High-Speed Serial Interfaces for Event-Driven Neuromorphic Systems

M. Jabłoński[1], T. Serrano-Gotarredona[2], and B. Linares-Barranco[2]

[1]AGH University of Science and Technology, Department of Automatics and Bioengineering, Krakow, Poland
[2]Instituto de Microelectrónica de Sevilla, IMSE-CNM (CSIC and Univ. Sevilla), SPAIN. bernabe@imse-cnm.csic.es

***Abstract*—Neuromorphic Engineering is the discipline of building sensory processing artificial systems inspired in the neural processing found in living beings. Biological neural brains show massive connectivity among neurons, which is not realistic to mimic using wires within silicon chips or between chips. Address-Event-Representation is a technology widely used among neuromorphic engineers to emulate such massive interconnectivity by time-multiplexing fast digital channels by transmitting "Address Events" between neurons that mimic the neural spikes transmitted in biology. Here we show on-going progress on bit-serial SATA AER inter-FPGA communications for multi-tile scalable neuromorphic systems.**

I. Introduction

Address Event Representation (AER) is a communication technique proposed in the early 90s [1], widely adopted by neuromorphic engineers. Originally, the AER technique would multiplex in time the spike activity of silicon neurons (in the range of a few Hz) into a high-speed asynchronous hand-shaked multi-bit parallel digital bus, transmitting each spike in a few nano seconds. Since then, AER based systems have grown in complexity and resources, requiring higher and higher communication bandwidths and logic resources.

With the growth of neuromorphic systems, processing and communication infrastructures performance requirements have increased. In AER-based systems the connectivity is one of the main performance bottlenecks. Interfaces with parallel digital buses have limited bandwidth due to bus frequency limitations, inter-bit jitter and skew, as well as bus length design restrictions. On the other hand, bit-serial interfaces allow for several Giga bits per second speeds as well as adjustable event bit lengths. The two-wire 4-phase handshake protocol, highly typical in parallel AER links, can be substituted by a flow control mechanism when there are LVDS links in both directions [2].

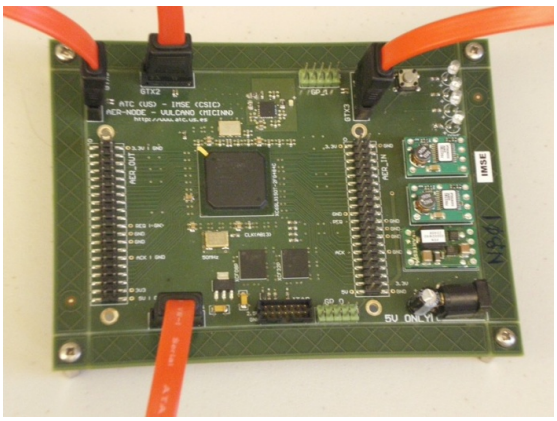
In this paper we show on-going progress on AER bit-serial links between FPGAs, tested at 1.5Gbps, capable of transmitting 32-bit Address Events at a rate of up to 37.43Meps (mega events per second). These links exploit parallel-to-serial transceivers available in Spartan-6 FPGAs. Events are generated by a 32-bit bit-parallel Test Pattern Generator (TPG) running at a higher uncorrelated clock, which interfaces to the transceiver through a clock-domain-crossing interface and high-speed synchronizers.

The link has been tested on a custom PCB that includes one Spartan-6 150t (with serial transceivers) FPGA and 4 bidirectional SATA connectors, which we call the "AER-Node" PCB, and is intended for grid assembly of many of them.

II. AER-Node PCB

The AER-Nodeboard was designed with the aim to demonstrate that spike-processing under AER is feasible and convenient for high speed frame-free vision, filtering, processing and actuation, using spikes from vision sensors to DC motors. Scalability is provided by four SATA connectors for bidirectional LVDS high-speed communications to enable a mesh of AER-Nodeboards [9]. Functionality can be increased with proper daughter boards, through two parallel 28-bit connectors and two 8-bit data connectors.

