

Embedding deserialisation of LHC experimental data inside Field Programmable Gate Arrays.

Ph. Busson, L.Dobrzynski, A. Karar, T. Romanteau for LLR, P.Moreira for CERN, J.L. Brelet, J.R. Macé for XILINX France, M. Défossez for XILINX Benelux, M. Crowley, C Gannon for XILINX DESIGN SERVICES (XDS)

LLR, 91128 Palaiseau, France – CERN 1211 Geneva 23, Switzerland – XILINX France, 78353 Jouy-en-Josas, France – XILINX Benelux, 2400 Antwerpen, Belgium – XDS, Dublin, Ireland
dobrzynski@cern.ch, romanteau@poly.in2p3.fr

Abstract

LHC experiments will make use of thousands of serial links in order to transfer digital data from the electronics sitting in the detectors to the off detector electronics located more than 100 meters away. Due to the high levels of radiation present in the detectors, CERN designed and developed a radiation hard serialiser, the Gigabit Optical Link (GOL) chip[1]. On the other hand, the off-detector electronics designed to process digital data received from the detectors will heavily rely on commercial programmable components like Field Programmable Gate Arrays.

Commercial components will be used to de-serialize the detector's data prior to processing with FPGAs. The Xilinx Company is now offering a new type of FPGAs (Virtex2Pro) which embed Multi Gigabit Transceivers (MGT[2]). The use of these components will allow a more powerful and compact design of the off-detector electronics processing boards. This paper will describe the results of tests performed to measure the performance of a link made with a GOL chip and a Virtex2Pro circuit.

I. INTRODUCTION

In High Energy Physics (HEP) experiments, high-speed links operating in the Gbit/s range are chosen for data transmission of both raw and specific trigger data. Low latency is required for the last one, to minimize the data storage inside the detector until the trigger decision is available. For the transmission of raw data a low latency is not required. It appears that a large reduction in number, size and cost of the data link boards can be achieved if a high integrated solution is possible for the deserialiser. The new Virtex2Po FPGA device from XILINX seems to be a good candidate for this application.

At a 1st stage, we performed a performance study where the GOL was connected to the commercial reference deserialiser TLKx501 device from Texas Instrument. These results will be used as a reference in this paper, but the details will not be reported

The present paper describes the work we performed in collaboration with XDS¹ to check the functionality of the GOL when connected to the MGT of the Virtex2Pro.

II. BASIC REQUIREMENTS

The evaluation focus on three measurements made on the communication link:

- **Bit-error rate:** While the MGT is synchronised to the transmitted signal, the number of bit- errors are counted and recorded periodically.
- **Loss of Synchronisation:** Under certain conditions, the Rocket IO may lose synchronisation with the signal transmitted by the GOL device. The number of times that the MGT loses synchronisation has been recorded. Statistics have been also collected regarding the resynchronisation time.
- **Link latency:** The Rocket IO receiver latency is measured.

All these measurements have been carried out in slow (800Mbits/s) and fast (1.6Gbits/s) modes.

The jitter of the clock provided to the MGT have been measured and the eye diagrams of the serial transmitted data observed.

A test environment, described in the next section, has been implemented for these measurements

III. TEST ENVIRONMENT

A. Overview

Figure 1 shows the test environment. The GOL board is used to transmit a known data sequence to the V2PRO board. The data sequence is received by the FPGA device on the V2PRO board. The FPGA device contains the MGT transceivers used to de-serialise the received stream. In addition, the FPGA contains dedicated logic and a PowerPC processor that are responsible for analysing the received data and making the measurements. The PowerPC is also responsible for offloading the measurements via an

¹ Xilinx Design Services based in Ireland

RS232 interface whereby they can be displayed on a hyper-terminal residing on a PC.

