

ETR 2010-01 May 31, 2010

Measuring propagation delay over a 1.25 Gbps bidirectional data link

P.P.M. Jansweijer {<u>peterj@nikhef.nl</u>}, H.Z. Peek {<u>henkp@nikhef.nl</u>}

Abstract

Measurement and control applications are increasingly using distributed system technologies. In such applications, which may be spread over large areas, it is often necessary to synchronize system timing and know with high precision the time offsets between parts of the system. In this note a method is described to determine the time offset between two nodes. These nodes are connected via an 8B/10B coded 1.25 Gbps bidirectional serial point to point communication channel, as for example is used by 1000BASE-X (Gigabit Ethernet).

The time offset is determined by measuring propagation delay using a marker signal. The signal is sent from a master to a slave node and back using serializer/deserializer (SerDes) functionality in (Virtex-5) FPGAs. The recovered clock at the slave node is used as the transmit clock of the slave so the complete system is synchronous. For a 1.25 Gbps serial communication channel the delay is known with a resolution of a single unit interval (i.e. 800 ps). This resolution can be further enhanced by measuring the phase relation between the transmit and receive clock of the master node.

The technique has been demonstrated to work over a single 10 km fibre that is used at two wavelengths, to facilitate a bidirectional point to point connection between master and slave node.

Nikhef, Department of Electronic Technology Science park 105 1098 XG Amsterdam, NL

Contents

1	Intr	oduction				
	1.1	General				
	1.2	Carrier synchronous Ethernet versus synchronous Ethernet				
2	Exp	perimental setup				
3	Tes	t setup in more detail				
	3.1	General				
	3.2	Overview				
	3.3	The 1.25 Gbps test setup				
	3.4	Complications due to dividing the 2.5 GHz shared PMA PLL centre frequency. 7				
	3.5	Closing the loop at the slave node using a digital PLL9				
4	4 Resets					
	4.1	Overview				
	4.2	Master node initiates a reset sequence				
	4.3	Slave node initiates a reset sequence				
5 Conclusions		clusions				
	5.1	General				
	5.2	Further work				
Appendix A						
	A-1	Dispersion and fibre choice				
	A-2	Compensate time offset for dispersion				
	A-2.1	General calculation				
	A-2.2	Measured and calculated values for the dispersion				
	A-2.3	Example calculation for time offset				
R	References					

1 Introduction

1.1 General

Measurement and control applications are increasingly using distributed system technologies. In such applications, which may be spread over large areas, it is often necessary to synchronize system timing and know with great precision the time offsets between parts of the system. An example of such an application is the KM3NeT telescope [1] to be built in the Mediterranean sea. In the framework of the design study for this telescope a technique was developed to measure propagation delay over a coded serial communication channel using FPGAs [2]. A first test setup was built to verify the principle of measuring propagation delay between a transmitter and a receiver using a coded serial communication channel operated at 3.125 Gbps. The transmitter and receiver reside in FPGAs on two separate development boards. This first test setup showed that it is feasible to measure propagation delay over a 100 km fibre with a resolution of one unit interval (i.e. 320 ps at 3.125 Gbps).

Another measurement and control application is being developed at CERN. Timing and computer networking needs to be distributed with high precision along the accelerator complex. A project was initiated to develop a new standard. This project is called White Rabbit [3] and it aims to develop a fully deterministic Ethernet-based field bus for general purpose data transfer and synchronization. It should be able to synchronize ~1000 nodes with subnanosecond accuracy over lengths of up to 10 km. The key technologies used are carrier synchronous Ethernet and precision timing protocol (PTP: IEEE 1588) [4].

Both applications (KM3NeT, White Rabbit) and also other measurement and control applications can profit from the results of the design study to measure propagation delay between a transmitter and a receiver. Therefore it was decided to build a second test setup that is able to measure the time offset between a master and a slave node using a bidirectional communication link based on the bit rate of 1000BASE-X Gigabit Ethernet [5] (1.25 Gbps).

1.2 Carrier synchronous Ethernet versus synchronous Ethernet

Synchronous Ethernet [6] uses the physical layer interface to pass timing from node to node in the same way timing is passed in SONET/SDH or T1/E1. Network synchronization in telecom systems is based on clock hierarchy. For reliability reasons, it is unrealistic to expect all global telecommunication networks to be synchronized to a single primary reference clock (PRC). Real networks use a flatter timing distribution structure with a number of PRCs running independently, creating plesiochronous (i.e. nearly synchronous) timing islands within the network.

