The 10G TTC-PON: challenges, solutions and performance

E.B.S. Mendes, S. Baron, D.M. Kolotouros, C. Soos and F. Vasey

a CERN, Geneva 23, 1211 Switzerland
b High Energy Physics Group, F3.303, Physics Department, University of Ioannina, Ioannina, 45110 Greece

E-mail: eduardo.brandao.de.souza.mendes@cern.ch

ABSTRACT: The TTC-PON (Timing, Trigger and Control system based in Passive Optical Networks) was first investigated in 2010 in order to replace the current TTC system, responsible for delivering the bunch clock, trigger and control commands to the LHC experiments. A new prototype of the TTC-PON system is now proposed, overcoming the limitations of the formerly presented solutions. A new upstream data transmission scheme relying on longer bursts is described, together with a high-resolution calibration procedure for aligning bursts in a time division multiplexing access. An error correction scheme for downstream data transmission is also discussed.

KEYWORDS: Detector control systems (detector and experiment monitoring and slow-control systems, architecture, hardware, algorithms, databases); Digital electronic circuits; Trigger concepts and systems (hardware and software); Front-end electronics for detector readout

Corresponding author.
Introduction to TTC-PON

In the Large Hadron Collider (LHC) experiments at CERN, the Timing, Trigger and Control (TTC) system is responsible for delivering the timing, trigger and control commands from the central processor to the detector sub-partitions. It consists of a point-to-multipoint unidirectional optical communication system in which the master can send data at 80 Mbps to up to 32 slave nodes [1]. In order to transmit the busy information (a single bit status indicating that the front-end buffers are in a warning state) from the slave nodes to the master, a separate electrical link is employed. This system presents several limitations for the future needs of the LHC experiments operation: a bidirectional link and a higher bandwidth are needed.

In the telecommunications industry, passive optical networks (PON) are a widely adopted solution in the Fiber To The Home (FTTH) deployments [2]. It consists of a point-to-multipoint bidirectional optical system in which a master node, called Optical Line Terminal (OLT), and several slave nodes, called Optical Network Units (ONUs), are present. Since the system is bidirectional, two directions of data transmission are distinguished: in the downstream direction (from OLT to ONUs) information is broadcast continuously at a given wavelength; in the upstream direction (from ONUs to OLT), information is transmitted with another wavelength making use of a multiplexing access scheme (commonly time division multiplexing, in which data transmission occurs in bursts) in order to allow the data transmission from several slaves.

The TTC-PON project was born from the idea of making use of the well-developed PON technology to propose an upgrade to the current off-detector TTC system, overcoming some of its limitations. TTC-PON is based on FPGAs for data transmission and reception since the system is located in the back-end of the experiments which is radiation-free. The TTC signals are then transmitted to the front-end of the experiments through the GBT (Gigabit Transceiver) link [3]. The architecture of the current TTC system and the TTC-PON system are shown in figure 1.

A first proof-of-concept of TTC-PON based on the 1G-EPON technology was implemented in 2010 [4] and fully characterized in 2012 [5]. In 2014 a first demonstrator of 10G-TTC-PON
was proposed [6] based on the 10G-EPON technology and was then deployed over the XG-PON technology [7] in 2015.

Despite an interesting high density time division multiplexing scheme allowing short waiting time for upstream transmission, the 2015 proposal presented some limitations in the upstream data transmission scheme (very low dynamic range, high sensitivity to temperature variations and need for customization of the commercial components) and in the downstream data transmission scheme (low split ratio). In addition, all the previously developed prototypes were using a single FPGA to emulate the full system. Even though the FPGA firmware was carefully partitioned, this solution was not close enough to the final application.

A new demonstrator was therefore implemented using individual FPGAs for each node of the system. Following a relaxation of the experiment specifications for upstream latency (to approximately 10 $\mu$s) and in order to overcome the limitations of the formerly proposed solution, a new upstream data transmission scheme consisting of longer bursts (125 ns) has been defined and a fine calibration procedure allowing high-resolution (0.417 ns) burst positioning has been implemented. In the downstream data transmission path, an error correction scheme was developed in order to allow a higher split ratio with comfortable margin. The challenges and proposed solutions of this new prototype are discussed in this paper, which is organized as follows: in section 2, an overview of the system is given; in section 2.1 the downstream path implementation and the error correction scheme is introduced and in section 2.2 the upstream path solution is discussed with an emphasis on the calibration procedure implementation.
2 System challenges and solutions

In the 2016 TTC-PON prototype, each node of the system (OLT, ONU) is built independently in a different Xilinx-Kintex7 FPGA evaluation board (KC705), which can be fully controlled via Ethernet. Figure 2 shows the current prototype and the corresponding block diagram.

