Multi-Gigahertz Source Synchronous Testing of an Optical Packet Switching Network

C.E. Gray[1], O. Liboiron-Ladouceur[2], D.C. Keezer[1], K. Bergman[2]
1 - Georgia Institute of Technology
2 - Columbia University

Abstract

A programmable FPGA-based digital logic circuit is enhanced with high-speed, positive emitter-coupled logic (PECL) to create a standalone tester for a self routing optical packet switching network. This tester serves as a test bed for evaluating high-speed electrical to optical conversion techniques and the difficulties associated with burst packet transmissions across the switching network. The demonstration system supports a four-bit wide bus at 2.5-4Gbps per channel. With wavelength-division-multiplexing (WDM) the bit-parallel message width is extendible to hundreds of channels. The project uses commercially available components to keep costs low, yet achieves performance characteristics comparable to (and in some ways exceeding) more expensive ATE. A CMOS FPGA-based logic core provides flexibility, adaptability, and communication with controlling computers while positive emitter-coupled logic achieves multi-gigahertz data rates with about ±25ps timing accuracy.

1. Introduction

The increasing demand for throughput in digital systems is revolutionizing the I/O protocols in use, leading many designs to incorporate source synchronous buses such as HyperTransport and PCI Express. However, aspects of the source synchronous interfaces such as high data rates, differential signaling, and the use of embedded or parallel clocks makes testing these I/O buses difficult with conventional automated test equipment (ATE) [1,5]. Some techniques have been explored to use multiplexing of ATE channels to achieve testing rates of 2-4 Gbps [4] and interfacing to source synchronous devices [5]. However these solutions are too unwieldy and expensive for testing of the data vortex.

The data vortex switch architecture is an optical switching fabric being developed at Columbia University [6] to address the needs of high-performance distributed computing systems. This network can be used for ultrahigh capacity, low-latency transfer of small, source synchronous data packets within clusters of supercomputers while remaining transparent with respect to the data transmitted. The architecture of a 12x12 port version of the network is shown in **Figure 1**. The topology is described using three parameters: cylinder, angle, and height. The input signals are injected at the outermost cylinder and travel to the appropriate height, angle, and ultimately exit at the destination node on the innermost cylinder.

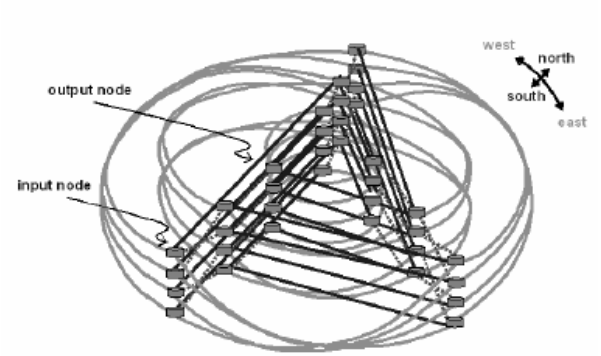

Figure 1. Optical Packet Switching Network (12x12 Data Vortex) [6].

Some of the main concerns with testing this network are related to the conversion of high-speed electrical signals that must be optically modulated and precisely aligned before being injected into the network. A similar conversion back to electrical signals must also be made at the destination. Injection and recovery of the data is made even more difficult due to the burst nature of the traffic. Additional concerns include time dispersion between the WDM channels and amplification noise within the network resulting in bit errors.

To solve the difficult testing requirements for this project, a self-contained miniature tester was developed, expanding upon the concept of a digital test core [8,9]. While such a tester does not have the range of features typically provided by traditional ATE, it can be designed to provide the specific test features needed for a particular application, often at a lower cost than that of commercial ATE.

A general introduction to the design approach for this miniature tester and initial transmission results were introduced at the 2005 DATE conference [10]. Expanded descriptions of the optical packet switching network, tester electronics, and packet transmission are given in more detail in sections 2-5. Section 6 presents new measurements of reception and deserialization of the source synchronous signals. Section 7 discusses some of the current results and possible future development.

