Passive Optical Networks for the Distribution of Timed Signals in Particle Physics Experiments

I. Papakonstantinou[a], C. Soos[a], S. Papadopoulos[a], J. Troska[a], F. Vasey[a], S. Baron[a], L. Santos[a], S. Silva[a], P. Stejskal[a], C. Sigaud[a], S. Detraz[a], P. Moreira[a], I. Darwazeh[b]

[a] CERN, Div. PH-ESE, 1211 Geneva 23, Switzerland
[b] University College London, Department of Electronics Engineering, Torrington Place, London WC1E 7JE, UK

jan.troska@cern.ch

Abstract

A passive optical network for timing distribution applications based on FPGAs has been successfully demonstrated. Deterministic latency was achieved in the critical downstream direction where triggers are distributed while a burst mode receiver was successfully implemented in the upstream direction. Finally, a simple and efficient protocol was introduced for the communication between the OLT and the ONUs in the network that maximizes bandwidth utilization.

I. INTRODUCTION

Optical links are deployed in a number of applications currently in the Large Hadron Collider (LHC) where both point-to-point (P2P) and point-to-multipoint (P2MP) topologies are exploited for data collection, timing distribution and control and management signal transmission. P2P links are mainly used in data read out systems, as the inherent bandwidth sharing property of the P2MP links makes them inadequate to be used in such applications. However, P2MP links are seen to offer advantages in cases when signals have to be broadcasted simultaneously to a number of destinations. This is the case for the Timing-Trigger and Control (TTC) system, [1] the part of which we are interested in is shown in Fig. 1, where triggers and commands are distributed downstream from the TTCex to a number of TTCrxs. Typically two variations of the TTC system are met depending on whether the TTCrxs are installed inside the detector or in the counting room, Fig. 1. Optical links are unicast in both cases and information is flowing only in the downstream direction from the TTCex to the TTCrxs. A separate "busy" electrical data link is used in order for the TTCrxs to communicate their status back to the TTCex but the "busy" link is usually slow to respond and it would be beneficial if the communication took place in real time. The objective of this work is to design a bidirectional optical link based on the commercial Passive Optical Network (PON) architecture to combine both downstream and upstream data in the same fibre while at the same time being able to meet the stringent latency and jitter requirements of the bespoke optical networks used in particle physics experiments.

II. PASSIVE OPTICAL NETWORKS

Passive Optical Networks (PONs) are Point-to-MultiPoint optical networks with no active elements in the signal's path
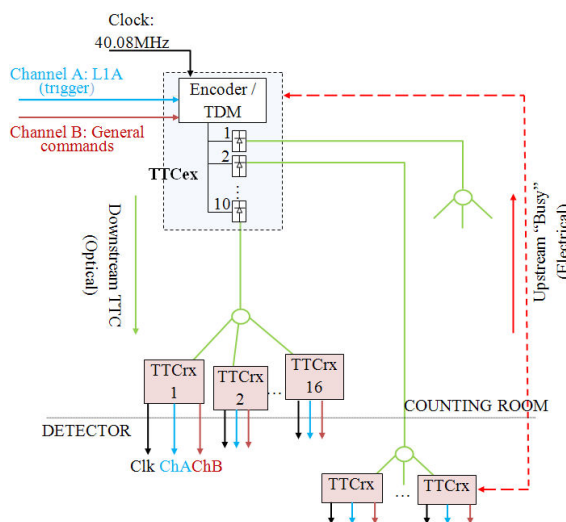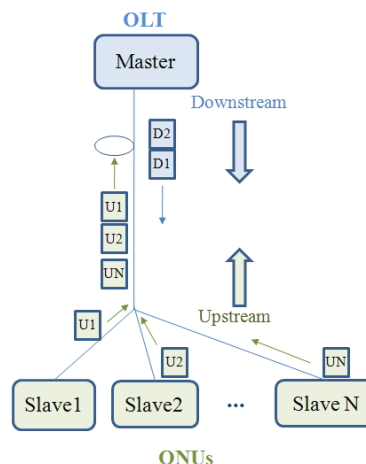

Figure 1: LHC TTC system.

