A point-to-point link for data, trigger, clock and control over copper or fibre

A. Tarazona, K. Gnanvo, S. Martoiu, H. Muller and J. Toledo

**ABSTRACT:** Upgrades of the LHC detectors target significantly higher event rates and higher bandwidth over point-to-point links. The Data, Trigger, Clock and Control (DTCC) is a new custom link protocol for data and control streams over different physical media, as copper or optical fibre. The DTCC link is implemented over 8b10b encoding. A version of the DTCC link over standard Category 6 cables is planned to be used with ALICE EMCal Calorimeters after its LS1 upgrade with a significant increase of the readout rate [1].

**KEYWORDS:** Data acquisition concepts; Control and monitor systems online; Modular electronics

1 Introduction
The Scalable Readout System (SRS) [2], developed within the RD51 collaboration, is intended as a general purpose multi-channel readout solution for a wide range of detector types and complexities. The scalable architecture is based on high-speed point-to-point links - with no buses involved, facilitating scalability without any loss of performance - in which the Data, Trigger, Clock and Control (DTCC) link is a key component. The DTCC link is FPGA-based. Therefore, the two ends that it connects must contain an FPGA.

The DTCC link hosts all digital communications between the Front-End Cards (FEC) [3, 4] and the Scalable Readout Unit (SRU) for large SRS systems (up to 82 kchannels). With digital front-ends, the DTCC link can also be used as interface between the FEC cards and the front-end electronics. Therefore, the DTCC link allows the establishment of a connection between all the elements in an SRS system. In this way, the user can manage each element in the chain from an online PC or farm, via 1, 10 GbE network ports or other specific application protocols, or even from a Passive Optical Network (PON) planned to handle trigger and slow-control frames.

2 The DTCC link
The DTCC link arose from the need to connect many front-end cards to only one card (Scalable Readout Unit) in order to gather all the data from all the detectors. This card can be seen as a network switch (see figure 1) but with added functions.
The DTCC link is a generic protocol, based on the RD51 experiments needs that can be used in any FPGA-based readout system because of its high versatility. It can be used in a variety of ways over standard Shielded Foiled Twisted Pair (SFTP) Category (Cat) 6A/7 cables, HDMI cables (in both cases with LVDS signalling) and also by means of optical fibre (using FPGA transceivers) if higher data rates and/or a longer distances are required. It carries clock, triggers, configuration/monitoring and data information between a readout unit and the front-end cards.
The main specifications that the DTCC link must comply are:

- Clock distribution (LHC clock) keeping a skew lower than 1 ns.
- Bidirectional.
- Dynamic synchronisation.
- Support different kind of frames; trigger, slow control (Ethernet frames) and data frames.
- 8b10b encoding.
- Fixed latency.
- Bit Error Ratio lower than $10^{-11}$ at 1 Gbps using a maximum cable length of 5 metres (length considered suitable to connect the front-end cards to the concentrator card, Scalable Readout Unit).

### 2.1 Implementation

The current implementations have been carried out over Ethernet cables, SFTP Cat6A/Cat7, as shown in figure 2. The Ethernet cables are based on four differential pairs. In the DTCC link, two of them are used as downstream lines and the other two as upstream lines. High-speed transmitters and receivers (SN65LVDS100 IC from Texas Instruments [5], RX and TX chips in figure 5) are used at both ends of the cables. These ICs are placed between the cable and the FPGA. They are simple standard ICs in order to reduce the cost. Therefore, they do not provide any kind of configuration to fine-tune the communication, such as pre-emphasis and de-emphasis. For this reason, the transmission quality depends on the type of cable used and its length. The choice of both the type of cable and its length is determined by the data rate of the data line (differential pair) and the desired Bit Error Ratio (BER).

As depicted in figure 2, there are two kinds of implementation over an Ethernet cable. One of them is called basic (figure 2a) and the other is called advanced (figure 2b). Both implementations are capable of distributing the clock to the downstream systems and of sending or receiving trigger, slow-control and data frames. The difference between them is that the basic one only supports cables of the same length while the advanced one supports cables of different lengths.
The two differential pairs used as downstream lines work in the same way in both implementations. Pair 1 is used to distribute the clock and pair 2 to send trigger and slow-control frames.