2. Design Overview - Test Bed Application

A miniature tester was designed to interface multi-gigahertz signals to optoelectronic components, bridging the high-speed electrical system to the optical network. The present system is intended as a test bed for the evaluation of various optical-to-electrical (O/E) and electrical-to-optical (E/O) techniques at a data rate of 2.5 Gbps per channel. However, the existing miniature tester has demonstrated a maximum data rate of 4.0 Gbps and similar techniques have yielded results up to 6.4 Gbps[11]. Various signaling

protocols are also being evaluated for the transmission of data packets through the data vortex.

For the test bed we create five high-speed data channels for both transmitting and receiving, to support a 4-bit parallel data word and source-synchronous clock. A framing bit and four lower-speed header bits for routing address information for the data vortex are also generated. The high-speed signals are indirectly modulated using an optical modulator on a continuous wave DFB laser, while the lower-speed signals are directly modulating DFB lasers. The optical signals are then multiplexed onto a single fiber and injected into an input port on the network. The data vortex supports dense wavelength-division-multiplexing (DWDM) which can be used to greatly expand the parallel data width in future design iterations. The end-application will require extending the word width to at least 64 or 128 bits. Combined with increasing channel data rates to a future target of 10 Gbps at each wavelength, the final aggregate data rate will be on the order of a Terabit-per-second.

The test bed electronics consist of high-speed PECL circuitry surrounding and controlled by an FPGA-based digital logic core (DLC). An example arrangement of the overall system is shown in **Figure 2**. Since the proposed use for the optical network is for high-performance distributed computing systems, testing of the network requires the creation of very precisely-timed logic signals to emulate the behavior of a parallel slice from a microprocessor-to-memory communication channel. For simplicity, the same tester board functions as transmitter and receiver, though both operations are physically and logically separated on the board.

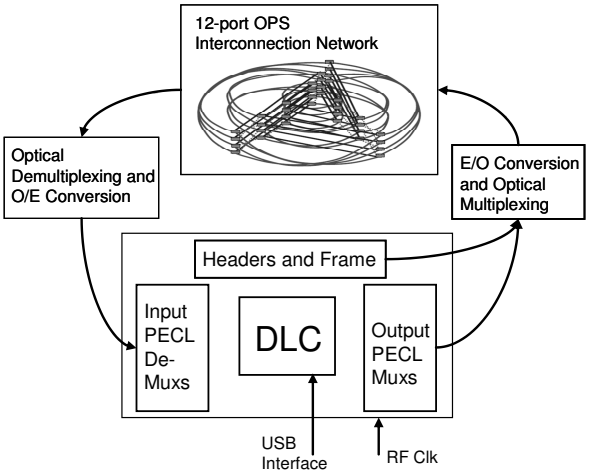

Figure 2. Optoelectronic test application (Data Vortex).

The test bed will also be used to demonstrate different communications protocols and evaluate the timing requirements of the optical network. An example of the desired packet structure is shown in **Figure 3**. Here four example data signals (each 32 bits in length) are synchronously produced in parallel to emulate part of a much wider data bus. These are precisely aligned in time with a source synchronous reference clock. The Frame bit signals when the data is valid and the four header channels carry the routing address data, which is used within the data vortex along with the frame to transparently route the message to the desired port.

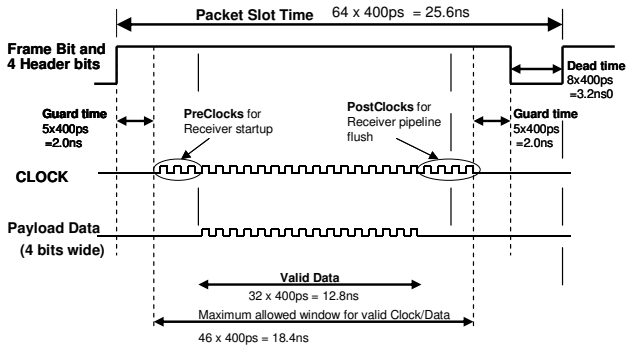

Figure 3. Test stimuli signals needed for the Optical Test Bed application [10].