(a)



(b)

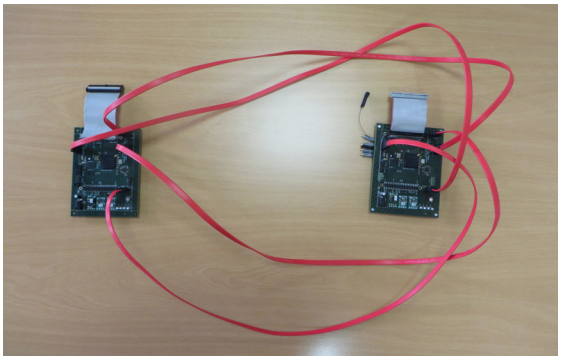


Fig.1:(a) AER-Nodeboard with 4 bidirectional SATA and 2 parallel AER connectors. (b) Test assembly where all 4 bidirectional SATA links are interchanging uncorrelated events in both directions.

For each bidirectional SATA link, Xilinx provides a parallel/serial transceiver wrapper (called here "GTP tile"), synthesized by the Core Generator utility.

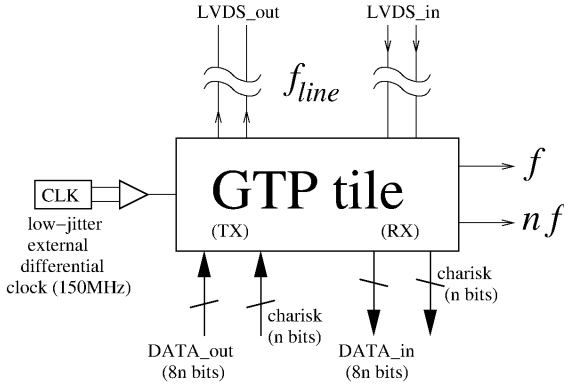


Fig. 2: "GTP tile" transceiver wrapper provided by Xilinx, showing input and output signals and related clocks.

Fig. 2 shows one such "GTP tile" together with its input and output signals and related clocks. All GTP tiles within the FPGA use an external very low jitter reference clock with differential clock lines. These differential clock lines feed into special fabric within the FPGA, and once inside, are transformed from differential to single ended before feeding the GTP tile. In our AER-Node boards we have used external clocks of either 100MHz or 150MHz. The 100MHz reference clock can be used for LVDS data rates of f_{line} = {1.0, 1.25, 2.0, 2.5}Gbps, while the 150Mhz one would provide f_{line} = {0.75, 1.5, 3.0}Gbps. Other reference frequencies are also possible. In the present work we used an AER-Node board with an external clock of 150MHz.

The GTP tile includes a transmitter side (TX) which reads 8n bits of parallel data (DATA_out) and serializes it out to the LVDS_out differential line. When using 8b/10b encoding [11] an additional n-bit flag 'charisk' signals whether the corresponding 8-bit word should be treated as a regular data or a comma command. The GTP tile also includes a receiver side (RX) which performs the opposite. The GTP tile can be synthesized to allow for different sizes of the parallel DATA_out/in word. The parallel data can have multiple of 8 bits, where these multiples can be either n = 1, 2, 4, or 8, thus allowing 'DATA_out/in' of sizes 8, 16, 32, or 64 bits. In this work we used always n = 4, which corresponds to 32-bit parallel data words.

The GTP tile internal PLLs transform the external reference clock into the high speed f_{line} frequency for LVDS bit-serial transmission. It also provides two reference clocks for user circuitry of frequency f and nf. When using 8b/10b encoding f_{line} = 10nf, and the 8n bits of 'DATA_out/in' and n bits of 'charisk' are read at the rising edges of clock f. For example, if using a 150MHz external reference clock and setting f_{line} = 1.5Gbps with n = 4, then f = 37.5MHz and 4f = 150MHz. Consequently, this would allow for a theoretical maximum event throughput of 37.5Meps (mega events per second), being each event of 32-bits. As we will see in Section IV, the maximum event throughput is slightly less, which is due to the need to periodically transmit commas for event alignment, clock correction and flow control.

III. Bit-Serial Bi-Directional AER Link

AER links between chips have traditionally been purely asynchronous [1], using bit-parallel ports together with two lines for 4-phase handshaking. A serial-word approach was also proposed [3], which would reduce the number of parallel bits by half. In the present approach we are using a clocked "GTP tile" that interfaces on one side through LVDS to another FPGA "GTP tile", and on the other side to in-FPGA clocked user circuitry. This clocked user circuitry may run at the same frequency f the GTP tile is reading (TX) DATA_out or providing (RX) DATA_in, or may run at a different frequency (either correlated to f and 4f, or totally uncorrelated). Ultimately, the FPGA user circuitry would interface to an external parallel AER device (such like an AER retina sensor [4], [5]) which is fully asynchronous.