The difference is in the upstream pairs (in Pair 4). Pair 3 is used to receive trigger, slow-control and data frames and Pair 4 is used as an extra data line (just in case the user wants to increase the throughput) in the case of basic implementation, while in the advanced one, pair 4 is employed to reflect the clock sent through Pair 1.

One of the main characteristics of the DTCC link is its low-skew clock distribution. In the basic implementation, the low-skew is achieved by using cables of the same length, and of course, taking into account the routing delays. For the case that the temperature changes, a DTCC link can be influenced by a temperature change in a different way than another DTCC link. This effect would increase the skew. However, this is unlikely to occur since the system is always in a temperature controlled room. Furthermore, the length of the cables is too short and a temperature change would affect all of them equally.

An approach to control and ensure a low-skew other than using cables of equal lengths is to employ the advanced implementation. A dynamic synchronisation (the phase adjustment is performed every few microseconds, see section 2.3) can be achieved by means of the differential pair 4 and a synchronisation method. This method is capable of measuring the phase difference between clocks and adjusting this phase difference until attaining the condition that the two clocks oscillate in-phase. This method will be addressed in section 2.3.

2.2 Operation
The DTCC link is based on an 8b10b encoding, a line code that maps 8-bit symbols to 10-bit symbols to achieve DC-balance and bounded disparity, and yet provide enough state changes to allow reasonable clock recovery. This means that the difference between the count of 1 s and 0 s in a string of at least 20 bits is no more than 2, and that there are not more than five 1 s or 0 s in a row.

This helps to reduce the demand for the lower bandwidth limit of the channel necessary to transfer the signal [6]. Another advantage of using this encoding is its control (or K) symbols. They do not have a corresponding 8b data byte and are used for low-level control functions, widely used in Fibre Channel or Gigabit Ethernet, for instance. In our case, these special or control symbols are used to distinguish among the different types of frames. As commented, trigger, slow-control and data frames can be sent (or received) by means of a single differential Pair in the DTCC link. That is, the control symbols mark each frame that is being sent (or received).

A simplified block diagram showing how the DTCC link is implemented in the FPGA is illustrated in figure 3. Pairs 2 and 3 are able to carry different types of frames, trigger, slow control and data frames. Each one of these pairs is divided into two channels, one is high priority and the other is low priority. Both channels are time multiplexed using the control symbols provided by the 8b10b encoding. The high-priority channel can interrupt any ongoing data transmission, which is later resumed after completion of the high-priority transmission. In this way, the high-priority channel provides the fixed latency necessary for timing information (like trigger and flow control).

The low-priority channel is sub-partitioned between DAQ data frames (data from the detectors) and Ethernet transmission for slow-control functions. Idle frames are used so that
the link remains synchronised when no frames are being sent. As commented on previous paragraphs, Pair 1 is used to provide the main clock to the slave cards and Pair 4 is used as an extra data line in the case of basic implementation, while in the advanced one, Pair 4 is employed to reflect the clock sent through Pair 1 (see figure 2 and figure 3). An automatic alignment process is carried out when the system is started or when the user decides to do it. Thus, the data are sampled in the optimal instant of time, in the centre of the eye diagram.

![Figure 3. DTCC implementation in the FPGA for basic implementation.](image)

Figure 3. DTCC implementation in the FPGA for basic implementation.

![Figure 4. Clock synchronisation. (a) Constant phase with respect to the master clock and (b) all the clocks in-phase.](image)

Figure 4. Clock synchronisation. (a) Constant phase with respect to the master clock and (b) all the clocks in-phase.

2.3 Synchronisation

The purpose of the synchronisation is to keep all the front-end cards connected to the readout unit within a certain clock skew, as shown in figure 4.