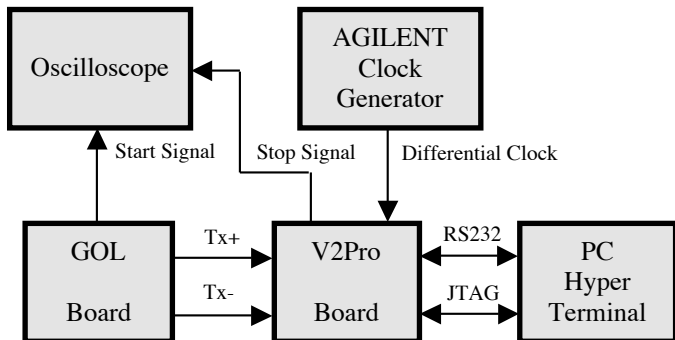


Figure 1: Test Environment setup

B. V2Pro Board

Xilinx has developed a dedicated high quality MGT characterisation board (ML320 Board). This board is a state of the art and provide the best possible test platform for our project. XDS have been granted an early access to it.

ML320 Board

The verified designs have been ported to the ML320 in order to run the final measurement tests. In addition to the V2PRO device, the board hosts also a UART driver that is used to support all test configuration and status functions.

A single differential clock running at 40MHz (slow mode) or 80MHz (fast mode) clocks the board and can be configured using ACE System.

C. GOL Board

The GOL board contains the GOL device and an Altera FPGA. It has the pattern generator providing the data sequence to be transmitted. The Altera device is programmed from a PROM. Several switches are used to change the operation modes of the board e.g. to switch between slow and fast mode.

D. Clocking

The V2PRO board requires a single differential clock. This clock will be 40MHz for slow and 80MHz for fast mode.

An Agilent data generator has been used to provide the two required clocks. The GOL board has its own on board clock.

These two clocks were not frequency or phase locked to each other.

E. Oscilloscope

A high performance oscilloscope has been used for probing the gigabit serial signals in order to help in any troubleshooting. This oscilloscope was used for the latency measurement between the GOL and the MGT, see section IV.B.

F. V2Pro FPGA

Figure 2 illustrates the implemented frame work inside the V2PRO FPGA. The design contains two MGTs. The MGT1² is configured to receive serial Ethernet data at 800 Mbits/s and MGT2³ is configured to receive serial Ethernet data at 1.6 Gbits/s. The SD+/SD- ports are connected to a pair of SMA connectors on the board and the FD+/FD- are connected to another pair of SMA connectors. The GOL serial data signals should be routed to one of these pairs depending on whether the slow or fast mode is chosen.

The MGT performs the serial to parallel conversion on the serial inputs and passes the data and its own status information to the measurement block. The measurement block performs the necessary measurements based on an analysis of the received data and the MGT status. It passes interim measurement data to the processor periodically. The processor compiles and processes the measurements and presents summary results on the RS232 UART interface whereby the measured data can be displayed. The processor program and data memory will reside in FPGA block RAM.

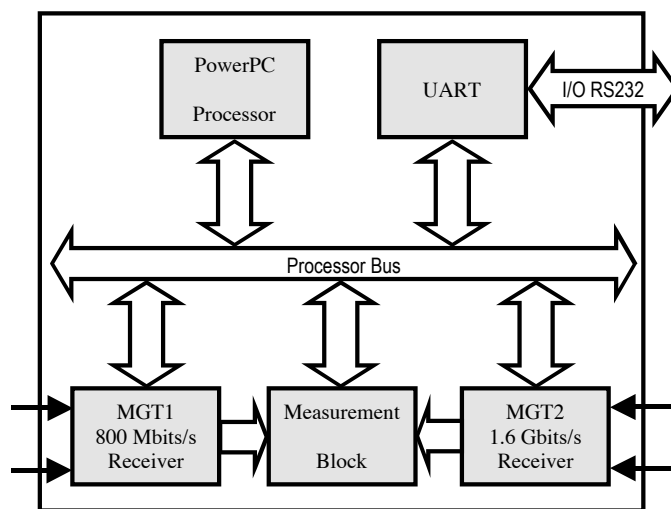


Figure 2: V2Pro Design Environment

The UART interface is bi-directional and it is conceived such that the user can enter commands on the hyper-terminal so that measurement tests can be flexibly and quickly configured and

² MGT identification number on LM320 board

³ MGT identification number on LM320 board

controlled. These commands would be interpreted and executed by the PowerPC by writing to the measurement block registers.

