A TTC upgrade proposal using bidirectional 10G-PON FTTH technology

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

CERN, 1211 Geneva, Switzerland
University of Ioannina, 45110 Ioannina, Greece

E-mail: dimitrios.kolotouros@cern.ch

ABSTRACT: A new generation FPGA-based Timing-Trigger and Control (TTC) system based on emerging Passive Optical Network (PON) technology is being proposed to replace the existing off-detector TTC system used by the LHC experiments. High split ratio, dynamic software partitioning, low and deterministic latency, as well as low jitter are required. Exploiting the latest available technologies allows delivering higher capacity together with bidirectionality, a feature absent from the legacy TTC system. This article focuses on the features and capabilities of the latest TTC-PON prototype based on 10G-PON FTTH components along with some metrics characterizing its performance.

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

1Corresponding author.
1 Introduction

The Timing-Trigger and Control (TTC) system is a crucial system dedicated to synchronization of experiment electronics to the LHC beam. Currently, it is a unidirectional network extensively deployed in all major detectors. It distributes the LHC bunch clock and the level-1 trigger accept decision (L1A) as well as individually addressed or broadcast commands to the various detector sub-partitions [1]. The lack of bi-directionality is currently compensated by implementing an extra network in charge of propagating the busy/throttle and other required signals up to the Central Trigger Processor. To match the needs for increased payload capacity and to provide bi-directionality, a new generation TTC system is being investigated for off-detector use, based on Passive Optical Network technology (PON). A PON is a bidirectional, single fibre, point-to-multipoint network architecture in which optical splitters are used to enable a master node or Optical Line Terminal (OLT) to communicate with up to 128 slave nodes or Optical Network Units (ONUs). It is based on already mature devices, as the PON is nowadays the most successful solution worldwide for deploying Fiber To The Home (FTTH, FTTx) networks [2]. A first TTC-PON demonstrator was built in 2010 during early investigations made at CERN, using commercial FPGAs (Field Programmable Gate Array) and 1-Gigabit Ethernet PON transceivers [3]. The first and very promising results motivated the work to explore the emerging XG-PON technology in order to better fit the user requirements in terms of latency and payload. Based on the backbone of the first TTC-PON demonstrator, a second demonstrator using 10G-PON devices was developed to include all the features of its predecessor with enhanced performance. In addition, it introduces some new features:
dynamic software partitioning targeting up to 128 nodes per network tree; large downstream payload; low downstream trigger latency; and embedded busy/throttle signals on the upstream path with bounded latency. The present phase of the TTC-PON project consists of tuning the 10G-PON technology to match our requirements. The protocol of the downstream path is optimized to allow fixed and low latency delivery of recovered clock as well as trigger and trigger types. On the upstream direction, all the parameters of the Time Domain Multiplexing protocol are being pushed far from the FTTx usual orders of magnitude in order to increase the individual bandwidth of each node and reduce the waiting time of busy/throttle signals. This required implementing custom synchronization and Clock and Data Recovery (CDR) schemes, especially on the upstream path. This article focuses on the features and capabilities of this new TTC-PON along with some metrics characterizing its performance so far, aiming to commission a first prototype within an experiment and test it onsite in 2015.

2 Current TTC

The existing TTC architecture in LHC is shown in figure 1a. It consists of TTCvi and TTCex cards which communicate with a number of custom receiver components (TTCrx), via a passive optical tree. The TTCex receives information from two channels activated by the trigger control system(Central/Local Trigger and TTCvi); Channel A contains the level 1 trigger accept (L1A) decision and channel B carries general purpose commands for the synchronization and calibration of the detector partitions. TTCex multiplexes the two channels in time, encodes them and uses the data to drive a bank of up to 10 Fabry-Perot lasers. Data are broadcasted through an optical fiber which is about 100 m long and are distributed to a maximum of 32 TTCrxs via an optical fan-out. The downstream data rate is 40 Mb/s and the bi-phase mark format is used to encode the data.