Here we want to explore the case where the "GTP tile" has to connect, inside the FPGA, to a synchronous parallel AER circuit clocked at the maximum possible frequency. For this, we synthesized a "Test Pattern Generator" (TPG) which we were able to clock at 375MHz. This clock was generated from a second Xtal on the PCB, which is uncorrelated to the low-jitter differential one shown in Fig. 2. Fig. 3(a) illustrates the test approach by showing a simplified diagram of the circuitry connecting to the TX part of the "GTP tile" in Fig. 2. There is a high speed Test Pattern Generator clocked at f_{TPG} = 375MHz, which provides a bit-parallel DATA of 32-bits through handshaking (Req and Ack). These handshaking signals go through high-speed synchronizers, each clocked by either f_{TPG} or 4f, as shown in the figure. The TX FSM clocked at 4f interfaces with the "GTP tile" on one side, and on the other side interfaces through a clock domain crossing circuitry with the high speed TPG. Since "GTP tile" reads and writes data according to the rising edges of f, the TX FSM will provide a signal "phase" for the circuits running at frequency 4f to signal when the rising edge of f is approaching. The TX FSM and clock domain crossing circuitry are both clocked by the same clock 4f, and interchange information about their internal status through signals "valid" and "busy". The "clock domain crossing" circuit in this case is rather simple and is given Fig. 3(b).

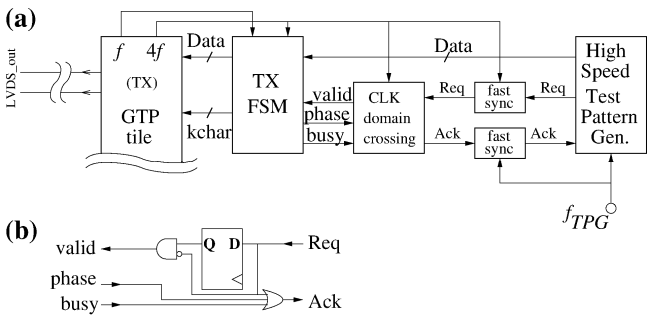


Fig. 3: (a) Interfacing circuitry between high speed clocked Test Pattern Generator and GTP tile. (b) Details of clock domain crossing circuit.

IV. Experimental Results

In this Section we provide initial experimental characterization results for the case of f_{line} = 1.5Gbps, 4f = 150MHz, and f = 37.5MHz.

A. Integrity of LVDS Signals in FPGA built-in GTPs

The quality of bit-serial transmission using GTP interfaces embedded in FPGAs depends on several factors. The primary source of errors is the frequency mismatch of the clocks involved and their jitter. Frequency mismatch (usually within a few tens ppms) can be successfully resolved by clock correction techniques based on elastic buffers and available within the GTP tiles [7], [8]. Physical properties of the link, such as quality of PCB traces, wires and connectors also impact signal integrity. A basic metric typically used for the evaluation of bit-

serial interfaces is BER (Bit Error Rate), which indicates the number of erroneous bits that have been transmitted during a test period - the lower the value the better the link. For this purpose Xilinx provides dedicated Test Pattern Generators based on 7-bit PRBS (Pseudo Random Bit Stream), which can be generated and verified inside a GTP transceiver if connected in loop-back mode.

A method frequently used to directly measure signal integrity is the Eye-Diagram analysis, which gives more detailed analog information on time and voltage margins for a particular link and its setup (pre-emphasis, equalization). It allows also determining the optimum sampling point to avoid violation of time-voltage margins embedded into an Eye-Diagram. This method can however have the drawback that the shape of the diagram can be affected by attaching the sampling probe, as this is an "invasive" method. For example, Fig. 4 shows measured eye-diagrams at two different sampling points.