Figure 2: A schematic representation of a Passive Optical Network

from the source to the destination. A master node, the Optical Line Terminal (OLT), communicates to a number of slave terminals, the Optical Network Units (ONUs), via a long feeder optical fiber and an optical splitter, Fig. 2. In the downstream direction (OLT→ONUs), PON is a broadcast network and so collisions cannot occur. Data are delivered to all ONUs which decide whether to further process them or to ignore them based on an address field. However, in the upstream direction (ONUs→OLT) a number of ONUs share the same transmission medium and so a channel arbitration
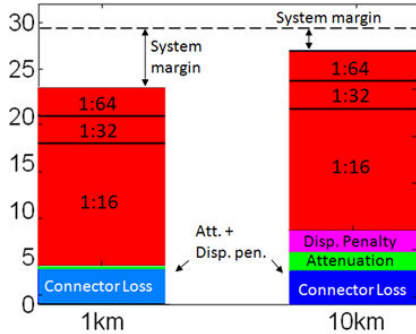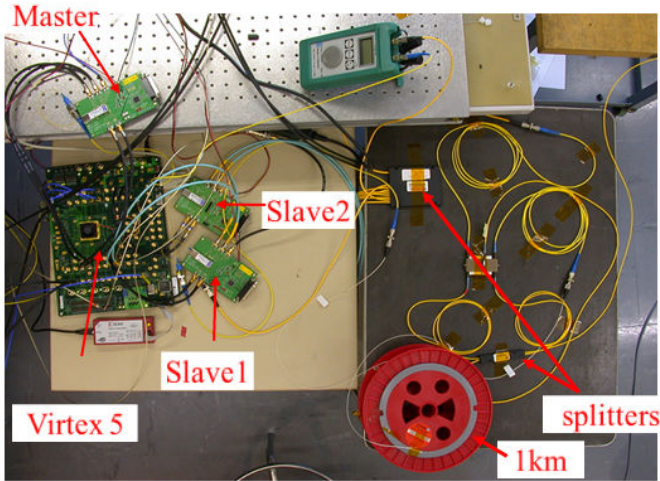
Figure 3: Optical link power budget diagram.

Figure 4: PON Demonstrator with one master (OLT) and two slave (ONU) nodes and 1km of fiber.

mechanism should be put in place to prevent collisions and to distribute bandwidth fairly among them. Time Division Multiple Access (TDMA) is the preferred multiplexing scheme in the first generation PONs as it is very cost effective while Dynamic Bandwidth Allocation (DBA) algorithms are employed for fairness. Typical commercial PON systems operate at 1.25Gb/s or 2.5Gb/s symmetric (downstream data rate equal to upstream), or asymmetric modes (downstream data rate higher than upstream).

III. PON DEMONSTRATOR

The aim of this project is to construct a PON demonstrator able to distribute trigger and command data with deterministic latency and fixed jitter while allowing the ONUs to communicate with the OLT in real time.

A. System Requirements and Specifications

A PON for TTC applications should meet the following requirement:

- System has to be able to deliver **synchronous** triggers and commands **continuously**
- **Latency has to be fixed** at both transmitting and receiving ends in the downstream direction
- A clock should be recovered from the downstream data with **low jitter**
- System should provide with the flexibility of both individually addressing or broadcasting to ONUs

- ONUs have to be able to respond in short time

Table 1: PON System Specifications

Property (General)	PON Demonstrator
Clock rate	40 MHz (ie LHC clock 40.08MHz)
Max distance	Up to 1000m
Encoding | Target BER	NRZ 8b/10b | <10-12
Splitting ratio	64
Frame Format	Commands + Trigger
BW Allocation Algorithm	Statistical Multiplexing
Property (Down|Up)	**PON Demonstrator**
Bit rate	1.6 Gb/s | 800 Mb/s
Latency	Fixed and Deterministic | To be determined
Received clock jitter	Able to drive a high-speed SERDES