G. PC / Hyper-terminal

The PC fulfils two functions. Firstly, it is used to download the V2PRO bit stream to the board. Secondly, it hosts a hyper-terminal from where the measurement tests will be controlled, configured and on which the measurement results will be displayed.

IV. MEASUREMENT AND RESULTS

A. Test Set-Ups

The tests described here after, rely on two specific set-ups:

GOL Connection Set-up

In this connection set-up, the communication link connects the GOL and the ML320 boards. The transmitted data rate is either 800 or 1600 Mbits/s depending on the BITS16 switch sitting on the Gol board. If the data rate is 800 Mbits/s then the transmitted data should be connected to the RX connectors for MGT1, otherwise MGT2 RX ports should be used.

This connection set-up allows bit error rate testing and latency measurement.

Loopback Connection Set-up

In this connection set-up, the GOL board is not required. In 800 Mbits/s mode, the TX ports of MGT9⁴ are connected to the RX ports of MGT9. In 1600 Mbits/s mode, MGT4⁵ ports are connected in the same way.

Each Rocket IO transmitter is driven by a pattern generator implemented in the FPGA logic, the same as the one used in the GOL board Altera device with the exception that some parts of the pattern contain (intentionally) wrong data words.

In this loopback configuration, the intention is that the same Rocket IO that transmits it, receives the transmitted data. This connection set-up is used for two reasons:

- It facilitates testing of the bit error measurement logic and software.
- It facilitates the measurement of resynchronisation time as will be described below.

B. Bit-Error Rate Tests

Basic Measurement Mechanism

A bit error counter (BEC) is implemented in the FPGA fabric. The receiver has the same pattern sequencer as the

one used at the transmitter. When the MGT is synchronised and the receiver's pattern sequencer is also locked to the received data sequence, then the BEC is incremented by X, where X is the number of bit errors within a word. X is obtained by comparing the output word of the receiver's pattern sequencer with the received data word from the MGT. The word width is 16 bit for slow mode and 32 bit for fast mode.

The GOL device remains in lock but consecutive sequences "S" are separated by one or more IDLE words. The bit error calculation is made over multiple sequences for a pre-specified time interval. The bit errors are only counted if they occur in a predefined range.

To specify a time interval over which BER measurements are made, there is a hardware timer implemented in the FPGA fabric. This timer is called the BEC Interval Timer. The timer creates a pulse signal periodically at a rate of one pulse per second, for example. This pulse causes the current result of the BEC to be stored in a memory-mapped register and a dedicated timer interrupt to be sent to the processor.

In response to the interrupt, the processor reads the BEC value from the memory-mapped register and streams the value to the UART so that it is displayed on the hyper-terminal.

It may be possible to make the BEC Interval Time programmable from the hyper-terminal. In addition, software could be used to compute an (running) average bit error count over multiple BEC intervals.

Pattern Loss of Lock Conditions

While the errors are being counted for a particular transmitted sequence, it may happen that errors are detected in each subsequent word or it may happen that the RXNOTINTABLE status bits from the MGT are active on many consecutive bytes.

If the MGT is synchronised (as indicated by RXLOSSOFSYNC), we have detected one of the following conditions:

- a framing error in the MGT where the byte alignment has been lost (indicated by consistent errors on RXNOTINTABLE)
- the pattern detector is not properly synchronised with the transmitted sequence, indicated by consistent mismatch between the receiver's pattern sequence and the received pattern sequence.

