

# 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](mailto:dobrzynski@cern.ch), [romanteau@poly.in2p3.fr](mailto: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 the Gigabit Optical Link (GOL) chip[1], a radiation hard serialiser. On the other hand, the off-detector electronics currently designed for processing digital data as 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 using FPGAs. The Xilinx Company is now offering a new type of FPGAs (Virtex2Pro) which embed Multi Gigabit Transceiver (MGT[2]). The use of such components will allow more powerful and compact design of the off-detector electronics processing boards. This paper will describe the results of tests performed for measuring the performance of a link made with a GOL chip and a Virtex2Pro circuit.

## I. INTRODUCTION

Today in High Energy Physics (HEP) experiments, high-speed links operating in the Gbit/s range were chosen for data transmission of both raw and trigger data. Low latency is required for trigger data links to minimize the amount of memory needed inside the detector and the overhead time between two trigger events. For the transmission of raw data this low latency is not required. It appears that a large reduction in number, size and cost for the data link boards can be achieved if a highly integration solution is possible for the deserialiser. The new Virtex2Po FPGA device from XILINX seems to be a good candidate for this application.

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

Present paper will describe the work we performed in collaboration with XDS<sup>1</sup> to check the functionalities of the GOL when connected to the MGT of the Virtex2Pro.

## II. BASIC REQUIREMENTS

The evaluation will focus on three measurements that will need to be made on the communications link. These measurements are:

- **Bit-error rate.** While the MGT is synchronised to the transmitted signal, the number of bit- errors will be counted and recorded periodically.
- **Loss of Synchronisation.** Under certain conditions, the Rocket IO may lose synchronisation to the signal transmitted by the CERN GOL device. The number of times that the MGT loses synchronisation will be recorded and statistics will be collected regarding the time it takes to resynchronise. It is desired to know whether the recovery time is constant or variable.
- **Link latency.** Ideally, the receiver latency of the Rocket IO is required. Again, it is desired to know whether the latency is constant or variable.

These measurements will be carried out in slow (800Mbits/s) and fast (1.6Gbits/s) mode and a test report will be produced.

If possible<sup>2</sup> within the duration of the project, the jitter on the clock provided to the MGT should be measured and eye diagrams of the serial transmitted data should be provided in the final test report.

A test environment will be set up to allow these measurements to be made. The next section describes the test environment.

---

<sup>1</sup> Xilinx Design Services based in Ireland

<sup>2</sup> Special oscilloscopes may be required to measure clock jitter and eye diagrams.

### III. TEST ENVIRONMENT

#### A. Overview

Figure 1 shows the test environment setup with all the boards and equipment required.

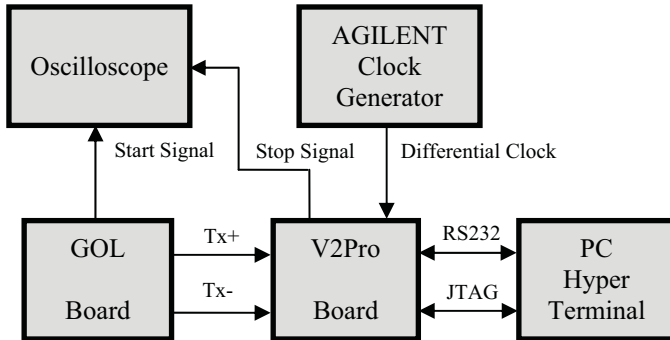


Figure 1: Test Environment

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 which together are responsible for analysing the received data and making the necessary measurements. The POWER PC is also responsible for offloading the measurements via an RS232 interface whereby they can be displayed on a hyper-terminal residing on a PC.

#### B. V2Pro Boards

There are two boards that can host the V2 Pro device to be used in the tests. Since the original definition of this project, Xilinx have developed a dedicated high quality MGT characterisation board. This board is state of the art and will provide the best possible test platform. XDS have been granted early access to this board especially for this project. The ML320 board is due at the start of August 2002. In the meantime, the AFX board will be used.

##### *ML320 Board*

The verified designs will be ported to the ML320 in order to run the final measurement tests.

As well as the V2PRO device, the board hosts a UART driver that will be used to support all test configuration and status functions.

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

#### C. GOL Board

The GOL board contains the GOL device and an Altera FPGA design that contains the pattern generator providing the data sequence that is ultimately transmitted. The Altera

device is programmed from a PROM. Several switches are provided to change the operational 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.

The GOL board has its own on board oscillator and requires no clocking input.

An Agilent data generator will be used to provide the two clocks required. These clocks will not be frequency or phase locked to each other.

#### E. Oscilloscope

A high performance oscilloscope will be available for probing the gigabit serial signals in order to aid in any troubleshooting that is required.

In addition, this oscilloscope will be used to allow measurement of latency between the GOL and the MGT, see section IV.B.

#### F. V2Pro FPGA

Figure 2 illustrates the design that will reside inside the V2PRO FPGA. The design contains two MGTs. MGT<sup>3</sup> is configured to receive serial Ethernet data at 800 Mbits/s and MGT<sup>4</sup> 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 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 onto 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.

The UART interface is bi-directional and it is intended that the user can enter commands on the hyper-terminal so that measurement tests can be flexibly and quickly configured and controlled. These commands would be interpreted and executed by the PowerPC by writing to the measurement block registers.