![Current TTC architecture.](a) Current TTC architecture.  
![TTC-PON architecture.](b) TTC-PON architecture.

**Figure 1.** Examples of TTC architectures.

The TTCrx acts as an interface between the TTC system and the detector partitions. Its function is to recover the LHC clock with deterministic phase and distribute it to the front-end detector electronics. The clock is de-skewed to compensate for variable particle times of flight and cleaned before distribution. TTCrx also de-multiplexes channels A and B and delivers the synchronization commands, the L1A trigger-accept decisions (with very low and fixed latency) and their associated bunch and event identification numbers to the front-end electronics. Although very efficient, this system is suffering from bandwidth limitations. It is also lacking bidirectionality and prevents users from sending feedback to the TCS (Trigger Control System). To tackle this problem
a separate electrical “busy/throttle” link delivers feedback on the status of the front-end readout buffers and of the data acquisition system to the trigger control system. In fact, the ‘busy’ signals from all detector partitions are merged in the fast merging modules (FMM) in the CMS case or in the ROD Busy modules in the ATLAS case, so that only one signal per detector partition finally reaches TCS. If a front-end buffer is close to overflow a “warning” signal is issued and TCS inhibits the L1A trigger-accept until the occupancy in the buffers falls below a predefined threshold and a “ready” signal is issued.

A similar architecture, presented in figure 1b, can be implemented to replace the current TTC architecture improving its performance. This architecture will be presented in the next section.

3 The 10G TTC-PON demonstrator

The bidirectional and point-to-multipoint architecture of passive optical networks theoretically allows building a TTC system without the drawbacks presented above. To prove it, a complete 1G TTC-PON demonstrator has been developed in 2010 during early investigations made at CERN, using commercial FPGAs and 1-Gigabit Ethernet PON transceivers as a proof of concept [3]. The 1G demonstrator was fully characterized in 2013 [4]. The very promising results motivated the work to explore the emerging XG-PON and 10G-EPON technologies in order to better fit the user requirements in terms of latency and payload. Based on the backbone of the first TTC-PON demonstrator, a second demonstrator using 10G-PON devices was developed to include all the features of its predecessor with enhanced performance (figure 2).

An equivalent TTCvi and TTCex instance has been realized (figure 1b) by combining the Kintex-7 FPGA and a commercially available 10G-EPON Optical Line Terminal (OLT) transceiver (RSL9988X-GGA by OESolutions) while three equivalent TTCrx instances have been realized by using the same FPGA and 10G-EPON Optical Network Unit (ONU) transceivers (RSN7877P-GGI by OESolutions). As both the OLT and ONU entities are using the same FPGA, a careful split of the logic and of the clock domains was required: each entity has its own reference clock and occupies a bounded area in the FPGA. The downstream direction (from OLT to ONUs) is based on a continuous optical serial link running at 11.2 Gbps, in the 1577 nm wavelength band carrying triggers, trigger types, clocks and other data synchronously broadcasted to all the branches of the network. The upstream direction (from ONUs to OLT), will mostly carry control, feedback and busy signals at 2.8 Gbps. Moreover the transmission medium is shared among the ONUs using

Figure 2. 10G TTC-PON prototype setup.