To achieve subnanosecond timing accuracy only a single timing island must be used with a single PRC as timing reference. The lower levels of the network hierarchy are carrier synchronous with the PRC. Hence the term carrier synchronous Ethernet is used.

2 Experimental setup

The test setup consists of two ML507 Xilinx development boards [7]. A Virtex-5 FPGA [8] is mounted on each board. One ML507 development board is designated as master node, the

other as slave node. Master and slave are connected via small form factor pluggable (SFP) transceivers and 10 km of fibre, creating a bidirectional link. A single fibre is used that is operated at dual wavelength. The slave node receiver recovered word-clock is used as the slave transmitter word-clock. Therefore the word-clock, recovered by the master receiver, is locked via the slave node, to the master transmitter word-clock (Figure 1). This is often referred to as carrier synchronous Ethernet.



Figure 1: Test setup overview

The test setup hardware facilitates fine time measurement which can be used for implementing the PTP protocol IEEE 1588 [4]. However it does not include actual time stamping and the software for PTP.

Due to the FPGA SerDes architecture, the serial communication is based on 16 bit words at the word-clock speed which are coded into 20 bit ordered sets. The coding for these ordered sets is based on Gigabit Ethernet definitions [5]. The transmitter serialization architecture contains a PLL that multiplies the word-clock to the serial bit rate (i.e. bit clock). Initially the receiver is using a local free running word-clock which is locked onto the transmitter word-clock after synchronisation onto the incoming bit stream. This lock occurs at a random bit position so the receiver recovered word-clock will have a phase offset with respect to the transmitter wordclock. However, this phase is known since the transmitter sends special code-groups during synchronization from which the receiver can recognize the word boundaries. The de-serialization architecture in the receiver aligns the data with a barrel shifter. When word synchronization is reached then the phase relation between the transmitter word-clock, the communication channel and the receiver recovered word-clock is determined by the number of barrel shifts (BitSlide) needed to properly align the receiver.

The time offset between master and slave node can be measured by sending a timing marker signal forth and back. A special code-group is used as a timing marker signal (see Figure 2). A "Start" pulse is generated when the master node sends a marker signal, which initiates a time-interval measurement. A "Stop" pulse is generated when the receiver in the master node has decoded the marker signal. An oscilloscope is used to measure the time between "Start" and "Stop" pulse. Total propagation delay fine time measurements done by the FPGAs are verified against the time interval measurements done with the oscilloscope.

Data	K28.5	D16.2	K23.7 Carrier Extend	Dx.Y BitSlide	K28.5	D16.2
	IDLE		Timing marker		IDLE	

Figure 2: 8B/10B coded timing marker signal

As described above, the phase relation between the master transmitter word-clock and the slave receiver recovered word-clock is determined by the "BitSlide" value. This value needs to be known in order to calculate the propagation delay forth and back, therefore it is added to the "BitSlide" value contained in the timing marker signal.

Since both the forward and backward path follow the same physical fibre, the time offset between master and slave is in principle half the propagation delay measured. However, the propagation speed in the single fibre is slightly different for different wavelengths (dispersion). This is a known small fixed factor for which can be corrected (see Appendix A).

Figure 3 shows the test setup with the master node connected to the slave node via 10 km of fibre.



Figure 3: Test setup

3 Test setup in more detail

3.1 General

In the following text and figures the bidirectional fibre link between master and slave node (see Figure 1) is split in two separate data paths. Figure 4 shows the test setup overview in more detail. The ML507 boards are shown as gray rectangles. On each board a Xilinx Virtex-5 FPGA (XC5VFX70TFFG1136) is mounted, which is shown as a blue rectangle. These FPGAs contain SerDes architectures which are called "GTX" (the yellow rectangles) in Xilinx nomenclature. A complete description of the GTX can be found in the "Virtex-5 FPGA RocketIO GTX Transceiver User Guide" [9]. Since the overview in Figure 4 is rather complex the chapters and figures below will zoom in onto sub parts where more details are explained.