The individual fiber-optic interconnections of the data vortex are sized to match a packet injection once every 25.6ns or 64 bit periods. The extra 32 bits of length relative to the embedded data is to accommodate a setup time before the data, a pipeline flush after the data, and guard times along either edge. The sizable guard times are required due to the routing nature of the data vortex. The network does not buffer packets, instead deflecting them away from occupied segments. A virtual buffer is created by allowing a packet to circulate around a cylinder until the exit port or routing path is available. A side effect, however, is a slight trimming at the extreme edges of the packet as each routing decision is made.

Due to these strict timing constraints and the need to characterize the system, the relative placement for leading and trailing edges for the data, frame, and header signals must be controlled with 10ps resolution in the optical test bed. A 10ns range for the placement of these edges is also required. Full control is maintained with the bit programming, allowing for a shift of the patterns within the valid window without adjusting the overall channel delay. This control also extends to the clock signal since it is implemented in the same fashion as the other data channels.

3. Test Bed Electronics and Digital Logic Core

A high level block diagram for the test bed electronics is shown in **Figure 4.** An FPGA-based digital logic core (DLC) forms the core of the design, connected to a personal computer (for configuration and control) and the PECL circuitry (for interfacing with the data vortex). Parallel data words are transmitted between the DLC and the high-speed PECL, which is used to multiplex or de-multiplex the higher speed signals transmitted through the data vortex itself. The DLC produces several hundred moderate I/O speed (100-400 Mbps) signals. These are formatted or multiplexed using PECL devices to create sub-nanosecond bit periods and multi-gigabit-per-second signals. An RF clock source (from an external instrument) provides a low-jitter timing reference. This serves as both a master clock (for synchronous applications) and as a reference for all timing-critical signals. In some applications, the RF clock is also provided to either the DLC and/or to the device under test (DUT). However, packet injection is handled for this network synchronously by

the tester and the data sampling clock is embedded in the packet alongide the data. A personal computer communicates through a Universal Serial Bus (USB) with the DLC, and provides high-level control of the tests (which otherwise are synthesized realtime in the DLC).

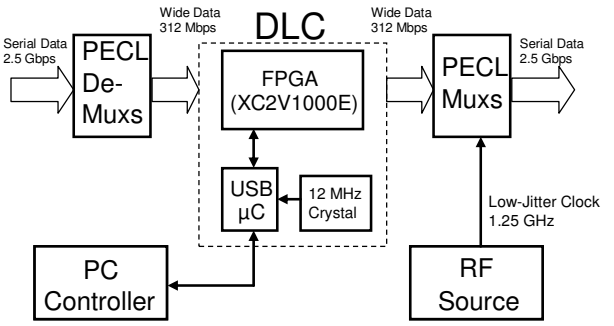

Figure 4. Using a programmable Digital Logic Core with high-speed PECL for testing an optical network.

Some of the key features of the Digital Logic Core design are illustrated in **Figure 4**. The central component is a 1-million gate FPGA (Xilinx XC2V1000), with over 200 I/O, each capable of running up to 800 Mbps. In addition, the DLC includes a specialized microcontroller chip for interfacing to a Universal Serial Bus (USB). Supporting these are a 12 MHz crystal oscillator for USB communications and a FLASH memory to store the FPGA programming information. The FLASH is programmed from a personal computer through an IEEE1149.1 (boundary scan) interface. Once programmed, it loads the personalization data to the FPGA upon power-up. The program can be changed by overwriting the FLASH. This feature is very useful for quickly adapting the DLC to handle new test applications, or to make corrections in an existing design. State machines encoded in the FPGA, together with higher-speed PECL multiplexers and sampling circuits, synthesize the desired tests in real-time.

