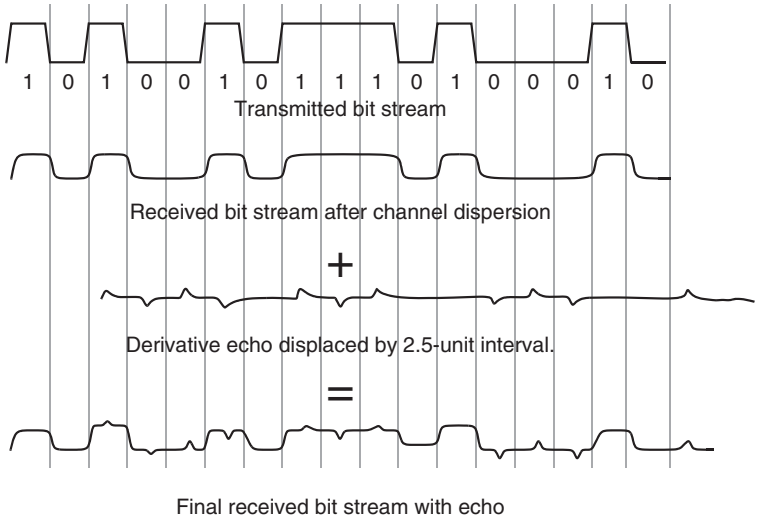

Figure 3
Problem of dispersion and echo in digital world.

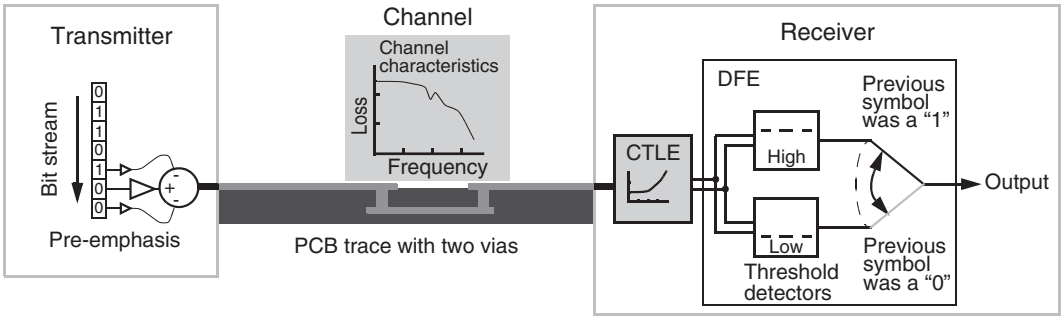

Figure 4
Three types of equalization used in our latest state-of-the-art backplane HSIO transceiver.

the frequency-dependent loss of the channel.

3) Decision feedback equalization (DFE): After a bit is received and passes through the circuitry to determine whether it is a "1" or a "0," it is used to set the voltage threshold for deciding whether the subsequent symbol is a "0" or a "1." Since a "1" causes the voltage of the subsequent symbol to be a little higher, DFE has an ISI-cancelling effect. The DFE circuit is represented by a simple switch in Figure 4.

DFE can be used to cancel echo by subtracting fractions of symbols received at any time in the past from the current symbol. Each unit interval (UI) delay in which a DFE circuit is added is called a *tap*. However, a significant amount of power and silicon area is needed to add DFE taps, so a more practical approach is to eliminate echoes by careful design of the PCB.

At Fujitsu Laboratories, we are using advanced three-dimensional (3D) electromagnetic field solvers to help design improved vias in PCBs. By carefully designing the surrounding wiring and power supply layers, we can enable electrical signals to move smoothly from layer to layer, thereby minimizing the echoes and eliminating the need for power-hungry DFE taps in our receivers. We are also exploring the use of PCBs made from more advanced materials that reduce the frequency-dependent losses. The use of such PCBs should enable us to reach speeds of 25 Gb/s per lane or higher by the middle of 2012.

5. Future interconnects

The laws of physics will eventually thwart our attempts to increase signaling speed on electrical wires. In the past, electrical interconnects have benefitted from process and voltage scaling, both of which drive down power requirements. But scaling is nearing the physical limits imposed by the manufacturing process and the physics of silicon itself, while higher signaling rates demand more complex circuits to overcome the channel impairments discussed above. What's next?