By using the basic implementation it is only possible to achieve a constant phase with respect to the master clock, as represented in figure 4a, as the delay is determined by the length of the cable. However, if the advanced one is used the phase of all the slave clocks can be adjusted so that all the clocks, both master and slaves, oscillate in-phase, as depicted in figure 4b.

The Digital Dual-Mixer Time Difference (DDMTD) method [7] is used for evaluating the phase differences. This is an all-digital version of the well-known Dual-Mixer Time Difference (DMTD) technique, where measured clocks are mixed with a tone of a similar frequency and filtered in order to obtain lower frequency signals with the same phase relationship. This method is
intended to stretch the clocks whose phase difference we would like to obtain in time by a certain factor in order to facilitate the measurement (further details in [7]). Thus, the phase difference can be estimated in a precise way in order to adjust the phase of the slave clock.

Figure 5. (downward way).

How can the phase difference be estimated in our system using the DDTMD method? If the scheme illustrated in figure 5 is analysed the following equations can be obtained: downward way:

$$\text{phase}_S = \text{phase}_M + d_{\text{PLLM}} \times \text{FPGA}_{\text{pin}} + \text{cable} + d_{\text{FPGA}} \times \text{DDMTD}_S \quad (2.1)$$

upward way:

$$\text{phase}_S + d_{\text{PLLS}} \times \text{FPGA}_{\text{pin}} + \text{cable} + d_{\text{FPGA}} \times \text{DDMTD} + \text{DDMTD} = \text{phase}_M + d_{\text{PLLM}} \times \text{DDMTD} \quad (2.2)$$

The delay introduced from the output FPGA to the RJ-45 connector (or from the RJ45 connector to the FPGA input) can be considered as part of the Ethernet cable. The lengths of the lines from FPGA to RJ-45 and from RJ-45 to FPGA are equal. The cable delays are equal in both directions because they (downstream and upstream lines) are differential pairs that make up the Ethernet cable. The delays represented in equation (2.1) and equation (2.2) are fixed once the FPGA logic has been placed by the synthesis tool. They can be obtained using the timing tools provided by the FPGA manufacturer.

Taking into account the following condition:

$$\text{phase}_M \times \text{phase}_S = 0 \quad (2.3)$$

We can obtain the necessary delay to be applied to the slave so that the two clocks oscillate in-phase. This process can be repeated as and when desired. Thus, the skew will always remain around 200 ps regardless of the length of the cables and the temperature changes.

If the DDMTD method is combined with the Precision Time Protocol (PTP) [9] the total delay introduced by each differential pair can be known dynamically [7] (not covered in this manuscript). In this case, the PTP is also used to solve a phase ambiguity produced from employing the DDMTD method.

2.4 Latency and skew

The link latency and skew is a key factor, especially for trigger and timing information. The less latency the system has, the better. In our system, the latency is defined as the passed time since a trigger signal is received in the master card and this is distributed to all the slave cards.
and finally it is received in each of the slave cards. The latency is given by the serialiser and deserialiser blocks, synchronisation block, PCB layout, connectors and length of the cable. Therefore, the highest latency of the system is obtained when 5 metres cables are used and obviously at low data rates (see table 1).

<table>
<thead>
<tr>
<th></th>
<th>2 metres cable</th>
<th>3 metres cable</th>
<th>5 metres cable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Avg</td>
<td>RMS</td>
<td>Avg</td>
</tr>
<tr>
<td>0.8 Gbps</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slave card 1</td>
<td>216.59</td>
<td>216.59</td>
<td>221.11</td>
</tr>
<tr>
<td>Slave card 2</td>
<td>216.52</td>
<td>216.52</td>
<td>221.18</td>
</tr>
<tr>
<td>Slave card 3</td>
<td>216.60</td>
<td>216.60</td>
<td>221.25</td>
</tr>
<tr>
<td>1.0 Gbps</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slave card 1</td>
<td>191.80</td>
<td>191.80</td>
<td>196.38</td>
</tr>
<tr>
<td>Slave card 2</td>
<td>191.73</td>
<td>191.73</td>
<td>196.43</td>
</tr>
<tr>
<td>Slave card 3</td>
<td>191.72</td>
<td>191.72</td>
<td>196.40</td>
</tr>
<tr>
<td>1.2 Gbps</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slave card 1</td>
<td>166.04</td>
<td>166.04</td>
<td>170.65</td>
</tr>
<tr>
<td>Slave card 2</td>
<td>166.11</td>
<td>166.11</td>
<td>170.79</td>
</tr>
<tr>
<td>Slave card 3</td>
<td>166.04</td>
<td>166.04</td>
<td>170.70</td>
</tr>
</tbody>
</table>