About 200 signals are available from the FPGA that serve as general-purpose I/O to support specific test applications. In principle, these are capable of running at 800 Mbps, although we typically limit them to 300 or 400 Mbps in order to maintain a sufficient design margin. In some applications, these signals can serve directly as I/O for testing the DUT. However, higher (multi-gigahertz) speeds are obtained using additional PECL multiplexers, timing generators, and sampling circuits (see Sections 4 and 6).

Figure 5 shows a cross-section view of the test bed electronics. The DLC is mounted on a secondary circuit board connected to the main PECL circuitry through a set of high-speed connectors. These connectors are impedance matched to accommodate the highest speed signals that the FPGA can generate and maintain signal integrity. This daughter-card configuration slightly increases the overall design size and signal latency, but is easily offset by the gains in testing flexibility. For instance, multiple DLCs can be kept on hand programmed for a variety of test applications. This also allows for upgrading the FPGA to newer, higher performance cores or adding external SRAM without a redesign of the PECL systems. Similarly, the PECL can be redesigned while reusing the DLC, which accounts for a significant portion of the system cost.

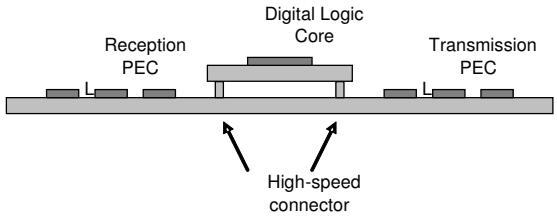

Figure 5. Cross-section view of Optical Test Bed electronics.

Figure 6 shows a current version of the test bed electronics. The DLC is mounted in the center of the board via high-speed socketed connectors. Visible at the lower right corner of the DLC is the USB connector for communications with the controlling PC. The PECL circuits for the transmitting and receiving functions can be seen on the larger, main board. The high-speed output data channels, including the source clock, are to the lower right while the slower speed frame and header signals are at the top. Input logic for the frame and high-speed data/clock is located to the right and the RF clock enters from the bottom. SMA connectors are used to interface differential PECL signals to removable electro-optic modules (not shown).

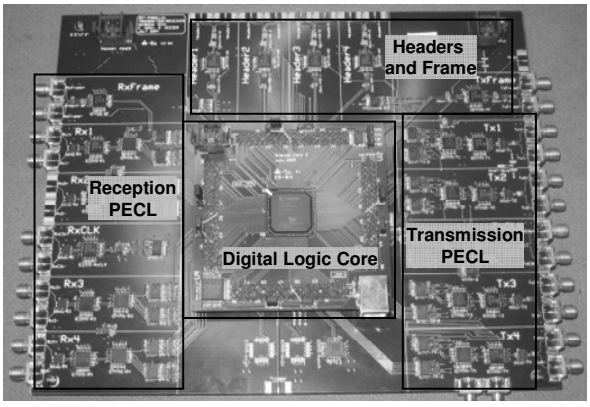

Figure 6. Transmitter and Receiver electronics used for the Optical Test Bed application.

4. Signal Transmission

A functional block diagram of the PECL circuitry for signal transmission is shown in **Figure 7**. The DLC controls programming of the delay over a parallel bus, frame/header formatting, and eight bits of parallel data for each of the high-speed channels. The frame, header, and parallel data bits are transmitted at 312.5 Mbps differentially and the high-speed channels are further multiplexed up to the target rate of 2.5 Gbps. All signals then pass through delay buffers before the output drivers, providing the ability to independently shift all channels by up to 10ns with 10ps resolution. This is useful for characterization of the data vortex, signal alignment for

testing various protocols, and ensuring timing accuracy across the parallel channels.