3.2 Overview

Figure 4 shows the components, necessary for measuring propagation delay over a 8B/10B coded serial communication channel. The barrel shift process used to determine the propagation delay from transmitter to receiver is described in detail in reference [2]. The signals in red trace the clock starting near the master node, to the slave node and back to the master node again. The path of the clock signal is described below.



Figure 4: Clock network

The system reference clock originates from a crystal oscillator that is located on the ML507 development board (designated to be the "master node") next to the Virtex-5 FPGA.

This clock is used as the primary reference clock for the serializer that is located in the Xilinx SerDes architecture. This clock is connected via two dedicated differential clock FPGA input pins. The reference clock is multiplied to the serial bit rate by the shared physical medium attachment (PMA) PLL. The bit rate is 20 times the word-clock rate and is fed to the parallel input serial output (PISO). The serializer sends the 8B/10B coded bit stream to the slave node where the clock data recovery (CDR) circuit in the receiver of the slave is extracting the receiver recovered word-clock from the incoming bit stream (RxUsrClk).

The backward clock path from slave node to master node is basically the same as the forward path. The reference clock for the transmitter in the slave node must be locked to the slaves receiver recovered word-clock. This is done by implementing a digital phase locked loop (DPLL) that drives the control voltage of a voltage controlled crystal oscillator (VCXO). The clock path continues from slave node to master node via the slaves shared PMA PLL and PISO, to the master receiver CDR and finally the master RxUsrClk.

The total propagation delay from master to slave and back is determined by:

- 1. Number of word-clocks
- 2. Number of barrel shifts necessary for word synchronization (bit clocks)
- 3. The phase between master transmitter word-clock (TxUsrClk) and the master receiver recovered word-clock (RxUsrClk).

When the bidirectional link is up and synchronized then the first two items have a constant delay expressed in bit clocks. The third item varies in time due parameters that influence the propagation delay. The main parameter is the temperature of the fibre. Another parameter that has an influence on the propagation delay is the pressure on the fibre when it is, for example, part of an undersea cable. The third item needs to be measured constantly to track the total delay. The total delay resolution can be much higher than the bit clock resolution.

3.3 The 1.25 Gbps test setup

Ironically, implementing the relative low bit rate of 1.25 Gbps is much more complicated than the implementation needed for a higher bit rate. In fact Figure 4 shows the test setup as it would appear for a 3.125 Gbps network. Figure 5 shows the test setup which is based on a 1.25 Gbps serial communication link.



Figure 5: Clock network for 1.25 Gbps

For a bit rate of 1.25 Gbps the shared PMA PLL needs to run on 1.25 GHz. However, the nominal operating range for the Virtex-5 shared PMA PLL is 1.5 to 3.25 GHz ([9], figure 5-1) which means that the desired operating frequency is out of range. Therefore the reference clock for the shared PMA PLL is set to twice the desired frequency (2.5 GHz) by feeding it with 125 MHz instead of 62.5 MHz. The next section describes how this complicates the architecture dramatically. Figure 5 shows a number of extra dividers ("/2", "PLL/2") and phase align circuits with respect to Figure 4.

3.4 Complications due to dividing the 2.5 GHz shared PMA PLL centre frequency

The serializer architecture within the GTX is fed with three clocks that all must originate from the same clock source (see Figure 6). The "Virtex-5 FPGA RocketIO GTX Transceiver User Guide" ([9]; figure 6-18: "Clock domains and alignment logic") may be helpful to understand the architecture.



Figure 6: Double frequency shared PMA PLL

TxUsrClk drives the physical coding sub-layer clock (PCS). This clock captures 16 bits of data from the FPGA fabric which are coded into two 10 bit code-groups that together form a 20 bit ordered set. The physical medium attachment (PMA) sub-layer clock feeds the 20 bit ordered sets to the serializer (PISO). Finally the serializer clock runs on a 20 times higher rate.

The clock that is used to serialize the bits, as well as the PMA clock is driven by the shared PMA PLL which is locked to the reference clock. The PCS clock, although indirectly, is derived from the same reference clock source. As explained in chapter 3.3 the shared PMA PLL runs on 2.5 GHz and a clock divider ("/2") is needed to create the 1.25 Gbps clock for serializing the bits. Yet another clock divider ("PLL/2") is needed to create the 62.5 MHz PMA clock. Because the outputs of these clock dividers start up randomly and are therefore not synchronized, there is no fixed phase relationship between the PCS and the PMA clock. Figure 7 shows that there are two phase relations that can exist between the PCS and PMA clock. This means that the latency between the transfers from the FPGA fabric to the actual bits on the serial output has one of two values, that differ 8 ns depending on which phase relation became active after power-up or reset.