The specifications of the system built are given in Table I. OLT and ONU transceivers were purchased from OESolutions [2] and were 1.25Gb/s EPON PX-20 standard compliant while the logic of our system was implemented on a Virtex 5, FPGA by Xilinx [3]. Power budget calculations, Fig. 3, revealed that we could comfortably support 64 ONUs in our network for 1km distance and so we designed our protocol to be able to support such a number of ONUs. However, due to restrictions to the number of evaluation boards and FPGAs we had at our disposal, we physically implemented a PON with 2 ONUs, Fig. 4, which were enough to allow us to demonstrate and to test all desired features.

IV. COMMUNICATION PROTOCOL

A feasibility study was conducted to evaluate the two commercial PON protocols, EPON (Etheret-PON) and GPON (Gigabit-PON), [4]-[5], and their potential to be used in our environment. The study concluded that none of the commercial protocols would be able to deliver the triggers with the strict timing requirements of the LHC experiments and so a custom protocol was devised that was addressing the following requirements:

- Synchronous delivery of a periodic trigger with clock rate 25ns, **(T)** field in Fig. 5.
- Auxiliary field to extend or to protect the trigger field, **(F)** field in Fig. 5.
- Broadcast or individual commands to ONUs, **(D1)** and **(D2)** field in Fig. 5.
- Arbitration of upstream channel to avoid collisions due to simultaneous transmissions from multiple ONUs, **(R)** field in Fig. 5.

A. Downstream Frames

Downstream is the most important direction for the network synchronization. According to the developed custom protocol, superframes are flowing in the downstream direction which consist of 65 subframes, Fig. 4. The beginning of each
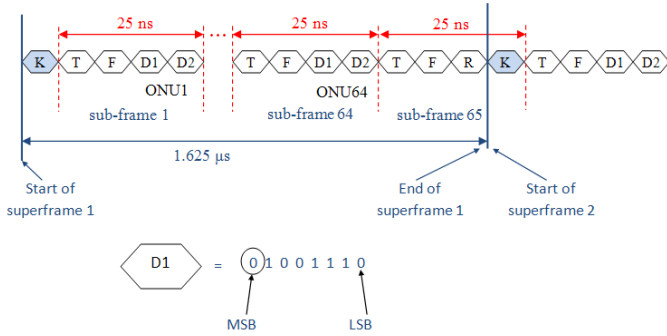
Figure 5: (a) OLT→ONU Upstream frame. Each field in the diagram corresponds to 1 byte. (b) Zoom in a D1 field to demonstrate how the distinction between broadcast or individually addressed ONUs is implemented.
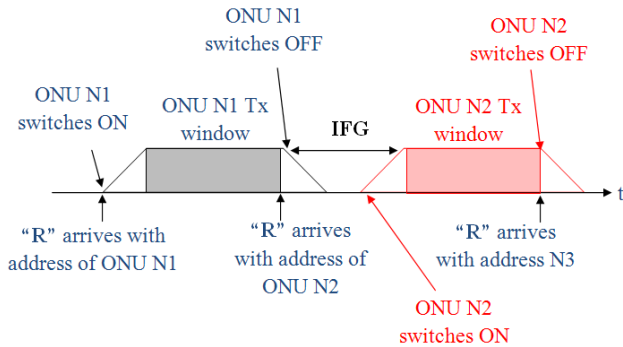