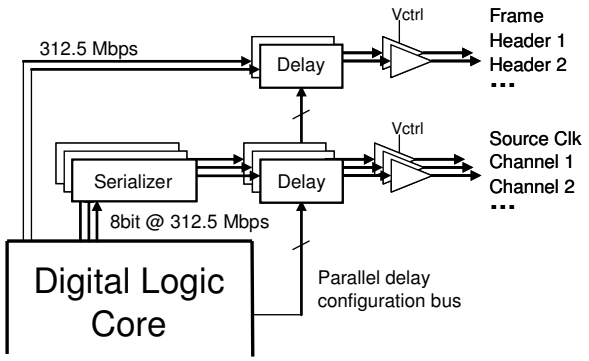

Figure 7. Transmission Logic Diagram.

Figure 8 shows an eye diagram with the output operating at the current project target rate of 2.5 Gbps. For this test, the output waveform is a pseudo-random bit pattern produced by a linear-feedback shift register (LFSR) encoded into the DLC. In addition to fast rise and fall times, the silicon germanium (SiGe) output buffers introduce very little jitter, which was measured at the crossover point. For a 2.5 Gbps signal, jitter was measured to be 46.7ps peak-to-peak, resulting in a usable eye opening of 0.88 unit intervals (UI). While the current data vortex has been designed for operations at 2.5Gbps, the test electronics have been used to generate sample signals at 4.0Gbps with similar results. The measured jitter at the crossover point was 47.2ps p-p with a usable eye opening of 0.81 UI and no visible signal attenuation. This bit rate is at the upper limit of some of the individual PECL components. We independently measured the 20% to 80% rise and fall times on a single edge and found them to be in the range of 70 to 75ps. The jitter on these individual transitions was only 24ps peak-to-peak (about 3.2ps rms).

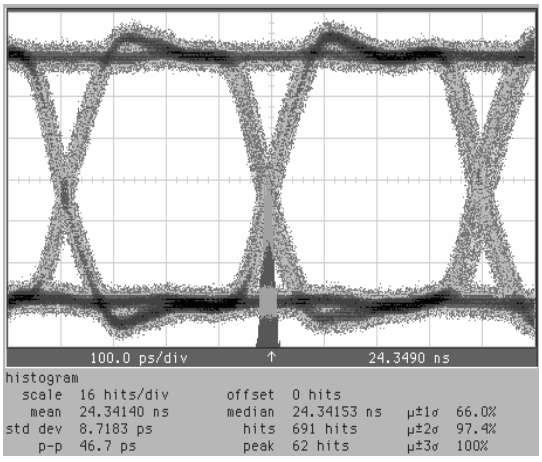

Figure 8. Example 2.5 Gbps eye diagram [10].

5. Electrical to Optical Interface

While a major focus of this design is creating a system that can characterize the data vortex as a whole, testing of current and future electro-optic components was an important design consideration. Therefore features were designed to help test these components and adjust to a wide range of output and input specifications. The most basic and previously discussed feature is the time shift/delay buffer. These buffers are used to align the output signals and accommodate for timing differences between the slow and high-speed E/O components.

A set of signals programmed for 10ps steps are shown in **Figure 9**. The delay chip uses a cascade of delay stages, each of which corresponds to a power of two multiple of 10ps (10ps, 20ps, 40ps, etc). A 10-bit configuration word selects which individual stages are activated and the total resulting delay is the signal offset. However, each bit has some percent error associated with it which results in some non-linearities as shown in the figure. To truly achieve 10ps timing accuracy, this error must be measured and calibrated. Once the actual delay incurred by the individual stages is known, the controlling software on the personal computer can calculate a more appropriate program value [7]. For instance, an uncalibrated setting for 150ps of delay (a configuration word of '0000001111') may actually result in a value of 180ps. The calibrated software can then drop the low two configuration bits, disabling the 10ps and 20ps stages, and come much closer to the desired 150ps target.