Table 1. Average (Avg) and root mean square (RMS) link latency (ns) using different data rates and length of cables (SFTP Cat6A). Test carried out with one master card and three slave cards.

The skew is the maximum time difference among trigger lines in the slave side. For example, in case that a master card distributes a trigger signal to three slave cards, the trigger signal will reach each slave card at a different instant of time. According to the link specification given in section 2, the maximum time difference among arrival times of the trigger signal have to be of 1 ns. The skew mainly depends on type of cable and on its length. It also depends on temperature changes. However, as noted in the section 2.1, this is unlikely to occur since the system is always in a temperature controlled room. To check the system, a temperature change was carried out by decreasing the room temperature 10 ºC. The skew variation was around 20 ps, which is a negligible amount to take into account.

Several measurements were performed to obtain the skew of system using both implementations, basic and advanced. The maximum measured skew was around 200 ps (see table 1).

2.5 Bit Error Ratio (BER) tests

In digital transmission, the transmitter, the receiver or the transmission media can cause bit errors. The BER estimates how well the digital transmission is performed. The BER of a digital communication system can be defined as the estimated probability that any bit transmitted through the system will be received in error. BER performance is limited by random noise and/or random jitter.

The quality of the BER estimation increases as the total number of transmitted bits increases. In the limit, as the number of transmitted bits approaches infinity, the BER becomes a perfect estimate of the true error probability [8]. However, it is impossible to transmit an infinite number of bits to reach an exceptional BER estimation in real BER testing, because the test
time would be infinite. A way to carry out BER estimation transmitting a finite number of bits would be to apply the following equation [8], based on the concept of statistical confidence levels.

\[
N = \frac{1}{\text{BER}} \left[ -\ln(1 - CL) + \ln \left( \sum_{k=0}^{E} \frac{(N \cdot \text{BER})^k}{k!} \right) \right]
\]

where \( N \) represents number of transmitted bits, \( E \) is the total number of errors detected and \( CL \) represents the confidence level (0 to 100 %).

Pseudorandom patterns were employed to estimate the robustness of the link. A test pattern emulates the type that is expected to occur during normal operation or stressful operation. These patterns are called pseudorandom bit sequences (PRBS) and they are classified by the length of the pattern. It is possible to obtain different results using dissimilar patterns. Longer patterns are more stressful that shorter ones. Therefore, different PRBS have been used to test the robustness of our system, in particular, PRBS7 (pattern length = \(2^7-1\)), PRBS15 (\(2^{15}-1\)), PRBS23 (\(2^{23}-1\)) and PRBS29 (\(2^{29}-1\)). Different data rates and lengths of SFTP Cat6A cable have also been used.

All tests were successful which were are based on equation (2.4) being \( CL = 0.99 \) and \( \text{BER}=10^{-12} \). This BER value was chosen because it meets the specifications (\(\text{BER} = 10^{-11}\)) and the necessary time to perform the BER testing is affordable. However, better BER values can be met, for example, \(\text{BER} = 10^{-13}\) using a SFTP Cat6A cable at 1 Gbps.

The obtained results validated the operation of the link since they meet the given specifications. It should be noted however that due to an excessive jitter the alignment process sometimes does not find the optimal sampling point which happens when a data rate of 1.2 Gbps and a cable length of 5 metres are used. Therefore, the BER target is not achieved. This issue will be addressed in section 3.