An equivalent TTCvi and TTCex instance has been realized (figure 1b) by combining the Kintex-7 FPGA and a commercially available 10G-EPON Optical Line Terminal (OLT) transceiver (RSL9988X-GGA by OESolutions) while three equivalent TTCrx instances have been realized by using the same FPGA and 10G-EPON Optical Network Unit (ONU) transceivers (RSN7877P-GGI by OESolutions). As both the OLT and ONU entities are using the same FPGA, a careful split of the logic and of the clock domains was required: each entity has its own reference clock and occupies a bounded area in the FPGA. The downstream direction (from OLT to ONUs) is based on a continuous optical serial link running at 11.2 Gbps, in the 1577 nm wavelength band carrying triggers, trigger types, clocks and other data synchronously broadcasted to all the branches of the network. The upstream direction (from ONUs to OLT), will mostly carry control, feedback and busy signals at 2.8 Gbps. Moreover the transmission medium is shared among the ONUs using
time domain multiplexing (TDMA) in the 1270 nm window — a much more challenging task requiring burst mode transmitter ONUs and a burst mode receiver in the OLT. Both datarates (Downstream and Upstream) were carefully chosen to satisfy the transmission of the bunch clock (BC), the FPGA specifications and the smooth operation of the optical transceivers. Different datarates (9.6 Gbps/2.4 Gbps) will be investigated in future studies to better match the datarates of other affiliated systems. The two link directions will be described in the next sections.

3.1 Downstream

In the Downstream direction the OLT broadcasts packets to all ONUs. The information passed can be addressed universally to all ONUs, individually to one of them using an address field embedded in the packets or to some of them using dynamic software partitioning.

3.1.1 Payload

A custom protocol was built in order to meet the needs of the LHC experiments and to serve as many ONUs as possible (128). With a bit rate of 11.2 Gbps and an 8B10B encoding scheme, the available payload is 8.96 Gbps, largely sufficient to broadcast triggers, trigger types, synchronous commands and configuration frames. Figure 3 shows a graphical representation of a downstream frame from the 10G TTC-PON versus the current TTC over the period of a BC. The transmitted frame is 28 Bytes long.

![Figure 3. 10G TTC-PON vs. current TTC payload.](image)

3.1.2 Performance

According to the specifications provided by the OLT and ONU manufacturer, the power budget allowed by such a PON system is of about 31 dB (@BER = 10^{-3}). A bit error rate (BER) test was run to verify this value using simultaneously three ONUs connected to the OLT. It was confirmed that this power budget was met by our system with small deviations, as shown in figure 4. As the BER target is 10^{-12}, the power budget of the downstream path is ≈ 25 dB. Considering that commercially available 1:128 splitters introduce an insertion loss of ≈ 26 dB, the current available OLTs will only allow a safe split ratio of 1:64. However the new generation PON systems will offer 4 dB more of power budget allowing us to target a split ratio of 1:128.

![Figure 4. Downstream BER vs. attenuation.](image)
3.2 Upstream

In the upstream direction, the prerequisites of the standard FTTx networks (TDMA, arbitration mechanism centrally managed at the OLT, asynchronous nodes, unequal branch lengths of the network tree) are inducing the standard FTTx protocol to have long bursts of data separated by long silent periods. This leads to extensive waiting time for each ONU to get the token. The TTC-PON system, however, can take advantage from its regular topology and specific requirements (synchronous ONUs, very low upstream payload) to gain in simplicity and to significantly reduce the waiting time for busy signals transmission. This section is presenting the way the standard FTTx protocol is modified to match the TTC-PON specificities.

3.2.1 Standard FTTx specification

Typically, an ONU data burst in FTTx framing can last up to $6 \mu s$ of which, 500 ns corresponds to the guard time (period during no ONU is transmitting), another 500 ns to the preamble and the rest $5 \mu s$ to user and encapsulation data. Network’s efficiency for 128 ONUs will be then around $\approx 65\%$ [5] while total waiting time will reach nearly 0.8 ms. The ratio between the average power level of one burst over the average power level of its preceding burst (dynamic range), of the order of 20 dBs should be tolerated.

![Figure 5](image.png)

Figure 5. a) FTTx upstream framing, b) 10G TTC-PON framing (not to scale).