The system is based on the XG-PON technology which allows a downstream transmission at 9.953 Gbps and an upstream transmission at 2.488 Gbps. Those values have to be slightly modified to 9.618 Gbps and 2.404 Gbps since the downstream data transmission rate has to be an even multiple of the LHC bunch clock frequency (40.078 MHz). For simplicity reasons, in the next sections, timing values are expressed with respect to a 40 MHz clock (i.e. 9.6 and 2.4 Gbps).

2.1 Downstream path

In the downstream direction, the OLT transmits continuously at 9.6 Gbps with a wavelength of 1577 nm. The frames are constructed on a LHC bunch clock period basis, appending an 8-bit header and using 24-bits for internal control communication between the OLT and ONUs (which can be potentially decreased to 8-bits).

In order to have a transmission synchronous to the 40 MHz LHC bunch clock, the reference clock (240 MHz) of the QPLL [8], responsible for generating the high-speed clock of the transceiver, has to be derived from a PLL that assures a fixed and deterministic phase at each reset or power up.

A particular requirement of our system in the downstream direction, which is normally not present in the telecommunications industry, is to have a fixed and deterministic latency after resets and power ups. The latency non-determinism in a high-speed serial link mainly comes from
two factors [4, 9]: (1) Elastic buffers are used to transfer data from the user clock domain to the transceiver PMA (Physical Medium Attachment) clock; (2) the receiver recovered low speed parallel clock can lock into any edge of the high-speed serial clock. The solution for (1) is to bypass the internal FIFO of the transceivers and perform a phase alignment procedure and (2) is solved by a frame alignment procedure that relies on a clock shift scheme making use of the rx_slide feature in PMA mode of GTX transceivers [8].

For simplicity reasons, initially an 8b10b encoding scheme was used in the downstream data transmission path, however, a limitation in terms of optical power budget was observed [10]. For the power budget calculation, the minimum OLT transmitter power is 3.68 dBm and an ONU receiver sensitivity of $-19.2$ dBm is adopted. The attenuation in the system for a 1:64 splitter is calculated as follows: the loss of a 1:64 splitter is the worst-case in the splitter datasheet (20.5 dB) and an additional loss of two connectors (0.5 dB/each) and 100 meters long fiber (0.5 dB/km) is assumed.

As shown in figure 3, the optical margin observed is 1.33 dB, which is an insufficient safety margin with a 1:64 split-ratio (we adopt 3 dB as safe margin, which is a relatively conservative approach).

### 2.1.1 Error correcting scheme

In order to increase the power budget in the downstream direction and allow a higher optical margin for a 1:64 network, an encoding scheme based on scrambling and forward error correction scheme is proposed. A self-synchronous scrambling approach is adopted to avoid the synchronization overhead of frame-synchronous scrambling or distributed-sample scrambling [11]; the drawback of self-synchronous scrambling is the error-multiplication phenomena. In order not to have error multiplication before decoding, scrambling and descrambling operations are performed before encoding and after decoding respectively; also, a low weight polynomial is chosen for the scrambling process.

Forward error correction is a technique in which the receiver is able to detect and correct errors without requiring re-transmission of the data, at the expense of bandwidth. Several methods exist in the literature, which can be classified in two big families: convolutional-codes and block-codes. Due to the frame-nature of our protocol, a preference was given to block-codes, notably: Hamming (single-bit error correcting), binary BCH (Bose-Chaudhuri-Hocquenghem, multiple-bit error correcting), and Reed-Solomon (multiple-symbol error correcting) codes [12]. In a block-code, a message of length $k$ is mapped to a codeword of length $n$ in a unique way by adding $n-k$ redundant bits/symbols. The mapping is done in such a way that the minimum distance between any two codewords is maximized.