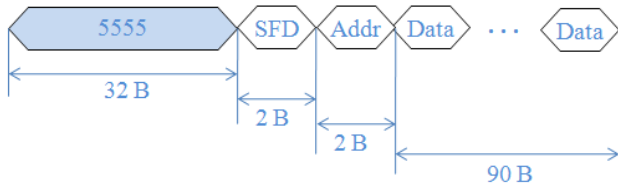Figure 6: Timing relationship between two successive ONU→OLT bursts
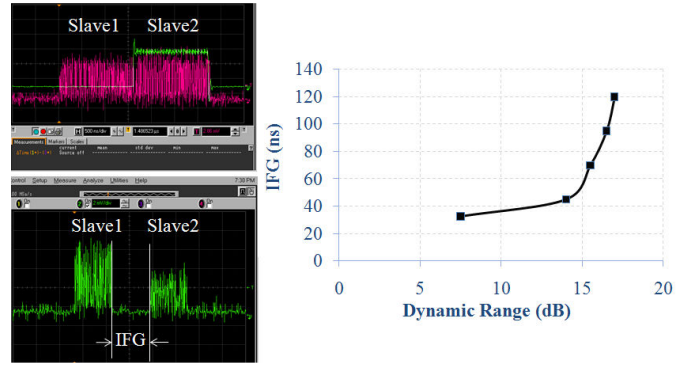

Figure 7: Upstream frame.

Figure 8: (a) Oscilloscope traces of bursts with different power (b) burst mode *Rx* dynamic range as a function of interframe gap (IFG).
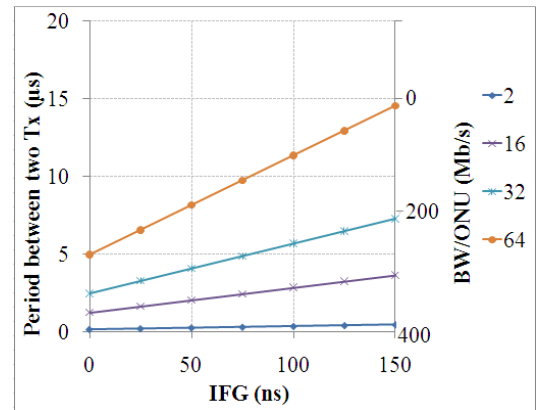

Figure 9: Waiting time between two successive transmissions from one ONU as a function of IFG.

superframe is signalled by a comma alignment character, <**K**>, which is used for synchronization and frame alignment. After, the <**K**> character the transmission of the first subframe begins. The first field of the subframe, <**T**>, carries the trigger information and is 1 byte long to provide the flexibility of assigning different triggers. The second byte, <**F**>, is an auxiliary field that might be used to either extend the trigger field or to protect it by means of forward error correction.

The last two characters in the subframe, <**D1**> and <**D2**>, carry commands intended for ONU1 only or commands broadcasted to all ONUs. The transmission duration of the four bytes (**T**, **F**, **D1** and **D2**) in each subframe is 25ns, at the at 1.6Gb/s downstream rate, corresponding to exactly one trigger period. Once the first subfame is finished the second subframe begins transmitting back-to-back. The structure of the second subframe is identical to the first one with the distinction that the D1 and D2 fields are now intended for ONU2 only unless if we operate in the broadcast mode. The distinction between individually addressed commands and broadcast commands depends on the most significant bit (MSB) in the D1 field, Fig. 5 (b). If this bit is "0" then we have a broadcast command if it is "1" then we have individual addressing. Sixty four such subframes are sent downstream,

as many as the number of supported ONUs, before the transmission of the last subframe that concludes the superframe. The 65th subframe is 3bytes long only, to restore the symmetry in the superframe and to allow the first trigger in the next superframe to be exactly 25ns apart from the last one. An important feature of the 65th subframe is that it finishes with an <**R**> character, which is used to arbitrate the occupation of the upstream channel as it will be explained in the next section.