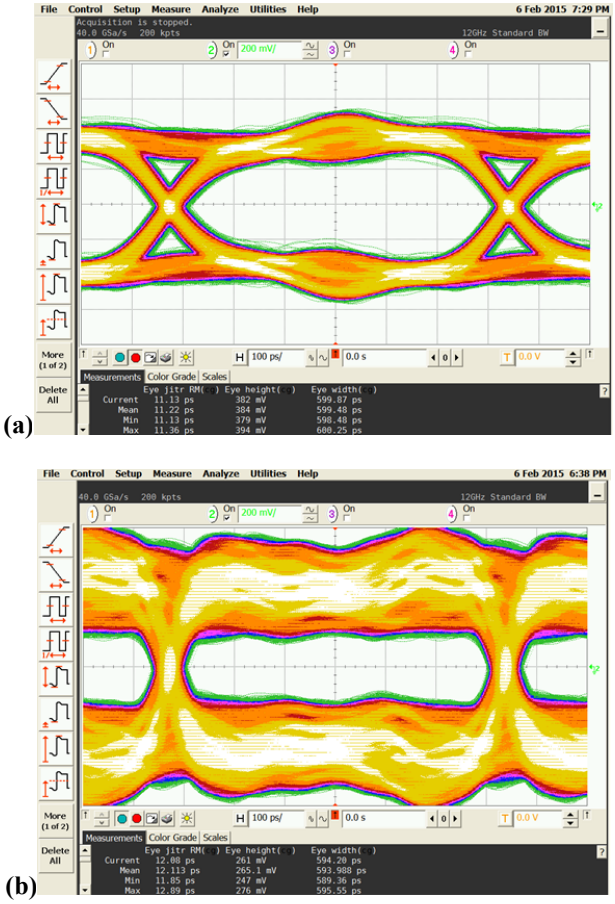


Fig. 4: Eye-Diagrams of LVDS link measured at different positions: (a) at FPGA RX input pins, (b) at inter-PCB SATA connector at TX side.

An alternative non-invasive technique for studying physical channel properties is the Bath-Tube analysis. A Bath-Tube curve can be acquired by measuring the BER value for a fixed period of time and sweeping a range of Sampling Point values. The result is a bath-shaped curve that usually yields very low values of BER (i.e. to 0.0) in the middle of the diagram, i.e. around 0.5UI (Unit Window). This corresponds to half of the period of the clock cycle, recovered from the received data bit-stream. The center of the region where the BER curve reaches 0.0 is the optimum Sampling Point for the setup of a particular loop-back link (see Fig. 5). Plots were collected with the ChipScope hardware debug interface and software plug-in iBERT [8] for Spartan6 devices.

Besides the Sampling Point, the following parameters also significantly impact transmission quality: Voltage Swing, TX-Preemphasis, and RX-Equalization. They should be adjusted to give minimum BER at a particular speed of the link. Fig. 5 shows some sample Bath-Tube diagrams for various RX-Equalization parameters. Fig. 6 shows the overlap of the bath graph with the eye diagram.

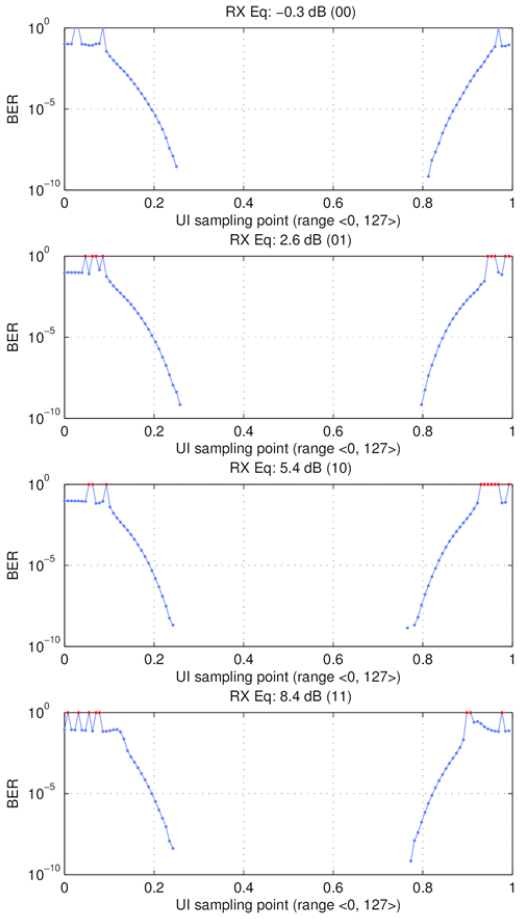


Fig. 5: Bath-Tube diagram for loop-back connection at 1.5Gbps, TX-Preemphasis 2.5dB, Voltage Swing 929mV