3.2.2 Customizing the FTTx protocol

Unlike the FTTx framing (IEEE802.3av), in the 10G TTC-PON there will be no token for the upstream channel arbitration. The firmware of each ONU runs synchronously to the OLT main clock and is identical to its neighbors, except that its talking window — defined by an offset value in an absolute counter — is different. The token is thus automatically passed without requiring the OLT arbitration. This requires careful synchronization and calibration routines presented in next sections. The line rate is set to 2.8 Gbps (see section 3.2.5). The proposed guard time for the TTC-PON will be of the order of 20 bits ($\approx 7.14$ ns) in order to reduce dead time. The preamble shrunk as well. Currently its length is 40 bits unencoded (14.28 ns) with the aim to reach 20 bits. The payload is significantly reduced also to 20 bits unencoded (7.14 ns). While it reduces the network efficiency to 20%, it dramatically reduces the total upstream waiting time. In the future, when testing and characterization are concluded, the payload will be 8B10B encoded, resulting to 16 bits (2 Bytes) of payload for the user. Taking into account all the above, each ONU burst will last 35.7 ns, including also the guard time. Therefore the total waiting time for 128 ONUs will be $4.57 \mu s$. This customization of the protocol is facilitated by the fact that the maximum dynamic range in the final network tree will be smaller than 6 dB.
3.2.3 Link synchronization

The implementation of short bursts necessitates the use of a link synchronization scheme: downstream and upstream paths are fully synchronous in terms of frequency. The 40 MHz BC clock is recovered from the serial data in the ONU receiver and a deskewing process based on the FPGA digital clock management (DCM) module makes it possible to fine tune its phase [6]. The recovered clock then, is used to drive a high-speed serializer for further distribution of data to the detector front-end components [3]. A second use of the recovered clock is to frequency and/or phase lock part of the logic and the transmitter part of the ONU’s MGT. The use of a common clock across the network can have many benefits: fine control of the bursts’ start time, easier recovery in the OLT side etc. The link synchronization can be achieved in the ONU with two different approaches: using an external clock synthesizer with deterministic phase (e.x. Si5338) that will clean the recovered clock and provide a high-quality clock reference for the transceiver; using a vendor specific (Xilinx) technique [7] to perform the same task without the need of any external components. The former principle has been validated using a climatic chamber and numerous reset cycles.

3.2.4 Calibration procedure

Running with the same clock, every ONU can have the same internal counter in order to know when is the right time to transmit information to the OLT. An offset value, different in each ONU and depending on its address, could theoretically ensure collision avoidance on the prerequisite that all fibers in the network tree have the exact same length.

In a real life setup however, fiber lengths cannot be guaranteed to be equal, they can even differ by several meters, resulting in collisions. This is the reason why a calibration process is needed: during initialization, a response time measurement takes place for each ONU (figure 7). The OLT gathers the information from all ONUs and sends individually an extra offset value. This value will be used by the ONU to adjust its transmission time (offset value previously mentioned) and allow a precise alignment of the bursts being Time Domain Multiplexed.

3.2.5 Very fast fine phase alignment

The last major concept designed to allow the use of such short bursts, is the “very fast fine phase alignment”. The bursts have random phases between them and the receiver in the OLT should
quickly adjust its decision point in time to be able to reconstruct the incoming data stream with a very low error rate. Using conventional CDRs would require up to 4.46 μs for clock recovery and phase acquisition [8], thus they can be described as “slow” for burst mode operation.

A workaround to this problem is to blindly oversample the incoming data: conventionally, on a serial line the transmitter and the receiver operate at the same frequency (parallel clock), same bus width and thus same line rate. This is not the way the TTC-PON operates: upstream, each ONU transmits data at 2.8 Gbps rate (20 bits @ 140 MHz). The receiver (OLT side) however, delivers the recovered serial datastream at a different rate. As it doesn’t need to recover the clock (ONUs are transmitting synchronously to the OLT), it can run at 11.2 Gbps (80 bits @ 140 MHz). Consequently, for each bit transmitted by an ONU, the OLT receives 4 bits. The received 80 bit sequence is then dissected into 20 groups of 4 bits. Three more copies of this sequence are also dissected and grouped in the same way but shifted by one bit (figure 8). For each group of each copy of the input data, a look-up table, carefully filled after thousands of iterations, decides whether it has received a “0” or “1” and if the sample contains any edges. The result is four 20 bit wide words, each of them with a different phase. Four parallel word aligners then try to align each word by searching for a comma (figure 9). Usually, only one or two aligners will detect this comma and output useful data. This procedure is continuously running, the phase being updated at each detected comma (i.e. at the start of each burst). Therefore, the fine phase alignment is automatically adjusted for each ONU on a continuous manner, and can track smooth variations of phase due, for example, to slight temperature variations.