In order to identify the best suited ECC (Error-Correcting Code) for our system, several parameters were taken into account: code efficiency, coding gain, latency and hardware complexity. BCH codes are good candidates for our system, due to the good random error correction capability and relatively low complexity [13]; four shortened BCH$(n,k)$ codes were implemented (table 1).
Table 1. Summary of implemented BCH codes.

| Original code     | Shortened code implemented | Error correction capability | User payload per Bunch clock |
|-------------------|----------------------------|-----------------------------|------------------------------|
| BCH (63, 57)      | BCH (40, 34)               | single-error                | 172 bits                     |
| BCH (127, 120)    | BCH (80, 73)               | single-error                | 187 bits                     |
| BCH (127, 120)    | BCH (120, 113)             | single-error                | 194 bits                     |
| BCH (127, 113)    | BCH (120, 106)             | double-error                | 180 bits                     |

Figure 4. BER measurements for different codes.

Figure 5. FEC tradeoff (decoder parameters).

using a systematic encoding approach in order to keep the DC balance properties of the scrambled data and simplify the decoder design.

We can observe in figure 4 the Bit Error Ratio (BER) measurement results from the four codes implemented in terms of Optical Modulation Amplitude (OMA). As a reference, the GBT-Reed Solomon code (GBT-RS(15, 11)) is also shown. For the single error correcting codes, we obtain a coding gain at BER = $10^{-11}$ of around 2 dB. It is interesting to observe that for the three single-error correcting codes, the correction capability does not change significantly with the codeword length. The double error correcting code provides us with a 3 dB coding gain. In order to analyse the other important parameters affecting our system, we take as figures of merit the design parameters of the FPGA implementation of the decoder: timing slack, latency and slice lookup tables (LUTs). The encoding process is much simpler than the decoding process and is thus not considered, as shown in figure 5. In terms of area (slice LUTs), the double-error correcting code is slightly more complex than the single-error correcting codes, however for modern FPGAs the overhead is still very small compared to the total number of available slice LUTs. We can also note that the decoding latency of the double-error correcting code is one clock cycle bigger than the single-error correcting code decoders, since an extra pipeline stage had to be inserted in order to meet timing due to a long critical path.

The most promising analysed code is the BCH (120, 113). This ECC can provide us with a large user payload (194 bits/bunch clock) and is able to provide us with a coding gain of 2 dB which is enough to reach the targeted 3 dB for the optical margin.
2.2 Upstream path

The upstream data transmission scheme is based on Time Division Multiplexing Access (TDMA). Each ONU has a time slot allocated to transmit 8b10b encoded information. The ONU burst is composed of four main fields: preamble, delimiter (comma), ONU address and payload as depicted in figure 6.

For TTC-PON, a very important figure of merit is the waiting time, which can be defined as the time between two transmissions from the same ONU; in our system, the waiting time affects the busy latency which should be as small as possible. For this reason, working with short bursts is highly desirable, which means trying to minimize as much as possible the overhead in the burst (gap and preamble).

In order to understand the need of the overhead, one should note the main difference between continuous-mode and burst-mode receivers. In continuous-mode transmission, the average amplitude of the incoming optical signal to the receiver does not vary rapidly with time. On the other hand, in burst-mode transmission, the difference of the average power of different ONUs can be high, and the receiver should be able to adapt to large power variations with a fast settling-time. The preamble is an alternating sequence of 0’s and 1’s that allows the burst mode receiver to adapt to the threshold of the incoming burst. In order to erase the threshold setting of the previous burst and prepare for a new threshold setting, the XG-PON standard recommends that the burst-mode receivers use an external reset between bursts, which is positioned in the gap. This reset signal is a timing critical signal coming from the MAC to the PHY layer. Studies on auto-generated reset burst-mode receivers are also present in the literature [14].

For the OLT XG-PON optical module used in the 10G TTC-PON prototype [15], a minimum reset duration of 25.6 ns is required, and a maximum settling time of 52 ns is necessary. Therefore, in the final scheme, a 125 ns long burst scheme is adopted. A payload of 64 bits per burst is available of which 16 bits are used for internal control and 48 bits are allocated to the final user.