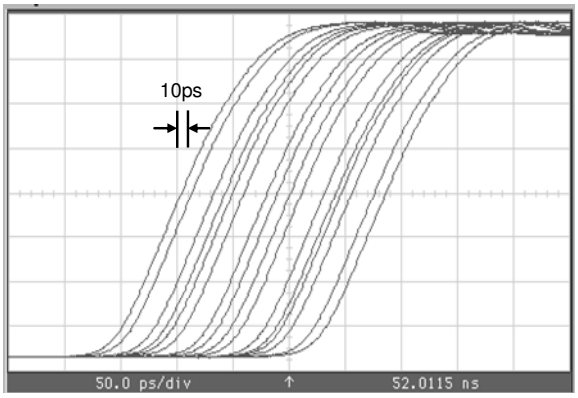

Figure 9. Time shifted signal in 10ps (programmed) steps.

The final output also uses receiver/drivers designed with the ability to adjust the reference voltage levels. Peak-to-peak output levels of the differential signal can be varied from 100mV to 750mV by changing an analog reference voltage. The current design uses a trimmer potentiometer, although an digital-to-analog converter could also be used for higher resolution control. Similar adjustment is available on the low logic level by varying the positive power supply slightly. By coupling these two values, a wide range of amplitude swings and midpoint bias values can be generated for characterizing the data vortex performance under non-ideal signal conditions and with various E/O modules [10].

An example packet generated by the current electronics is shown in **Figure 10**. The signals follow the format and timing as presented in **Figure 3** earlier. The data is 32 bits in length, shown below as an alternating pattern of '1010' over the entire length, while the clock extends before and after the data. Since the clock is not continuous between packets, these extra transitions are required to setup the serializer and

flush the final data word to the FPGA. The packet slot time is 64 bit periods or 25.6ns, and the data/clock signals are roughly centered between the 7/8^{ths} of the packet where the frame is valid. This leaves time on either edge of the data for the proposed guard times. However, due to the programmable nature of the packet structure, the whole pattern can be shifted an integer number of bit periods in either direction independent of the frame/headers.

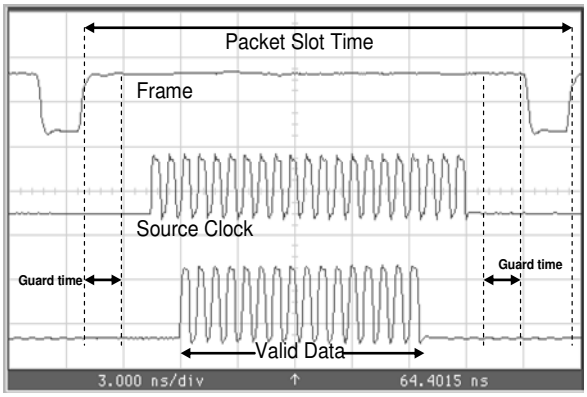

Figure 10. Test stimuli signals used for the Optical Test Bed application.

6. Signal Reception and Deserialization

A block diagram of the receive circuitry is shown in **Figure 11**. All signals are passed through the same type of high performance buffer/drivers that are used in the output stage. The signals are then immediately passed through another delay buffer which allows for deskewing the channels individually. This helps to ensure timing accuracy or allows the simulation of non-ideal conditions within the data vortex. The deserializer is enabled by a valid frame and is clocked by the delayed source clock. The deserializer sends a parallel clock to the DLC to sample the parallel word after every 8 bits of incoming data.

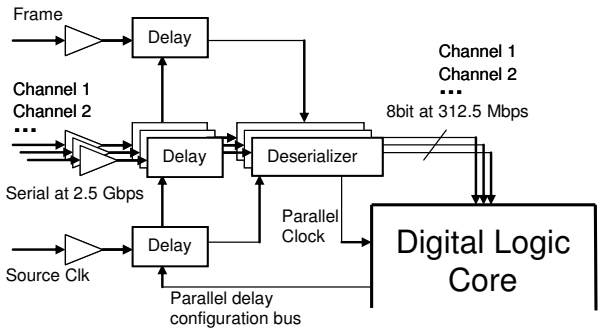

Figure 11. Receive Logic Diagram.