2.6 Transmission media

Throughout this paper the DTCC link is implied to work over Ethernet cables. The reasons for this are cost reduction and suitability for our system. However, another very useful approach would be possible, to use the DTCC link over optical fibre. Nowadays, the FPGA includes a lot of high-speed transceivers at an affordable price. Furthermore, by using optical fibre, higher data rate (limited by the internal transceivers of the FPGA) and longer lengths can be achieved. Therefore, this approach would be the optimal for new designs. In figure 6 a fixed-latency optical fibre approach is represented, based on [7, 10–12] and in the sections above.

The main difference between using optical fibre or Ethernet cables for the DTCC links is that there are only two available paths, up or down. Then, the clock needs to be embedded in the data and then recovered. However, the rest of logic can be reused. Both the upstream and downstream paths are split into two channels, low and high-priority. In this way, this approach achieves the same functionality as the copper approach, by being able to send/receive trigger, slow control and data frames. The DDMTD algorithm and Precision Time Protocol are included as well.

Thus, the temperature changes can be compensated in order to keep all the clocks synchronised, which is critical when long fibres are used. In addition, the latency can be known
at any time with exceptional accuracy, which can be very useful when the packets need to be marked with a timestamp.

![Diagram showing a scheme to estimate the phase difference.](image)

**Figure 6. Scheme to estimate the phase difference.**

### 3 Jitter measurements

Jitter measurements are very important in high-speed serial data systems since they allow the characterization of the performance of the system. Excessive jitter can increase the BER of a communication system and also lead to a violation of timing margins, causing circuits to behave improperly.

The jitter measurements were taken by using a high-speed oscilloscope (RTO1044 from Rohde & Schwarz [13]) and a stringent PRBS, in particular, a PRBS29. All the lines, clock, trigger and data, were evaluated in both master and slave card, in order to check the degradation of the signal.

The most important line is the clock line, since the rest of lines depend on it. Thus, the less jitter the clock line has, the better the behaviour of the rest of the lines. The following tables represent the obtained jitter in each line, clock line, trigger line and data line. The shown results are valid both for the basic and the advanced implementation.

According to the tables, table 2 to table 4, SFTP Cat7 cables present better jitter figures than SFTP Cat6A cables although the differences between them are very small. It was expected due to the higher bandwidth that SFTP Cat7 cables support. Better jitter figures are also observed in the trigger lines than in the data lines. This can clearly be seen in figure 7, the eye diagram corresponding to the data line is closer than the corresponding to the trigger line. This is due to the source clock being taken from the clock line in the receiver (slave cards) without being regenerated.

Therefore, a considerable amount of jitter is added to the clock signal, since the transmission media is the main source of jitter. A way of cleaning this clock would be to use a jitter cleaner.
chip. If this is done, the performance of the trigger and data lines should be similar and solve the issue commented in 2.5.

The measured jitter, which is about what we expected, allows us to meet the BER design goal for short cables (less than or equal to 5 metres).

(a) Period jitter.

<table>
<thead>
<tr>
<th>Cable</th>
<th>Master Card</th>
<th>Slave Card</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stats</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\mu$ (ns)</td>
<td>$\sigma$ (ps)</td>
</tr>
<tr>
<td></td>
<td>RMS (ns)</td>
<td></td>
</tr>
<tr>
<td>SFTP Cat6A</td>
<td>24.999</td>
<td>13</td>
</tr>
<tr>
<td>SFTP Cat7</td>
<td>25.000</td>
<td>20</td>
</tr>
</tbody>
</table>

(b) Cycle-to-cycle jitter.

<table>
<thead>
<tr>
<th>Cable</th>
<th>Master Card</th>
<th>Slave Card</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stats</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\mu$ (fs)</td>
<td>$\sigma$ (ps)</td>
</tr>
<tr>
<td></td>
<td>RMS (ns)</td>
<td></td>
</tr>
<tr>
<td>SFTP Cat6A</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>SFTP Cat7</td>
<td>0</td>
<td>29</td>
</tr>
</tbody>
</table>

Table 2. Jitter in the clock line. This line carries a 40MHz clock. Average ($\mu$), root mean square (RMS) and standard deviation ($\sigma$).