Figure 7: Two phase relations that can exist between PCS and PMA clock

Xilinx provides a way to overcome this behaviour by implementing a "Tx phase alignment circuit" [9]. When this circuit is activated then the PMA clock is aligned with TxUsrClk restoring the fixed phase relation between the transfers from the FPGA fabric to the actual bits on the serial output. A state machine supplies the necessary control signals for the phase alignment circuit after the PLLs in the system acquired lock.

3.5 Closing the loop at the slave node using a digital PLL

The slave node receiver recovered word-clock that is the source of RxUsrClk is used as the slave transmitter word-clock (TxUsrClk), such that the master receiver recovered word-clock is locked, via the slave node, to the master transmitter word-clock (Figure 1).

The most obvious feedback is to directly couple RxUsrClk to TxUsrClk as is shown in Figure 8. The figure also shows that the clock for the serializer (PISO) originates from the reference clock input (follow the blue dashed lines). The 125 MHz crystal oscillator and TxUsrClk are two different clock domains, however the clocks must originate from a single source otherwise buffer over- and under-flows occur.



Figure 8: The slave node TxUsrClk directly fed by RxUsrClk

In order to lock the phase of the transmitted bit stream to the received bit stream of the slave node, there must be a phase lock between the RxUsrClk and the transmitter reference clock. The next logical step is to multiply the RxUsrClk by 2 and use this as the reference clock (see Figure 9). Unfortunately this creates a deadlock. Initially the clock from the shared PMA PLL is used to set the centre frequency of the PLL in the CDR circuit at the nominal bit rate of the serial data stream. When a stable serial bit stream is received the phase detector input of the PLL in the CDR circuit switches from the local free running clock to the edges of the received bit stream. So without the shared PMA PLL clock there is no stable RxUsrClk.



Figure 9: Another way to generate the slave node TxUsrClk

To avoid this deadlock there should first be a stable reference clock in order to lock onto the incoming bit stream. This can be accomplished with the circuit as shown in Figure 10 where a VCXO is set at its centre frequency and is used to supply a stable 125 MHz clock to the reference clock. First the VCXO is used as a reference clock for the CDR in the receiver. Once the CDR is locked on the incoming bit stream a stable receiver recovered word-clock (RxUsrClk) is generated. After obtaining this lock the CDR does no longer use the VCXO, which then is used to drive the TxUsrClk that is phase locked onto the RxUsrClk by closing the loop.



Figure 10: The solution. Use of a VCXO that is locked to RxUsrClk

The phase comparator, loop filter and DAC controller are based on Xilinx Application note Xapp854 [10] although a different kind of phase detector is used and some adjustments are made in the filter coefficients. These blocks reside in the FPGA while the actual DAC and VCXO reside on a small board outside the FPGA (see Figure 11). The DPLL circuit needs to be optimized further but it was possible to prove the feasibility of measuring propagation delay between a master and a slave node with it.



Figure 11: The DAC and 125 MHz VCXO are external components

4 Resets

4.1 Overview

There are two scenarios to be considered with respect to resetting the system. In the first scenario the master node is reset and the slave node acts accordingly. The second scenario occurs when the slave node is reset and the master node must take appropriate action.

4.2 Master node initiates a reset sequence

Figure 12 shows the scenario where a reset is initiated on the master node. The numbers in the figure show in what order the reset sequence takes place:

- 1. The reset signal for the master is being asserted. In the test setup this is done by a push button.
- 2. After the reset signal is de-asserted, the master node shared PMA PLL is locking onto the reference clock and the internal architecture of the GTX is being reset. When these actions are completed then the signals "Shared PMA PLL Locked" and "ResetDone" get asserted respectively.

3. In the mean time the external "PLL/2" of the master node also had a chance to lock onto the reference signal. When the lock signal of the "PLL/2" is asserted then a state machine that drives the phase align circuit is activated such that the phase of the PCM clock and the PMA clock of the master node transmitter are aligned (see also Figure 7). Now a stable stream of IDLE code-groups is sent to the receiver in the slave node.