B. Upstream Transmission and Frames

The <**R**> character carries the address of the next ONU to occupy the upstream channel. In the example shown in Fig. 6, an <**R**> characters arrives that contains the address of the ONU N1. Although the <**R**> character is received by everybody, only ONU N1 starts switching its laser on. After an initial period required for the power of the laser in the ONU N1 to stabilise, it starts transmitting its data in a predefined time window before it switches its laser off. Precautions have been taken to leave a gap without transmission between two successive transmissions from different ONUs, the interframe gap (IFG), to allow to the burst mode receiver at the OLT to get ready to accept a new burst.

The upstream frame is shown in Fig. 7. It starts with a long transition rich field (alternated 1s and 0s) to allow to the burst mode receiver to successfully recover the average transmission level and to set its decision threshold. It then contains a comma <**K**> character for frame alignment
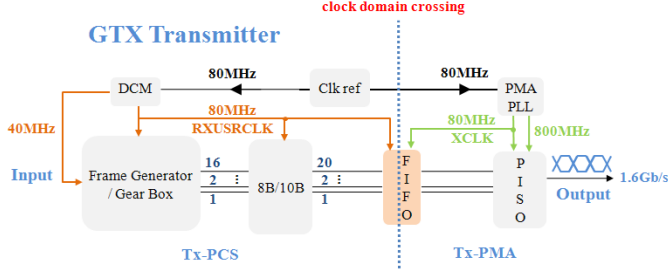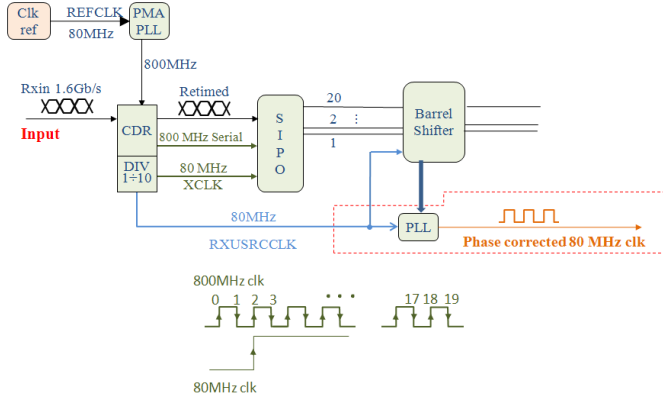
354

Figure 10: OLT transmitter implementation in FPGA.
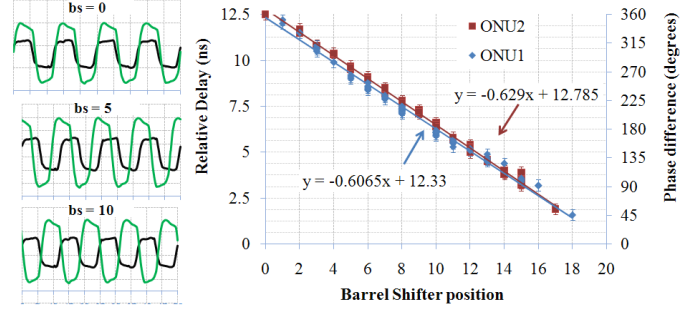

Figure 11: ONU receiver implementation in FPGA.

Figure 12: (a) Oscilloscope traces showing the phase difference between a reference clock (green line) and the recovered by the Rx parallel clock (blue line) for different barrel shifter values. (b) Relative delay between reference and recovered clocks as a function of the barrel shifter position for the two ONUs implemented in our system for 200 test cases.

followed by the address of the ONU and the transmitted data. The IFG affects the amount of bandwidth that is available in the upstream for pure data transmission and is closely related to the dynamic range of the receiver the maximum difference between the powers from two successive bursts for error free operation. Figure 8 shows experimental results of the IFG as a function of the dynamic range. According to Fig. 7, the larger the power difference between two bursts arriving at the OLT Rx, the larger the IFG required to maintain errorless operation. It is therefore advised to design a PON network whose branches are balanced, in terms of optical loss, to keep the IFG as small as possible and thus to maximize the upstream bandwidth.