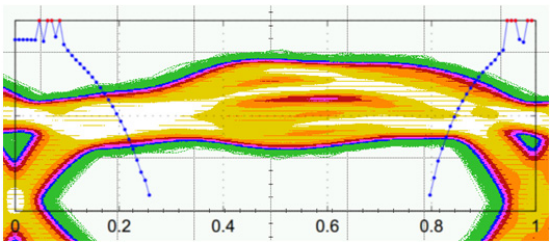


Fig. 6: Bath-Tube chart embedded into Unit-Window of the Eye-Diagram, for 1.5Gbps, TX-Preemphasis 2.5dB, voltage swing 929mV.

B. Event Error Rate (EER) Characterization

For our bit-serial inter-FPGA link it is interesting to characterize the "Event Error Rate" (EER), as this would

include all possible error sources, not only physical, but also logical, as the proper operation of the flow control protocol, the clock correction scheme, or the multi-byte event alignment [2]. In order to test EER we used the TPG discussed in Section II clocked at a very high clock frequency of f_{TGP} = 375MHz. The test setup involving two AER-Node PCBs (NB1 and NB2) is shown in Fig. 7. The parallel-serial transceiver link in each FPGA includes the "GTP tile", the "TX block" (clocked at 4*f*, the faster user clock provided by "GTP tile") and an "RX block" (clocked at *f*, the slower user clock provided by "GTP tile"). Besides the TPG, there are also two "Synchronous Test Pattern Checkers" (STPC) clocked at frequency *f*. The STPC at the end of the chain in each FPGA tests EER for the whole transmission chain, while the STPC between "GTP tile" and "TX block" tests clock domain crossing issues.

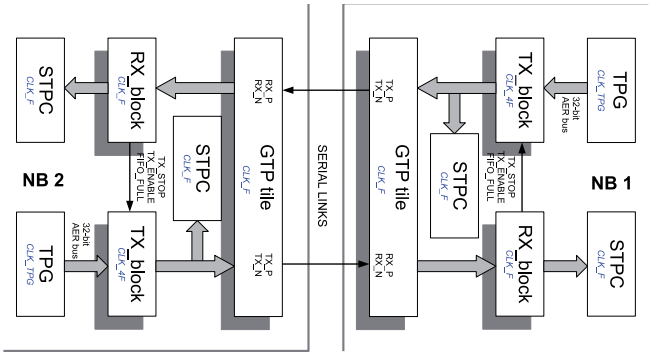


Fig. 7: Test setup consisting of two AER-Node PCBs (NB1 and NB2), each with one bit-serial link ("GTP tile", "TX block" and "RX block") together with a "Test Pattern Generator" (TPG) clocked at very high frequency f_{TPG} = 375MHz and two "Synchronous Test Pattern Checkers" (STPC) clocked at the slower clock provided by "GTP tile" *f* = 37.5MHz.

Table I: Measured Errors, 32-bit Event Throughput and skew between pairs of AER-Node Boards on all 4 links

	Internal STPC errors	EER	RX Measured Throughput (eps)	Clk skew (ppm)
NB2 → NB1	0	0.0	37426439	-16.134
NB1 → NB2	0	0.0	37427647	16.133
NB4 → NB3	0	0.0	37428143	-29.409
NB3 → NB4	0	0.0	37425942	29.408

Long term tests of serial connections at 1.5Gbps were performed concurrently between two AER-Node PCBs for each all 4 serial links, as connected in Fig. 2(b), during 20 hours or over. No errors were detected at none of the STPCs, as shown in Table I: columns "Internal STPC errors" and "EER" (which corresponds to the errors counted at the STPC at the end of the paths). The throughput measured at the internal STPCs was always 37,427,043 eps, which corresponds exactly to the expected TX side generated effective data rate (excluding the COMMA density we were using). However, the measured throughput varies slightly on the path ends STPCs (RX side), which are the ones shown in Table I under "RX Measured Throughput", because these rate numbers are measured using the clocks at the destination AER-Node board whose clock differs by a few tens of ppms. The measured difference between the TX and RX throughputs is given under column "Clk skew" in ppms.