Figure 12: Reset behavior and PLL locks when the master node initiates a reset sequence

- 4. Due to the reset, the transmitter in the master node interrupted the transmission of the bit stream. As a result, the CDR in the receiver of the slave node lost lock. This loss of lock of the CDR forces the bit slide state machine (see Figure 13) to assert a CDR reset pulse to make sure that the CDR will obtain a fresh lock onto the stable bit stream that is being sent from the master node transmitter. The slave node transmitter will be forced to transmit IDLE code-groups until valid packets are received from the master node.
- 5. After the CDR of the receiver in the slave node obtained a lock and valid comma characters are recognized then the bit slide state machine starts searching for the proper word alignment, using a barrel shifter. This eventually leads to the "BitSlide" value for the slave.
- 6. Since the CDR of the receiver of the slave node obtained a lock onto the incoming bit stream the slave node receiver recovered word-clock is now also present. Next, the DPLL containing the VCXO will be locked onto this receiver recovered word-clock.
- 7. When a DPLL lock is established and both the shared PMA PLL and "PLL/2" in the slave node show a locked status then the state machine that drives the phase align circuit of the transmitter in the slave node is activated. This causes the phase of the PCM clock





Figure 13: Bit slide state machine

- 8. The CDR in the receiver of the master node lost lock since the transmitter in the slave node has been interrupting its bit stream. When a loss of lock occurs then the bit slide state machine asserts a CDR reset pulse to make sure that the CDR will obtain a fresh lock onto the stable bit stream that is now being sent from the slave node transmitter.
- 9. After the CDR in the receiver of the master node obtained a lock and valid comma characters are recognized then the bit slide state machine starts searching for the proper word alignment, using a barrel shifter. This eventually leads to the "BitSlide" value for the master. The number of barrel shifts needed by the slave node to obtain word synchronisation is encoded in the timing maker signal (see Figure 2) so the master node is able to calculate the total amount of barrel shifts (bit clocks) needed for both the slave node and the master node to obtain word synchronisation.

From this point the bidirectional link is up and synchronized. The master node sends a data packet and this is a sign to the slave that the reset sequence has completed and it may send data packets to the master node. The total propagation delay from master to slave and back is determined as described in chapter 3.

4.3 Slave node initiates a reset sequence

Figure 14 shows the scenario where a reset is initiated on the slave node. The numbers in the figure show in what order the reset sequence takes place:

1. The reset signal for the slave node is being asserted. In the test setup this is done by a push button.



Figure 14: Reset behaviour and PLL locks when the slave node initiates a reset sequence

- 2. After the reset signal is de-asserted, the slave node shared PMA PLL is locking onto the reference clock and the internal architecture of the GTX is being reset. As a consequence the slave node GTX transmitter stops sending which causes the master node receiver to lose lock. When this happens the master node transmitter is forced to send IDLE code-groups continuously. The receiver of the slave node asserts "Shared PMA PLL Locked" and "ResetDone" respectively. Note that the reference frequency for the slave node is the centre frequency of the VCXO that is part of the DPLL which is not in a lock state yet.
- 3. The CDR in the receiver of the slave node has been reset so the CDR will obtain a fresh lock onto the stable bit stream that is being sent from the master node transmitter.
- 4. After the CDR in the receiver of the slave node obtained a lock and valid comma characters are recognized then the bit slide state machine (see Figure 13) starts searching for the proper word alignment, using a barrel shifter. This eventually leads to the "BitSlide" value for the slave.
- 5. Since the CDR in the receiver of the slave node obtained a lock on the incoming bit stream there is now also a stable receiver recovered word-clock. The VCXO will now be

controlled such that a lock occurs onto the receiver recovered word-clock. When the lock signals of the "PLL/2", the "shared PMA PLL" and the DPLL are all asserted then a state machine that drives the phase align circuit is activated. This causes the phase of the PCM clock and the PMA clock to be aligned (see also Figure 7). Now a stable stream of IDLE code-groups is sent to the receiver in the master node.