The signal timing relation at the deserializer is shown in **Figure 12**. A few clock edges (not shown) are required after asserting the frame bit to initialize the device. The chip is configured to operate on both the rising and falling edges of the input clock, recovering the serial bits and displaying the output as a parallel 8-bit word six clock periods (12 bit periods) after the first bit of data. Subsequent 8-bit parallel words appear every fourth clock period until the complete payload has been recovered by the FPGA. Due to the piplined nature of this device, an extra two clock periods are required after the final data bit (**see Figure 3**) to 'flush' the system, propagating the fourth and final data word onto the outputs and the corresponding parallel clock to the FPGA.

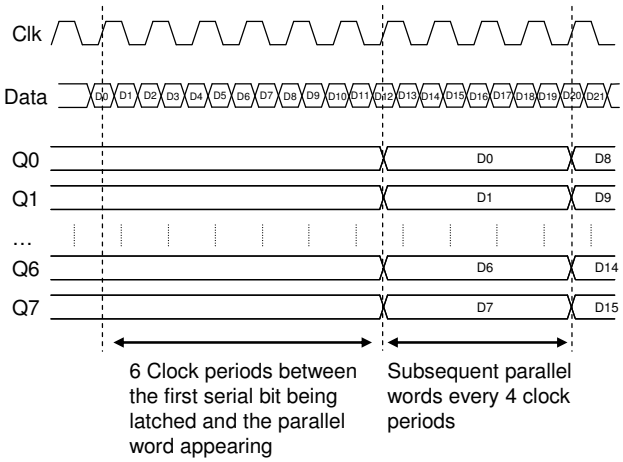

Figure 12. Signal timing of the deserialization process.

Figure 13 shows a single data channel being demultiplexed from the 2.5 Gbps serial stream to 8 bits of parallel data. The upper signal is programmed as '10101010 11111111 01010101 11111111' operating at 2.5 Gbps. The upper and lower row of arrows are indicating the first and second bit of each 8-bit word, which are shown as the middle and lower signals respectively. These 312.5 Mbps signals are read in parallel by the DLC and stored in the local memory or used in real-time testing. (Note: The parallel data has been time-shifted in the below figure. **Figure 12** shows a more representative timing relation between the signals.)

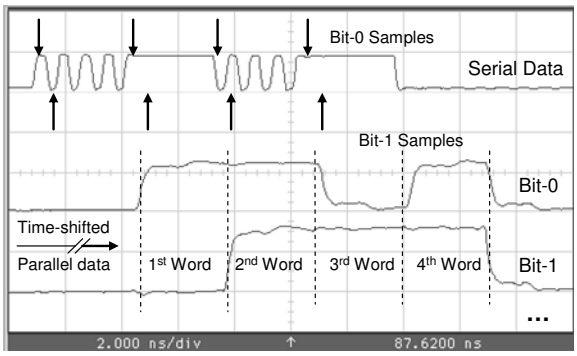

Figure 13. Demultiplexed serial data stream.

Of the 350ps available in the usable eye openings, there is only a only 270ps timing window after accounting for the setup and hold times required by the deserializer and cumulative channel-to-channel skew (**see Figure 14**). The window was measured by skewing the received clock, using the tunable delay, relative to the serial data and sampling the full message length of 32 bits across all 4 channels until no errors were detected in the message. This data was gathered in the ideal case, looping back the electrical data directly from the output buffers and into the recieve circuitry. Signals passed through the electro-optical components and routed through 5 hops in the optical packet switching network exhibited a 150ps timing window due to additional dispersion across the WDM bit-parallel message [12]. As with other source synchronous systems, this loss of timing precision is a central focus for future improvements as it directly impacts our ability to push to faster signaling rates. It also effects the extent to which the parallelism (data width) can be expanded and the scalability of the OPS network with respect to fiber length.