There are also three other main challenges in the upstream implementation: (1) the transmission of ONUs has to be synchronous to the OLT clock, which requires a local synchronization procedure, that is, the downstream recovered clock is cleaned by means of an external PLL and used to drive the transmit PLL (CPLL in our design [8]) of the FPGA transceivers; (2) A classical CDR scheme cannot be used due to the excessively high time-to-lock, and a $\times 4$ oversampling scheme is proposed as depicted in [6]; (3) The token for the time division multiplexing arbitration is not passed by the OLT but each ONU has a fixed offset to transmit information, which requires a fine resolution calibration as explained in the next section.
The upstream data transmission scheme described above was tested performing Burst Bit Error Ratio (BBER) tests as described in [5]. The BER curves of figure 7 (left) were measured during nine days with a target BER $= 10^{-11}$ and a confidence level of 95% [16]. The worst observed sensitivity at BER $= 10^{-11}$ is $-24$ dBm (OMA) and the receivers are within the specifications of the vendor at BER $= 10^{-4}$. In order to estimate the power budget, a pessimistic projection of the worst measured BER curve is performed and a sensitivity of $-22$ dBm (OMA) is adopted. The power budget is calculated for a minimum ONU transmitter power of 2.79 dBm (OMA), that results in a power budget of 24.79 dB.

From these measurements, the power budget is enough to provide us with a margin higher than 3 dB for a 1:64 split-ratio considering a 100 meters long fibre.

Dynamic range tests were also performed by varying the power of ONU2 with respect to the power of ONU1 thanks to an extra optical attenuator. No penalty was observed for up to 12 dB of power difference between the bursts as depicted in figure 8.
2.2.1 Time division multiplexing arbitration

In TTC-PON the upstream bandwidth allocated to each ONU is fixed and the number of ONUs is known before a run is initiated. Those characteristics make it possible to have a fixed time slot for the transmission of each ONU during a run and therefore the OLT does not have to pass the token every time an ONU is allowed to transmit. Instead, a broadcast signal (heartbeat) is sent every $T_{HB} = \text{Number\_ONUs} \times 125\,\text{ns}$ seconds; two headers are used in downstream transmission, the regular header sent every 25 ns, and the heartbeat header sent every $T_{HB}$. The heartbeat serves as a time reference signal to the ONUs which reset an internal counter every time a heartbeat is received. As long as the heartbeat signal arrives at all the ONUs at the same instant, all the counters will have the same state and by giving an appropriate offset to each ONU, the bursts are aligned in time. If the heartbeat signal does not arrive at all the ONUs at the same time — which could be a consequence of different fibre lengths of different ONUs — the ONUs internal counters will have different states and therefore a collision could potentially happen.

This limitation is overcome with a calibration procedure before a run is started. In a classical PON system, a similar ranging procedure is also proposed. The calibration consists of performing a roundtrip latency measurement for each ONU and compensating the transmission offset depending on the roundtrip value; the procedure is arbitrated by the OLT, a given ONU is put in calibration mode (transmits continuously and sends back the received header, acting as a header-mirror, other ONUs are turned-off) and the distance between the transmitted heartbeat header and the received heartbeat header is measured. A resolution of up to one upstream unit interval (0.416 ns) can be achieved relying on the barrel shifter value used for data alignment in the oversampler. The calibration procedure is illustrated in figure 9.

### 3 Conclusions

The 2016 XG-PON based prototype for TTC-PON overcomes the main limitations of the previously proposed prototypes. Thanks to the implementation of an error correcting scheme, a split-ratio of 1:64 can now be comfortably met with a margin higher than 3 dB. In the downstream direction, the user can make use of up to 194 bits per bunch clock (which corresponds to 7.76 Gbps). In the upstream direction, a new protocol relying on a 125 ns burst length was implemented, providing the user with 48 bits per burst. The upstream latency is now 8 us for 64 ONUs, and thanks to the use of the burst mode receiver reset, the dynamic range can be as high as 12 dB, allowing a non-balanced optical tree.
The design migration to the FPGA platforms Kintex Ultrascale from Xilinx [17] and Arria10 from Altera [18] is foreseen for Phase-1 upgrades.