- 6. The transmitter in the slave node interrupted the transmission of the bit stream. As a result, the CDR in the receiver of the master node lost lock. When a loss of lock occurs then the bit slide state machine in the master node asserts a CDR reset pulse to make sure that the CDR will obtain a fresh lock onto the stable bit stream that is now being sent from the slave node transmitter.
- 7. After the CDR in the receiver of the master node obtained a lock and valid comma characters are recognized then the bit slide state machine starts searching for the proper word alignment, using a barrel shifter. The number of barrel shifts needed by the master node to obtain word synchronisation will be the "BitSlide" value for the master. The number of barrel shifts needed by the slave node to obtain word synchronisation is encoded in the timing maker signal so the master node is able to calculate the total amount of barrel shifts (bit clocks) needed for both the slave node and the master node to obtain word synchronisation.

From this point the bidirectional link is up and synchronized. The master node sends a data packet and this is a sign to the slave that the reset sequence has completed and it may send data packets to the master node. The total propagation delay from master to slave and back is determined as described in chapter 3.

5 Conclusions

5.1 General

The test setup shows that it is feasible to measure the time offset between a master and a slave node that are connected via 10 km of fibre. When the bidirectional link between master and slave node is up and synchronized then the total propagation delay from master to slave and back is a constant delay expressed in bit clocks, that has a resolution of 800 ps (@ 1.25 Gbps). The total propagation delay fine time measurements done by the FPGAs were verified against the time interval measurements done with an oscilloscope.

The total delay resolution can be much higher than the bit clock resolution when the phase between master transmitter word-clock (TxUsrClk) and the master receiver recovered word-clock (RxUsrClk) is properly measured. By measuring this phase relation, delay variations in the communication channel can be continuously tracked such that these variations can be automatically compensated. Measurements done with the oscilloscope show that it is realistic to expect that the total delay resolution can be enhanced by at least an order of magnitude with respect to the 800 ps bit clock timing.

5.2 Further work

The test setup hardware facilitates fine time measurement, however it does not include actual time stamping and the software for PTP. This means that time stamping functionality must be added that includes master TxUsrClk/RxUsrClk phase information.

A media access control (MAC) service interface must be added and the PTP time stamps must be synchronized to the data packets. This should be organized as described in IEEE P802.3bf [11].

It is desirable to include all functionality in an IP-core such that it is easy for users to implement a timing aware Ethernet interface. However the functionality presented is highly dependent on the actual FPGA SerDes architecture, which makes it difficult to create a single universal IP-core.

The DPLL in the slave node must be optimized in order not to introduce jitter that deteriorates the timing measurement resolution.

The propagation delay dependency for fibre temperature and pressure should be further evaluated.

Appendix A

A-1 Dispersion and fibre choice

To communicate in two directions, the bidirectional communication channel between the master and the slave node is routed over a single fibre that is operated at dual wavelength. The test setup uses two small form factor pluggable (SFP) modules: the Optoway SPB-7610G transmits on 1310 nm and receives on 1550 nm while the Optoway SPB-7710G receives on 1310 nm and transmits on 1550 nm.

Using the test setup showed that the fibre choice is important. The first available fibre was 10 km of type "G655" which is optimized for 1550 nm. The connection between master and slave was established although it appeared not very stable. Investigation pointed out that this was not a matter of attenuation. Excessive jitter on the 1310 nm receiver was causing problems, as can be seen in Figure 15. Remember that the used fibre is optimized for 1550 nm which means that the chromatic dispersion at 1550 nm is low but inevitably the chromatic dispersion on 1310 nm is much higher.



Figure 15: Excessive jitter on the 1310 nm receiver

The Optoway devices use two different laser types for 1310 and 1550 nm. The 1310 nm device uses a multi-mode laser diode transmitter while the 1550 nm laser uses a single-mode distributed feedback (DFB) laser diode. The optical power spectra of both lasers are shown in Figure 16.



Figure 16: Power spectra for the two types of laser used

From this figure it is clear that the output spectrum width for the 1310 nm laser is much wider (specified: 2.5 nm RMS max.) than the spectrum of the 1550 nm DFB laser (specified: 1 nm RMS max.). This fact, and the fact that the fibre used has optimized chromatic dispersion at 1550 nm explains the excessive jitter that is observed in Figure 15.

This leads to the choice of a fibre that is optimized for 1310 nm, such as type "G652", which has a low chromatic dispersion on 1310 nm (< 3 ps/km.nm) but inevitable the chromatic dispersion at 1550 nm is much higher (< 18 ps/km.nm). Using type G652 fibre one would expect jitter on the 1550 nm receiver, however the output spectrum width for the 1550 DFB laser is much smaller which results in a small amount of jitter due to chromatic dispersion. Figure 17 shows the received waveforms when a type G652 fibre is used.