In either case, the receiver's pattern sequence is halted and bit errors are no longer recorded. The receiver's pattern detector must wait for the start of the next transmitted sequence before relocking and continuing the error counting process. The loss of lock condition is detected when 8 consecutive bytes are received in error where a byte error is defined by mismatch or RXNOTINTABLE.

If there are many occurrences of this loss of lock condition during a BEC measurement interval then the BEC measurement may be considered invalid. For this reason,

⁴ MGT identification number on LM320 board

⁵ MGT identification number on LM320 board

the number pattern detector loss of lock conditions is also counted by the hardware and the PowerPC may read this count at the end of the measurement interval.

Results

The GOL connection set-up is used to connect the GOL data lines to either MGT2 (1600Mbits/s) or MGT1 (800 Mbits/s). The bit error rate test was conducted for both data-rates. The results are reported in table:

Table 1: Bit Error Test Results

Data Rate	Test Duration (hours)	Bit Error Rate	Synchronisation Losses
800Mbits/s	17 : 00 ⁶	0	0
1600Mbits/s	67 : 00 ⁷	0	0

Bit error rate is computed by the equation $BER = B/(D*T)$ using the following quantities:

- The time elapsed T.
- The uncoded data rate D i.e. 640Mbits/s for slow mode and 1280Mbits/s for fast mode
- The number of bit errors encountered during the time elapsed i.e. B.

Analysis

The bit error rate testing has shown that a perfect communications channel can be established and maintained over very long periods of time. The good results are likely to be due to:

- The high quality transmission signals coming from the GOL device.
- The high quality of the ML320 board design in terms of power supply network and the track layout.
- The high quality of the REFCLK clock source on the ML320 board.

The only factor that will have to be significantly different in the real application will be the transmission quality. In the real application, the transmission signals will exhibit considerable levels of jitter. At this time, it was not possible to vary the jitter of the transmitted signals but this is definitely a future possibility.

C. Link Latency Tests

The MGT receiver latency is of interest. However, there is not easy to measure just the receiver latency. It was then measured the end-to-end latency: L_TOTAL. The

⁶ Equivalent to a BER limit of $6.6 \cdot 10^{-14}$ at 95% of CL

⁷ Equivalent to a BER limit of $8 \cdot 10^{-15}$ at 95% of CL

MGT receiver latency L_RX can then be obtained by subtracting the transmitter latency L_TX and the cable latency L_CABLE from L_TOTAL:

$$L_{RX} = L_{TOTAL} - L_{TX} - L_{CABLE}$$

L_TX is obtained from reference[1], L_CABLE is given by the cable length and L_TOTAL is obtained by the following way:

When the Altera device provide the value FFFF to the GOL device for transmission it can produce a TX_START signal, then when the receiver FPGA device detects this FFFF, it can produce a RX_START signal. The TX_START and RX_START signals are then probed by an oscilloscope and the time between them can be measured to give L_TOTAL. Repeated manual measurements have been made to determine the latency and study its variations.

Alternatively, it may be possible to route the TX_START signal from the GOL board to the V2PRO board and the FPGA can then be used to make repeated automatic measurements.

Results

To execute the measurement, the GOL connection set-up was used. An oscilloscope was used to probe the TX_START signal from the GOL board and the RX_START signal from the ML320 board. The time difference between the rising edge of TX_START and RX_START has been measured

The results are quoted in Table 2. The time difference between TX_START and RX_START was found to be variable but bounded below and above. The table quotes the minimum and maximum time difference observed.

Table 2: Latency Results

Data Rate	Minimum (ns)	Maximum (ns)
800Mbits/s	732	763
1600Mbits/s	414.4	434.2

A detailed analysis of these results is given in reference[4].

For comparison table 3 gives the GOL device latency [1]. The measured latency for the GOL+TLK is about 5 clock cycles.