CONCLUSIONS

We have shown correct operation of bit-serial AER links between custom made PCBs for multi-tile PCB assembly interconnected through 4 SATA links. Initial experiments operating at 1.5Gbps are reported. The links are stimulated through asynchronous handshaking techniques inside each FPGA by Test Pattern Generators clocked by uncorrelated clocks. These uncorrelated clocks were set at the maximum possible frequency we were able to set, 375MHz, in order to test the channel at its maximum possible throughput. Exhaustive tests were performed to verify the correct operation, including eye diagram measurements, bath-tube measurements, bit error rates, and event error rates.

ACKNOWLEDGMENTS

This work has been supported by Spanish grants (with support from the European Regional Development Fund) BIOSENSE (TEC2012-37868-C04-02/01), Andalusian grant NANO-NEURO (TIC-6091), Polish grant from the Ministry of Science and Higher Education of the Republic of Poland AGH UST no. 11.11.120.612, EU CHIST-ERA grant PNEUMA (PRI-PIMCHI-2011-0768), EU FP7 grant HBP (The Human Brain Project) 604102, and H2020 EU grant ECOMODE 644096.

REFERENCES

[1] M. Sivilotti, "Wiring considerations in analog VLSI systems with application to field-programmable networks," *Ph.D. Dissertation*, Computation and Neural Systems, California Inst. Technol., Pasadena, CA, 1991.

[2] T. Iakymchuk, A. Rosado, T. Serrano-Gotarredona, B. Linares-Barranco, A. Jimenez-Fernandez, A. Linares-Barranco, and G. Jimenez-Moreno, "An AER handshake-less modular infrastructure PCB with 8 2.5Gbps LVDS serial links." *Proc. of the IEEE Int. Symp. Circ. and Syst.* (ISCAS), pp. 1556–1559, June 2014.

[3] K. Boahen, "Point-to-Point Connectivity Between Neuromorphic Chips Using Address Events," *IEEE Trans. on Circ. and Syst. Part-II*, vol. 47, no. 5, pp. 416-434, May 2000.

[4] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128x128 120dB 15us latency asynchronous temporal contrast vision sensor," *IEEE J. Solid State Circuits*, 43(2) 566-576, 2007.

[5] T. Serrano-Gotarredona, B. Linares-Barranco, "A 128x128 1.5% Contrast Sensitivity 0.9% FPN 3µs Latency 4mW Asynchronous Frame-Free Dynamic Vision Sensor Using Transimpedance Preamplifiers," *IEEE Journal of Solid-State Circuits*, vol. 48, No. 3, pp. 827 – 838, 2013.

[6] R. Serrano-Gotarredona, et al. "CAVIAR: A 45k-Neuron, 5M-Synapse, 12G-connects/sec AER Hardware Sensory-Processing-Learning-Actuating System for High Speed Visual Object Recognition and Tracking", *IEEE Trans. on Neural Networks*, vol. 20, No. 9, pp. 1417-1438, Sep. 2009.

[7] DS160. Spartan-6 Family Overview, v.2.0. Xilinx, October 2011.

[8] UG386 Spartan-6 FPGA GTP Transcievers. Advance Product Specification, v.2.2. Xilinx, April 2010.

[9] C. Zamarreño-Ramos, A. Linares-Barranco, T. Serrano-Gotarredona, B. Linares-Barranco, "Multicasting Mesh AER: A Scalable Assembly Approach for Reconfigurable NeuromorphicStructured AER Systems. Application to ConvNets", *IEEE Trans. on Biomedical Circ. and Syst.*, vol. 7, No. 1, pp 82-102. Feb. 2013.

[10] Daniel B. Fasnacht, Adrian M. Whatley, Giacomo Indiveri. "A Serial Communication Infrastructure for Multi-Chip Address Event Systems". *Circuits and Systems (ISCAS), 2008 IEEE International Symposium on*.

[11] P. A. Franaszek, et al., "Byte oriented DC balanced (0,4) 8b/10b partinioned block transmission code," US Patent 4,486,739, Dec. 4, 1984.