Figure 17: Waveforms at the receiver using a G652 type fibre

A-2 Compensate time offset for dispersion

A-2.1 General calculation

The time offset between master and slave node can be calculated by measuring the propagation delay P from master to slave node and back. Since both the forward and backward path follow the same physical fibre, the first estimation of the time offset between master and slave is in principle half the propagation delay measured:

(1)
$$t_{offset} = \frac{P}{2}$$

However, calculating the exact time offset is more complicated due to chromatic dispersion and internal FPGA delays. Figure 18 shows the various delays that contribute to the measured propagation delay.



Figure 18: Measured propagation delay broken down into various sub delays

In the above figure, the FPGA delays in the master and slave transmitter and receiver are taken to be equal ($\delta_{mTx} = \delta_{sTx} = \delta_{Tx}$ and $\delta_{mRx} = \delta_{sRx} = \delta_{Rx}$), which is justified since the functionality of the SerDes architectures in the master and the slave node is identical. Figure 19 shows that δ_{Tx} is in the order of 110 ns while Figure 20 shows that δ_{Rx} is in the order of 350 ns.



Figure 19: FPGA delay δ_{Tx} from Start pulse to electrical Tx terminal on the SFP



Figure 20: FPGA delay δ_{Rx} from electrical Rx terminal on SFP to Stop pulse

For simplicity the transmit and receive delays are combined in one variable, i.e. $\delta = \delta_{Tx} + \delta_{Rx} \approx 460$ ns.

The reference propagation delay t_{pd} in Figure 18 is taken to be the t_{pd} due to λ_1 . Therefore it is labelled $t_{pd\lambda_1}$. The chromatic dispersion is not zero and is different for each wavelength. From Figure 18 it can be seen that the measured propagation delay P equals:

(2)
$$P = 2.t_{pd\lambda 1} + t_{disp} + 2.\delta$$

The time offset between master and slave can be written as:

(3)
$$t_{offset} = t_{pd\lambda 1} + \delta$$

The dispersion time t_{disp} [ps] can be written in terms of the chromatic dispersion coefficient $D(\lambda)$ [ps/km.nm] which is a function of λ in [nm] per kilometre distance x (see equation(4) and (5)). This is a known and fixed factor which can be corrected.

(4)
$$\Delta t_{disp} = D.x.\Delta\lambda[ps]$$

thus:

(5)
$$t_{dsip} = x \int_{\lambda_1}^{\lambda_2} D d\lambda$$

Let $V_{\lambda 1}$ be the propagation velocity [km/ps] for λ_1 then:

(6)
$$t_{pd\lambda 1} = \frac{x}{V_{\lambda 1}}$$

When equation (5) and (6) are fed into equation (2) then x can be written in terms of P:

(7)
$$x = \frac{V_{\lambda 1} \cdot (P - 2\delta)}{2 + V_{\lambda 1} \cdot \int_{\lambda_1}^{\lambda_2} Dd\lambda}$$

The time offset between master and slave can now be calculated using equations (6) and (7) as input to equation (3):

(8)
$$t_{offset} = \frac{P + V_{\lambda 1} \cdot \delta \cdot \int_{\lambda_1}^{\lambda_2} Dd\lambda}{2 + V_{\lambda 1} \cdot \int_{\lambda_1}^{\lambda_2} Dd\lambda}$$

A-2.2 Measured and calculated values for the dispersion

The two wavelengths used in the test setup are 1310 nm (λ_1) from master to slave node and 1550 nm (λ_2) from slave to master node. A measurement shows that the actual dispersion is 15150 ps, for the fibre used (length: 10.7 km).

There are two methods for calculating the dispersion. The first method is already formulated in equation (5) and can be calculated as shown below. A type G652 fibre has a dispersion coefficient that is approximated by equation (9) [12] in which S_0 is the zero dispersion slope in [ps/nm².km] and λ_0 the zero dispersion wavelength in [nm].