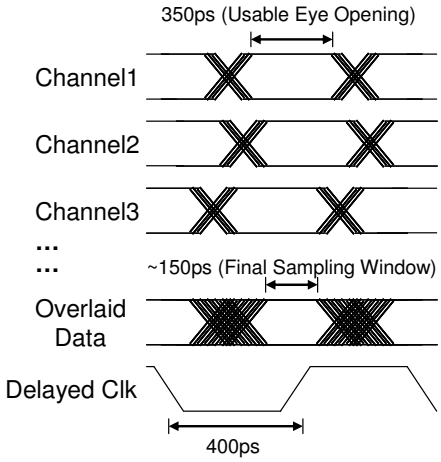

Figure 14. Effect of cumulative channel skew relative to the source synchronous clock.

7. Conclusions

In this paper we have demonstrated how a CMOS FPGA-based DLC can be used as the central controlling logic of a multi-gigahertz source synchronous test source and receiver for an optical packet switching network. The DLC was used to format and align packets for burst injection into the network at 2.5Gbps per data channel, and completely recover the four bit-parallel, source synchronous data channels at the destination with a 150ps timing window.

Many of the techniques explored for this project are also directly applicable to general-purpose electrical testing. The current miniature tester can serve with little or no modification as a relatively low cost, high-precision pattern generator and real-time comparator. It is already being used to explore some techniques to multiplex signals to 8Gbps or beyond. Work is also being done to incorporate a DLC into an existing system that currently uses ATE channel multiplexing. By imbedding a DLC and corresponding PECL circuitry to generate some of the high-speed signals required by the DUT, potentially hundreds of ATE channels could be freed up for other testing needs.

References

1. S. Lomaro, “Testing high-speed serial interface technology: is your test solution in synch?”, *2003 IEEE/CPMT/SEMI International Electronics Manufacturing Technology Symposium*, pp. 437- 438, July 2003.
4. D. C. Keezer, D. Minier, M.C. Caron, “Multiplexing ATE Channels for Production Testing at 2.5 Gbps,” *IEEE Design & Test of Computers,* Vol. 21, No. 4, pp. 288-301, July-Aug. 2004.
5. D.C. Keezer, D. Minier, P. Ducharme, "Source-Synchronous Testing of Multilane PCI Express and HyperTransport Buses," *IEEE Design & Test of Computers*, Vol. 23, No. 1, pp. 46-57, Jan-Feb. 2005.
8. J.S. Davis, D.C. Keezer, “Multi-Purpose Digital Test Core Utilizing Programmable Logic,” *Proc. of the Intl. Test Conf.*, pp. 438-445, October 2002.
9. J.S.Davis, D.C. Keezer, K. Bergman, O. Liboiron-Labouceur, “Application and Demonstration of a Digital Test Core: Optoelectronic Test Bed and Wafer-level Prober,” *Proc. of the Intl. Test Conf.*, pp.166-174, Sept./Oct. 2003.
6. A. Shacham, B.A. Small, O. Liboiron-Ladouceur, K. Bergman, "A Fully Implemented 12x12 Data Vortex Optical Packet Switching Interconnection Network," *J. Lightwave Technol.* 23 (10) 3066-3075 (Oct 2005).
10. D.C. Keezer, C. Gray, A. Majid, N. Taher, “Low-Cost Multi-Gigahertz Test Systems Using CMOS FPGAs and PECL,” pp. 152-157, *Proc. of Design, Automation, and Test in Europe*, March 2005.
11. A.M. Majid, D.C. Keezer, "An Improved Low-Cost 6.4 Gbps Wafer-Leve Tester" *Proc. of the 6th IEEE Electronics Packaging Technology Conference (EPTC)*, pp.814-819, December 2005.
12. O. Liboiron-Ladouceur, C. Gray, D.C. Keezer, and K. Bergman, "Bit-Parallel Message Exchange and Data Recovery in Optical Packet Switched Interconnection Networks," accepted for publication in the next available issue of IEEE Photonics Technology Letters, 2006.