<sup>3</sup> MGT identification number on LM320 board

<sup>4</sup> MGT identification number on LM320 board

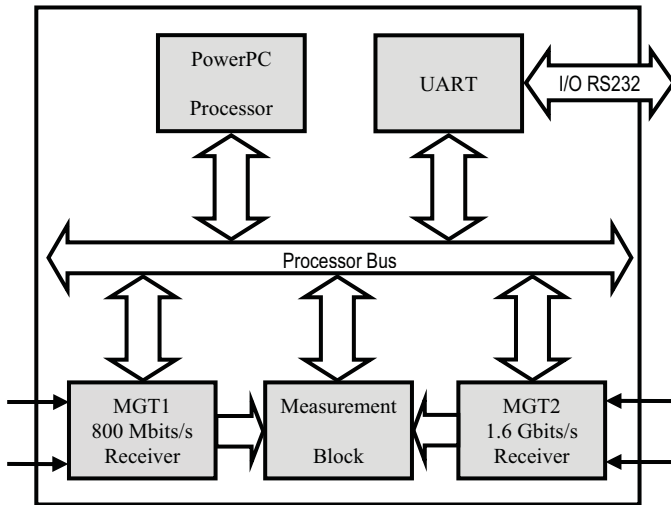


Figure 2: V2Pro Design Environment

### 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 whereby the measurement tests will be controlled and configured and the measurement results will be displayed.

## IV. DETAILED MEASUREMENT AND RESULTS

### A. Test Set-Ups

The tests documented in the following sections rely on two specific set-ups.

#### 1) GOL Connection Set-up

In this connection set-up, the communications link is between the GOL board and the ML320 board. The transmitted data rate can be 800 or 1600 Mbits/s depending on the BITS16 switch. If the data rate is 800 Mbits/s then the transmitted data should be connected to the RX connectors for MGT9, otherwise MGT4 RX ports should be used.

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

#### 2) Loopback Connection Set-up

In this connection set-up, the GOL board is not required. In 800 Mbits/s mode, the TX ports of MGT9<sup>5</sup> are connected to the RX ports of MGT9. In 1600 Mbits/s mode, MGT4<sup>6</sup> ports are connected in the same way.

Each Rocket IO transmitter is driven by a pattern generator implemented in the FPGA logic. This is the same pattern generator that is used in the Altera device in the GOL board

with the exception that some parts of the pattern are errored.

In this loopback configuration, the intention is that the same Rocket IO that transmitted 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

#### 1) Basic Measurement Mechanism

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

GOL device remains in lock but consecutive sequences are separated by one or more IDLE words. Bit error calculation is made over multiple sequences for a pre-specified time interval. The bit errors are only counted if they occur in the range (S(2),S(65535)). The operation is as follows.

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 RXLOSSOF SYNC), we have detected one of the following conditions:

<sup>5</sup> MGT identification number on LM320 board

<sup>6</sup> MGT identification number on LM320 board

- a framing error in the MGT whereby 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 RXNOTINTTABLE.

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

#### Description

The purpose of this test is to measure the bit error rate. Note, as described in [3], bit errors are only accounted for when they occur inside the framed data sequence.

The GOL connection set-up is used connecting the GOL data lines to either MGT4 (1600Mbits/s) or MGT9 (800 Mbits/s).

#### Results

The bit error rate test was conducted for both data-rates. The results are reported in the following table.

Table 1: Bit Error Test Results

Data Rate	Test Duration	Bit Error Rate	Synchronisation Losses
800Mbits/s	17 : 00 <sup>7</sup>	0	0
1600Mbits/s	67 : 00	0	0

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

- The time elapsed T.
- The unencoded 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. It is desired to know its value and if it is constant or variable. However, there are difficulties with measuring just the receiver latency and it is more feasible to measure end-to-end latency  $L_{TOTAL}$ . The MGT receiver latency  $L_{RX}$  can be obtained by subtracting the transmitter latency  $L_{TX}$  and the cable latency  $L_{CABLE}$  from  $L_{TOTAL}$ . That is,  $L_{RX} = L_{TOTAL} - L_{TX} - L_{CABLE}$ . The figures for  $L_{TX}$  can be obtained from [1]. The  $L_{CABLE}$  figure can be derived from the cable length i.e. delay/metre for standard cable is 5ns/metre.

$L_{TOTAL}$  can be measured by the following means. The Altera device will pulse a TX\_START signal when it is providing the value FFFF to the GOL device for transmission. At the receiver, the FPGA device will detect the FFFF value on the received data provided by the MGT. When this occurs, the FPGA device will pulse a RX\_START signal. The TX\_START and RX\_START signals will be probed by an oscilloscope and the time between them can be measured to give  $L_{TOTAL}$ . Repeated manual measurements will be made in this fashion to determine whether the latency is constant or variable.

Alternatively, it may be possible to route the TX\_START signal from the GOL board to the V2PRO board and the FPGA can be used to make repeated automatic measurements. For this option, it may be necessary to have a very fast clock in order to get good time resolution.

#### 1) Description

The GOL connection set-up is used. In addition, an oscilloscope is used to probe the TX\_START signal on the GOL board. On a second oscilloscope channel, the RX\_START signal from the ML320 board is observed. The time difference between the rising edge of TX\_START and RX\_START is measured.

<sup>7</sup> It is acknowledged that a longer test needs to be run.

## 2) Results

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

## 3) Analysis

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

The other information that we have concerns the GOL device latency. This is specified in GOL Reference Manual and is re-conveyed in Table 4.

Table 4: 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 RXLOSSOF SYNC[1:0] port. When this MGT port changes from zero to non\_zero, 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 off 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 is 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 number<sup>8</sup> 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

The greatest 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 our applications. The first results of the projects reported in this paper demonstrate the GOL-MGT is good combination for many LHC experiment requirements. Our future work in this domain will be to find specific solution, always based on the use of FPGA fabric and its embedded cores.

<sup>8</sup> Approximately 50 measurements were made on each mode.

Moreover, we think that it will be useful to perform a precise characterisation of clock jitter and the integrity of 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