Another important parameter in PON networks is the time that an ONU has to wait before it is able to occupy the transmission medium. Fig. 9 shows the waiting time between two consecutive transmissions from the same ONU as a function of the IFG and for different number of supported ONUs. The waiting time increases linearly with the IFG which is another reason to prefer balanced PONs that require minimum IFGs. At the same time as we add more ONUs in the system and we increase the IFG, the available bandwidth per ONU for data transmission reduces. Figure 9 reveals an interesting trade-off: On one hand we want to be able to design a network with as many client ONUs served by a single OLT as possible to reduce the cost of the system; On the other hand, the greater the number of supported ONUs the longer the waiting time. A balance between cost and waiting time must therefore be found.

V. TRANSCEIVER DESIGN IN VIRTEX 5 FPGA

This section introduces the transmitter and receiver designs for the upstream and downstream datastreams with emphasis given on the steps taken to achieve deterministic latency.

A. OLT Transmitter

The transmitter at the OLT is implemented based on the GTX transmitter of the Virtex 5, a more detailed description of which can be found in [3]. Latency issues at the Tx generally arise when data cross clock domains such as the Tx-PCS and the Tx-PMA in our case (Fig. 10). These two domains are clocked by the RXUSRCLK and the XCLK correspondingly, two clocks that are not phase aligned but have to be for the correct operation of the serializer block (PISO). The default method to phase align these two clocks is by using an elastic buffer (FIFO) which introduces a non-deterministic latency. Instead, we operate the GTX transmitter in advanced mode where we completely bypass the elastic buffer and use the PMA PLL to adjust the phase of the XCLK so that it matches the phase of the RXUSRCLK. The total latency through the transmitter was measured to be 75ns.

B. ONU Receiver

The ONU receiver design is shown in Fig. 11. The 1.6Gbit/s serial datastream is presented at the input of a CDR (clock and data recovery) circuit. The CDR recovers the clock from the incoming bitstream, retimes the data and passes them on to the next stage which is a serial-to-parallel circuit (SIPO). In addition, a divider generates the FPGA receiver parallel clock which is also fed to the SIPO and which affects the time that the parallel data leave from the SIPO. The operation of the divider is the most vulnerable part in the receiver with regards to achieving deterministic latency. This is because the 80 MHz parallel clock can lock on any edge of the serial 800 MHz clock when the receiver is reset introducing non-deterministic latency.

The latency issue that the divider introduces is solved by implementing a barrel shifter after the SIPO (Fig. 11). In order to identify the relative phase of the parallel clock compared to the serial clock, we take advantage of the <K> character in the downstream superframe and the fact that the order with which the parallel data exit from the parallel lines of the SIPO is affected by the operation of the divider as well. To make this point more explicit in the speculative scenario where the parallel clock started from the first edge of the serial clock, the first bit of the <K> character should come out from the first parallel line of the SIPO, the second bit from the
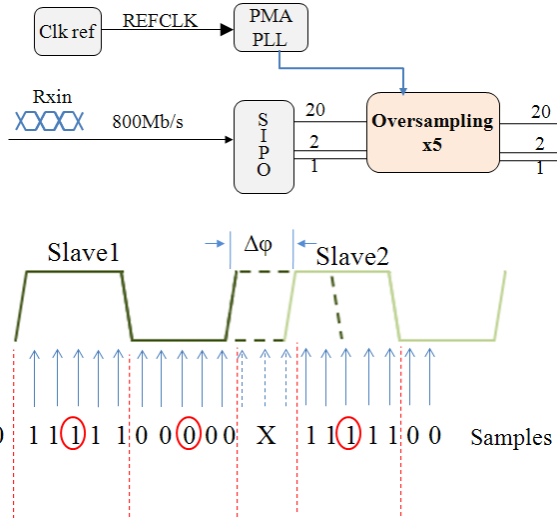
Figure 13: Burst mode receiver operating on oversampling mode.

second line and so on. However, if the parallel clock was delayed compared to the first edge of the serial clock then the first bit of the <K> character would be transferred to a different output line of the SIPO. The job of the barrel shifter is to identify which line exactly the first bit of the <K> character came out from and to feed this information to a PLL to perform the phase correction task. This last phase correction step has not yet been implemented.