![Figure 8. ONU TX width vs. OLT RX width.](image)

![Figure 9. Very fast fine phase alignment.](image)

### 3.2.6 Performance

In order to characterize the upstream path and quantify the efficiency of the phase alignment method described above, repetitive and concurrent Burst-Bit Error Rate Tests (B-BERT) were performed...
with three ONUs transmitting, using the customized TDMA protocol previously described [4]. Each B-BERT was targeting $10^{-11}$ and was performed applying an increasing attenuation on the common fiber (see figure 2), starting from 26.5 dB up to 35.5 dB in steps of 0.25 dB. A reset cycle of both ONUs and OLT was performed between each B-BERT sequence, resulting in slightly different phase relationships between bursts (of the order of a few ns). Performing more than a hundred consecutive B-BERTs consequently allowed accumulating statistics on the system performance. The results are presented in figure 10. Each graph — one per ONU — was concurrently built and represents one hundred consecutive B-BERTs separated by system resets. These results show the excellent behavior of the very fast phase alignment technique based on oversampling that was previously described: the OLT managed to reach a BER of $10^{-11}$ on each of the bursts of each of the ONUs at the start of each test sequence. The spread between the best and the worst BER plots for each ONU (of a few dB) can be explained by the starting conditions of the optical modules of the ONUs (laser biasing, extinction ratio, power level) that slightly change between tests. However, even the worst plot of these hundreds of measurements still comfortably allows a split ratio of 1:128, which was the original target of this TTC-PON prototype.

![Figure 10. Upstream B-BER measurements.](image)

### 4 Summary

The objectives of the presented study were to describe a TTC upgrade proposal using bidirectional 10G-PON FTTH technology, the methods used and modifications applied to the available standards, and the performance obtained under some stress conditions, both at the physical and at the protocol layers. This new system offers major improvements over its predecessor: low and deterministic latency, high quality recovered clock as well as high capacity in the downstream path. Exploiting this last feature, dynamic software partitioning can be implemented allowing to issue “Trigger types” addressing different parts of the sub-detectors. Moreover bidirectionality is introduced, a feature absent from the current TTC, using a customized upstream protocol. This allows feedback from the sub-detectors and integration of the “busy/throttle” link (FMM or ROD) into the optical link with a maximum dead time of less than 5 µs.
References

[1] J. Varela, *Timing and synchronization in the LHC experiments*, talk given at the 6th Workshop on Electronics for LHC Experiments, Krakow Poland, 11–15 Sep 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] 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.

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

[5] P. Begovic, N. Hadziahmetovic and D. Raca, *10G EPON vs. XG-PON1 efficiency*, talk given at the 3rd International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), Budapest Hungary, 5–7 Oct 2011, http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6078887&tag=1.

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

[7] Xilinx Inc., *All digital VCXO replacement for gigabit transceiver applications*, XAPP589, v2.2 (2011), http://www.xilinx.com/support/documentation/application_notes/xapp589-VCXO.pdf.

[8] Xilinx Inc., *Kintex-7 FPGAs data sheet: DC and AC switching characteristics*, DS182, v2.11 (2014), http://www.xilinx.com/support/documentation/data_sheets/ds182_Kintex_7_Data_Sheet.pdf.