New PON protocols are now entering into the market: XGS-PON which is based on a symmetric data-rate (10 Gbps) transmission scheme [19]; NGPON-2 which is based on a wavelength division multiplexing to allow a data transmission at 40 Gbps in downstream direction [20]. A feasibility study of TTC-PON based on those protocols is also foreseen for Phase-2 upgrades.

References

[1] J. Varela, Timing and synchronization in the LHC experiments, talk given at the 6th Workshop on Electronics for the LHC Experiments, Krakow, Poland, 11–15 September 2000 http://ttc.web.cern.ch/TTC/LEB00Sync.pdf.

[2] T. Koonen, Fiber to the home/fiber to the premises: what, where and when?, Proc. IEEE 94 (2006) 911.

[3] P. Moreira et al., The GBT-SerDes ASIC prototype, 2010 JINST 5 C11022.

[4] I. Papakonstantinou et al., A fully bidirectional optical network with latency monitoring capability for the distribution of timing-trigger and control signals in high-energy physics experiments, IEEE Trans. Nucl. Sci. 58 (2011) 1628.

[5] D.M. Kolotouros et al., Metrics and methods for TTC-PON system characterization, 2014 JINST 9 C01015.

[6] D.M. Kolotouros, S. Baron, C. Soos and F. Vasey, A TTC upgrade proposal using bidirectional 10G-PON FTTH technology, 2015 JINST 10 C04001.

[7] ITU, G.987.2: 10-Gigabit-capable passive optical networks (XG-PON): Physical media dependent (PMD) layer specification, (2010) https://www.itu.int/rec/T-REC-G.987.2-201001-S/en.

[8] Xilinx, 7 Series FPGAs GTX/GTH Transceivers. User Guide. UG476 (v1.12), (2016) https://www.xilinx.com/support/documentation/user_guides/ug476_7Series_Transceivers.pdf.

[9] R. Giordano and A. Aloisio, Fixed-latency, multi-gigabit serial links with Xilinx FPGAs, IEEE Trans. Nucl. Sci. 58 (2011) 194.

[10] S. Baron and E. Mendes, Recent Developments in the TTC-PON system, talk given at the EP-ESE Electronics Seminars, CERN, Geneva 23, Switzerland, 31 May 2016 https://indico.cern.ch/event/465344/.

[11] B.G. Lee and S.C. Kim, Scrambling Techniques for Digital Transmission, Springer-Verlag (1994).

[12] E. Berlekamp, Algebraic Coding Theory, revised edition, World Scientific (2015).

[13] X. Zhang, VLSI Architectures for Modern Error-Correcting Codes, CRC Press (2016).

[14] X. Qiu, [OFC 2013 Tutorial OW3G.4] Burst-mode Receiver Technology for Short Synchronization, in proceedings of the 2013 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC), Convention Center, Anaheim, U.S.A., 17–21 March 2013 http://ieeexplore.ieee.org/document/6533160/.

[15] Ligent, LTF7222 XG-PON1 OLT Transceiver, (2016) http://www.ligentphotronics.com/html/Products/XGPON1_XGSPON/20160912_9.html.
[16] J. Redd, *Calculating statistical confidence levels for error-probability estimates*, (2000) http://www.lightwaveonline.com/articles/print/volume-17/issue-5/features/calculating-statistical-confidence-levels-for-error-probability-estimates-53462167.html.

[17] Xilinx, *UltraScale Architecture GTH Transceivers. User Guide. UG576 (v1.4)*, (2016) https://www.xilinx.com/support/documentation/user_guides/ug576-ultrascale-gth-transceivers.pdf.

[18] Intel, *Arria 10 Transceiver PHY User Guide. UG-01143*, (2016) https://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/hb/arria-10/ug_arria10_xcvr_phy.pdf.

[19] ITU, *G.9807.1: 10-Gigabit-capable symmetric passive optical network (XGS-PON)*, (2016) https://www.itu.int/rec/T-REC-G.9807.1-201606-P/en.

[20] M. Carroll, Z. Ye and D. Remein, *FSAM & ITU-T Activities on Next-Generation PON Stage-2 (NG-PON2)*, in *IEEE 802.3 NGEPON ad hoc meeting*, Beijing, China, March 2014 http://www.ieee802.org/3/ad_hoc/ngpon/public/mar14/remein_ngpon_01_0314.pdf.