Table 3: GOL Device Latency Figures

Mode	Minimum (ns)	Maximum (ns)
Slow	68	78
Fast	54	64

D. Resynchronisation Time Tests

The MGT indicates that synchronisation has been lost if there is a non-zero value on the RXLOSSOFFSYNC[1:0]

port. When this MGT port changes from zero to nonzero, a hardware timer is commenced. The timer stops when the MGT port returns to zero. The timer is a counter implemented in the FPGA fabric. It runs of the 40MHz clock and it is 32 bit wide.

When the timer stops, an interrupt goes to the processor. The processor reads the value of the timer and resets it. The value read is the recovery time of the MGT and this is displayed on the hyper-terminal.

In order to obtain this measurement, a mechanism is required so that the MGT is forced into a desynchronised state. Some ideas are as follows:

- The processor will write to a register and this register will cause the MGT REFCLK to be disabled. When the RXLOSSOF SYNC port changes to non-zero, the REFCLK is re-enabled so that resynchronisation becomes possible again.
- The Agilent clock source frequency can be adjusted until the MGT is forced into a desynchronised state. The problem with this is that the user adjusting the frequency does not know when exactly to restore the clock source to the nominal frequency in order to allow resynchronisation to occur.

Because this is an unusual application, it was necessary to consult internal Xilinx MGT experts for advice on this (Sean Koontz).

1) Description

The loopback connection set-up is used for this test. The transmitter data lines of the Rocket IO are routed back into the receiver of the same Rocket IO. From the hyper-terminal user interface, the user can force the transmitter to inhibit the transmitter, causing the transmission line to be held in a fixed state for a short period of time. This causes the receiver to lose synchronisation. As soon as the transmitter becomes uninhibited, a timer initialises and starts counting. The timer increments until the receiver resynchronises. When this occurs, the timer stops counting and the value of the timer, i.e. the resynchronisation time is reported to the hyper-terminal. This process can be repeated ad infinitum and different results should be obtained each time. The maximum resynchronisation time is bounded by the interval between transmitted comma characters.

2) Results

A number⁸ of measurements were made over time in both fast and slow mode. For each measurement, the receiver was made to lose synchronisation and the resynchronisation timer value was reported to the hyper-terminal. A different result was obtained for each measurement. In addition, each result was in the range [321,30771] where the units are in 40MHz clock cycles.

3) Analysis

The results were as expected. The resynchronisation time is dependent on where in the transmitted frame the desynchronisation occurs. Thus, if the point of desynchronisation is random, then the resynchronisation time is random. In addition, since resynchronisation must wait for a COMMA character (i.e. the start of the next frame) to arrive at the receiver, then the resynchronisation time must be bounded by the interval between COMMA characters. The COMMA character interval (frame size) is 32777 clock cycles for both slow and fast mode. Thus, it would be expected that the results be bounded in the range [0,32777].

V. CONCLUSION

Integration of deserialisers in FPGA devices is a key factor to reduce the global cost of the off-detector electronic for LHC experiments. The fast development of industrial standard chips (FPGA) for the telecom market is a good opportunity to try to use them for HEP applications. The first results of the projects reported in this paper demonstrate that the GOL-MGT is a good combination for many LHC experiment requirements. Our future work in this domain will be to find specific solutions, always based on the use of FPGA fabrics and their embedded cores. Moreover, we think that it will be useful to perform a precise characterisation of the clock jitter and the integrity of the Giga-bit data link.

VI. REFERENCES

- [1] GOL Reference Manual, v1.2, May 2002, CERN Microelectronics Group
- [2] Rocket I/O Transceiver User Guide, UG024, v1.3, June 2002, Xilinx
- [3] Requirements Specification (rs01-2250-0001), v1.3, Xilinx Design Services
- [4] Test Report (tr01-2250-0001), v1.3, Xilinx Design Services

⁸ Approximately 50 measurements were made on each mode.