(a) SFTP Cat6A cable.

<table>
<thead>
<tr>
<th>Stats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jitter p-p (ps)</td>
</tr>
<tr>
<td>Jitter 6$\sigma$ (ps)</td>
</tr>
<tr>
<td>RMS (ps)</td>
</tr>
</tbody>
</table>

(b) SFTP Cat7 cable.

<table>
<thead>
<tr>
<th>Stats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jitter p-p (ps)</td>
</tr>
<tr>
<td>Jitter 6$\sigma$ (ps)</td>
</tr>
<tr>
<td>RMS (ps)</td>
</tr>
</tbody>
</table>

Table 3. Jitter in trigger line (measured in reception, slave card), line carrying PRBS29 data at 1 Gbps.

(a) SFTP Cat6A cable.

<table>
<thead>
<tr>
<th>Stats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jitter p-p (ps)</td>
</tr>
<tr>
<td>Jitter 6$\sigma$ (ps)</td>
</tr>
<tr>
<td>RMS (ps)</td>
</tr>
</tbody>
</table>

(b) SFTP Cat7 cable.

<table>
<thead>
<tr>
<th>Stats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jitter p-p (ps)</td>
</tr>
<tr>
<td>Jitter 6$\sigma$ (ps)</td>
</tr>
<tr>
<td>RMS (ps)</td>
</tr>
</tbody>
</table>

Table 4. Jitter in data line (measured in reception, master card), line carrying PRBS29 data at 1 Gbps.
4 Functional verification

During August 2013, in the RD51 facilities at CERN, several tests were carried out with data from 8k gas detector channels distributed over 8 FEC cards. FECs were connected via DTCC links to the SRU readout unit, which acted as a switch and sent data upstream via 1 GbE (see figure 1a). The readout clock and trigger were sent to all the FEC cards in order to synchronise the transmission of data to the online system. As DAQ system, the ALICE DATE online system was used.

We were able to configure the 8 FEC cards and the readout chips connected to them through DTCC links, by using slow control frames (Ethernet frames). Once the whole system was configured, the acquisition was started. The trigger signal was distributed by sending trigger frames from the SRU readout unit to the all FEC cards. A total of 800,000 events were sent from the SRU readout unit to the DAQ system via 1 GbE. The trigger rate was close to 600 Hz.

Tests were successful as the SRU readout unit was able to gather the data from all the FEC cards and then to send them to the DAQ with no error. In addition, the acquisition rate was very encouraging (lower than 100 Hz in typical SRS applications).

These tests were a great step forward since the DTCC link allows us to build large readout systems.

5 Conclusions

A fixed-latency link capable of carrying Data, Trigger, Clock and Control information has been described and demonstrated. The link is based on an 8b10b encoding and runs over copper although an optical fibre approach has been proposed. One of the leading characteristics of this link is its low-skew clock distribution, around 200 ps. The data rate of link is 1 Gbps using 5 metres SFTP Cat6A/Cat7 cables of length with a BER1 lower than 10^{-12}. SFTP Cat7 cables present a slightly better performance than SFTP Cat6A cables as the jitter levels are lower.

The developed implementations satisfy the design requirements. However, a small issue has been found when the link works close to the rate limit (1.25 Gbps, max speed of the LVDS I/Os in a Virtex 6 FPGA) and the length of the cable is of 5 metres. This issue is due to an excessive jitter in the data line, since the source clock is taken from the clock line in the receiver (slave cards).

6 Future work

The jitter issue commented in section 3 has been solved in the new designs by adding a jitter cleaner chip in the receiver, which will be tested in the coming months.
The optical fibre approach proposes in section 2.6 has not been tried yet due to the fact that our current hardware does not have the appropriate components for it. Therefore, the DTCC link over optical fibre will be performed in the coming designs.

As the tests were successful the next steps are to:

1. use a 10 GbE connection between the SRU readout unit and the DAQ system.
2. connect more FEC cards to the SRU readout unit and reach an acquisition rate of 1 kHz.

References