(9)
$$D(\lambda) \cong \frac{S_0}{4} (\lambda - \frac{\lambda_0^4}{\lambda^3}) [ps / nm.km]$$

The integral of equation (5) can now be solved:

(10)
$$\int_{\lambda_1}^{\lambda_2} Dd\lambda = \frac{S_0}{8} \left(\lambda^2 + \frac{\lambda_0^4}{\lambda^2}\right) \Big|_{\lambda_1}^{\lambda_2} [ps/km]$$

There is no accurate value for S_0 and λ_0 in the datasheet of the fibre used [13]. However reference [12] provides more precise values for the type of fibre that is used ($S_0 = 0.086$ ps/nm².km and $\lambda_0 = 1313$ nm). The result of the integral from 1310 to 1550 nm is 2060 ps/km. The fibre has a length of 10.7 km so from equation (5) it follows that the dispersion equals 22042 ps.

The second method to calculate the dispersion is based on the different refractive indexes $(n_1 \text{ and } n_2)$ for the wavelengths used as is show in equations (11) and (12).

(11)
$$t_{pd\lambda} = \frac{xn}{c}$$

(12)
$$t_{disp} = t_{dp\lambda 2} - t_{pd\lambda 1} = \frac{x(n_2 - n_1)}{c}$$

The datasheet of the fibre used [13] specifies $n_{1310} = 1.467$ and $n_{1550} = 1.468$, i.e. not accurate enough for computing the offset. Reference [12] specifies the refractive indexes an order of a magnitude more precise ($n_{1310} = 1.4677$ and $n_{1550} = 1.4682$) such that the dispersion is calculated to be 17833 ps using equation (12). Table 1 gives a summary of measured and calculated results for the dispersion.

t _{disp}	[ps]
Measured	15150
Calculate using equation (10)	22042
Calculate using equation (12)	17833

Table 1: summary of measured and calculated results for t_{dsip}

It can be seen that these values do not correspond very well which is due to the limited accuracies of the values that are specified in the fibre datasheets. It should be noted that the datasheet values used in equation (10) are maximum values (not typical).

A-2.3 Example calculation for time offset

Equation (8) can be fed with realistic values. For a 10.7 km long fibre the measured propagation delay is 105.88 us thus the mean propagation velocity is $4.9.10^6$ ps/km. Due to fibre length, the propagation delay P dominates with respect to the second term in the numerator so the second term may be omitted. When the measured value for the dispersion ($t_{disp} = 15150$ ps over 10.7 km) is used then this means that the integral of D(λ) equals 1416 ps/km. Feeding the above values into equation (8) leads to equation (13) where the corrected time offset between master and slave node is calculated using the example values above:

(13)
$$t_{offset} \approx \frac{P}{2 + V_{\lambda 1} \cdot \int_{\lambda_1}^{\lambda_2} Dd\lambda} \approx \frac{P}{2 + \frac{1416}{4.9 \cdot 10^6}} \approx \frac{P}{2.000289}$$

This means that the corrected time offset (13) only differs a small amount compared to the uncorrected value given by equation (1). Although the difference is small it is significant and predictable. For a 10 km link it translates to a time offset difference in the order of 10 bits (@ 1.25 Gbps).

References

- [1] <u>http://www.km3net.org/</u>
- [2] P.P.M. Jansweijer, H.Z. Peek, *Measuring propagation delay over a coded serial communication channel using FPGAs*, Nucl. Instr. and Meth. A (2010), <u>http://dx.doi.org/10.1016/j.nima.2010.04.126</u>
- [3] <u>http://www.ohwr.org/projects/show/white-rabbit</u>
- [4] <u>http://ieee1588.nist.gov/</u>
- [5] IEEE Std 802.3, "Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications", chapter 36
- [6] http://www.embedded.com/215600261
- [7] <u>http://www.xilinx.com/products/devkits/HW-V5-ML507-UNI-G.htm</u>
- [8] <u>http://www.xilinx.com/products/virtex5/</u>
- [9] <u>Xilinx UG198 Virtex-5 FPGA RocketIO GTX Transceiver, User Guide</u>
- [10] http://www.xilinx.com/support/documentation/application_notes/xapp854.pdf
- [11] IEEE P802.3bf (in preparation), Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications. Media access control (MAC) service interface and management parameters to support time synchronization protocols.
- [12] http://www.photonics.byu.edu/FiberOpticConnectors.parts/images/smf28.pdf
- [13] <u>http://communications.draka.com/sites/eu/Datasheets/SMF%20-%20Enhanced%20Single-Mode%20Optical%20Fiber%20(ESMF).pdf</u>