Figure 12 demonstrates the operation of the barrel shifter concept by comparing the phase of a fixed reference clock with the phase of the recovered at the ONU clock for different barrel shifter values and for both ONUs supported by our system. The relative delay between reference and recovered clocks follows a linear trend. The slope of the two lines is 606ps and 629ps for the two ONUs correspondingly close enough, to within experimental error, to the expected value of 625ps that corresponds to the period between two consecutive edges of the serial clock. Based on these measurements, the barrel shifter concept will allow us to correct the latency at the receiver.

C. OLT Burst Mode Receiver

The burst mode receiver in the OLT, Fig. 13 (a), requires a 5x oversampling circuit. Burst mode oversampling works by blindly sampling the incoming datastream at a multiple of the bit rate and making a decision based on the sample that is closest to the center of the bit, [6]. This method is preferred over the usual implementations that use PLLs to recover the clock since PLLs typically have a large time constant and therefore are impractical to be used in high speed serial applications that involve bursts. The oversampling circuit generates 5 samples for each received bit (Fig. 13 (b)) and then tries to identify the transition region between bits. It is therefore important to provide a sufficient number of transitions in the datastream, a requirement satisfied by the long <5555> field transmitted in our upstream frame (Fig. 7). A decision circuit collects all samples from a predefined

window of incoming bits and implements a majority voting algorithm to identify the sample which is most likely to be closest to the center of the bit. If a burst from a second ONU arrives then it will be out of phase with the previous burst, Fig. 12 (b). In this case, the decision circuit will identify the new transition regions and adjust its decision sample.

VI. Future Developments

In order to complete our demonstrator system we will carry out the following implementations.

- The system will migrate onto two FPGA platforms, one for the OLT and one for the ONUs.
- The Barrel shifter position will be used to feed a PLL in order for the latency of the receiver at the ONU to become constant.
- Currently, we measured a jitter at the recovered parallel clock at the ONU of 166ps pk-to-pk and 36ps RMS which is worse than our specifications. An external PLL will be used to clean the jitter from the recovered clock.

VII. Conclusions

Our work has shown that bidirectional optical links based on Passive Optical Networks are excellent candidates for future TTC distribution systems. Optical links with fixed latency in the downstream direction and potentially low jitter where demonstrated while at the same time information was allowed to flow in the opposite direction through the same optical fiber. In future systems a ranging mechanism might be implemented through which the round trip time between the OLT and each ONU can be calculated. In this case, we need to ensure that the latency in the upstream direction is deterministic as well.

VIII. Acknowledgement

IX. References

[1] J. Troska, E. Corrin, Y. Kojevnikov, T. Rohlev, and J. Varela, "Implementation of the Timing, Trigger and Control System of the CMS Experiment," *IEEE Trans. Nucl. Scienc.,* vol. 53, pp. 834-837, Jun 2006.

[2] www.oesolutions.com.

[3] Virtex 5 GTX transceiver Available online. http://www.xilinx.com/support/documentation/data_sheets/ds022-1.pdf.

[4] EPON Standard is part of the Ethernet IEEE802.3 standard and is available online from IEEE. http://standards.ieee.org/getieee802/802.3.html.

[5] GPON is the ITU-T G.984 standard available online from ITU. http://www.itu.int/rec/T-REC-G/e.

[6] J. Kim, and D.-K. Jeong, "Multi-Gigabit-Rate Clock and Data Recovery Based on Blind Oversampling," *IEEE Comm. Magazine*, Vol. 41, pp. 68-74, Dec. 2003.