For distances greater than 100 m, electrical signaling has been supplanted by optics, with parity having been reached for distances of 10–100 m. Engineers are capable of producing extraordinarily low-loss high-bandwidth channels using optical fibers, thanks in part to the pioneering work in purifying glass by the 2009 Nobel Laureate Charles Kao.[3] The next frontier for optics is the backplane. Short low-loss backplane optical channels eliminate the need for equalization, thereby simplifying circuits and reducing energy use.

There are several challenges related to the use of optical backplanes.

1) The vertical cavity surface emitting laser (VCSEL) light sources they use require a significant amount of power to laze. This handicap becomes less of a problem as signaling rates increase because the overhead power is amortized over more bits.

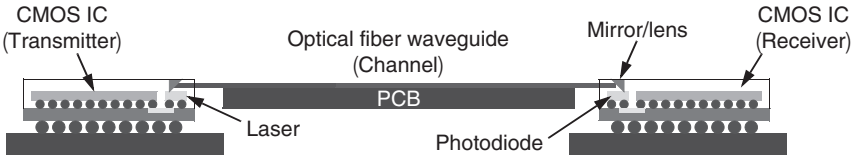

Figure 5
Optical HSIO link.

2) The optical components, packaging, and channels are extra components that add to the cost of the backplane.
3) Servers need hundreds to thousands of lanes of high-density interconnects, so the footprint of existing electro-optical modules is too large.
4) It has not been possible to modulate a signal onto a laser or receive a signal from a photodiode without using expensive-compound semiconductor integrated circuits (ICs).

At Fujitsu Laboratories, we are developing electro-optic conversion circuits using mainstream microprocessor complementary metal oxide semiconductor (CMOS) IC technology. Indeed, it should be possible to have an optical port on a microprocessor in the future. We have already demonstrated the ability to use CMOS technology at a bit rate of 40 Gb/s per lane for single lane telecommunications transponders.[4] We are currently working to adapt this technology to the server backplane through a combination of innovative CMOS circuit design and high-density, low-cost optical component integration and packaging (**Figure 5**).

6. Conclusion

Exponential growth in Internet traffic demands exponential growth in the bandwidth of the computer interconnects that serve up the bandwidth. Through CMOS process scaling and deployment of advanced adaptive equalization circuits and backplane channel design, Fujitsu Laboratories researchers have advanced the state-of-the-art to 10 Gb/s per lane at a respectable energy requirement of 20 pJ/bit, with 25 Gb/s on the way, thereby maintaining leadership on the roadmap. In addition, our researchers are starting to focus on optical interconnects to extend the bandwidth to 40 Gb/s and above while reducing the required energy to below 5 pJ/bit.

References

1) Minnesota Internet Traffic Studies. *http://www.dtc.umn.edu/mints*
2) Y. Hidaka, et al.: A 4-Channel 10.3 Gb/s Transceiver With Adaptive Phase Equalizer for 4-to-41 dB Loss PCB Channel. *Intl. Solid State Circuits Conference Digest of Technical Papers*, pp. 346–348, Feb. 2011.
3) K. C. Kao, et al.: Dielectric-fiber waveguides for optical frequencies. *Proc. of the Institution of Electrical Engineers*, Vol. 113, No. 7, pp. 1151–1158 (1966).
4) N. Nedovic, et al.: A 3-Watt 39.8–44.6 Gb/s Dual-Mode SFI5.2 SerDes Chip Set in 65nm CMOS. *IEEE Journal of Solid State Circuits*, Vol. 45, No. 10, pp. 2016–2029 (2010).

William W. Walker
Fujitsu Laboratories of America, Inc.
Mr. Walker is engaged in the research of high-performance CMOS memory circuits, high-speed electrical transmission I/O circuits, and high-speed optical transmission I/O circuits.

Yasuo Hidaka
Fujitsu Laboratories of America, Inc.
Dr. Hidaka is engaged in the research and development of high-speed electrical transmission I/O circuits for server systems.