

Real-time FPGA Implementation of Transmitter Based DSP

Philip M. Watts^(1,2), Robert Waegemans⁽²⁾, Yannis Benlachtar⁽²⁾, Polina Bayvel⁽²⁾, Robert I. Killey⁽²⁾

⁽¹⁾ Computer Laboratory, University of Cambridge, UK, philip.watts@cl.cam.ac.uk

⁽²⁾ Optical Networks Group, Dept of Electronic and Electrical Engineering, University College London, UK, rkilley@ee.ucl.ac.uk

Abstract Considerations for implementing transmitter based DSP for optical communications applications at 10 Gb/s and above are discussed including examples of linear and non-linear EPD compensation and OFDM generation.

Introduction

There has been huge recent interest in digital signal processing (DSP) for optical communications due to the advantages of reduced cost, increased flexibility and low sensitivity to aging and environmental effects when compared to optical techniques. This paper discusses the implementation of transmitter based DSP, in particular electronic predistortion (EPD) and generation of advanced modulation formats such as optical orthogonal frequency division multiplexing (OFDM). Application specific integrated circuits (ASIC) for EPD at 10 Gb/s are commercially available [1]. However, the cost of a set of masks for a 45nm CMOS design is around \$1million (and rising with each CMOS generation) creating a serious barrier for university based researchers wanting to demonstrate real-time viability. Other options such as shuttle runs in which several research designs are implemented on a single mask suffer from disadvantages such as inflexibility and considerably greater investment in development time and design tools. An alternative, discussed in this paper, is the implementation using field programmable gate arrays (FPGA) which are low-cost and reprogrammable. FPGAs have been used to demonstrate real-time operation at reduced bit-rates, for example in receiver based applications, digital coherent receiver [2] and OFDM [3], and precoding applications [4]. In this paper, considerations for the implementation of 20 GSample/s DSP suitable for generation of 10 Gb/s signals on FPGAs are discussed. Examples of implementation for EPD and generation of OFDM are given. In addition, the potential for achieving higher bit rates and increased EPD transmission distances are discussed.

Field programmable gate arrays (FPGA)

FPGAs are reprogrammable logic devices. They consist of an array of cells each of which contains look-up tables (LUTs) for implementing combinational logic and flip-flops for implementing sequential logic along with a network of reconfigurable interconnections. As DSP at optical line rates requires very high off-chip bandwidths, high-end FPGAs containing high speed serial transceivers are required. While these transceivers are equipped with coding, scrambling and error checking circuits

necessary to generate common communication standards (e.g. Ethernet, fibre channel), these are bypassed in the applications discussed in this paper, i.e. transceivers are used purely as time division multiplexers (TDM). These high-end FPGAs also contain hard-wired circuit blocks such as clock managers, block static random access memory (SRAM) and DSP multiplier accumulator (MAC) units.

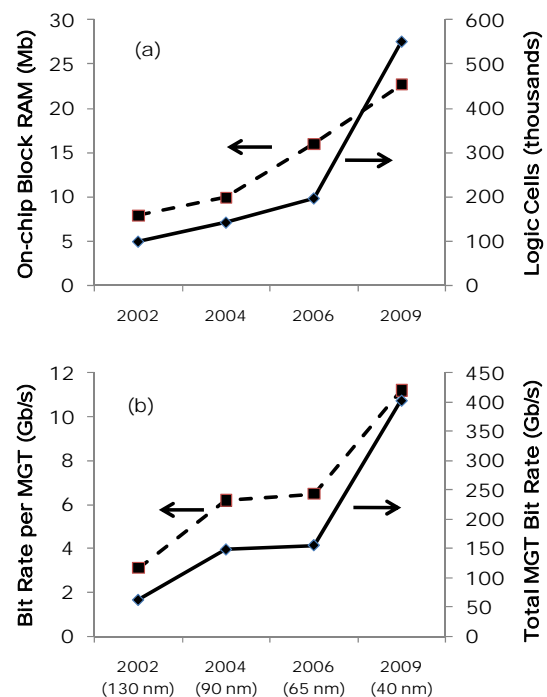


Fig. 1: Growing power of FPGAs 2002-2009 with CMOS generations indicated (a) logic cells and RAM (b) off-chip transceiver bandwidth [5]

FPGAs tend to use the latest CMOS process and so benefit from Moore's law in offering increased programmable logic and other circuit blocks in each new generation. Figure 1 compares the number of logic cells, RAM and transceiver bandwidth available in four generations of FPGAs between 2002 and 2009 [5]. The comparison is between the largest device in each generation which has sufficient numbers of transceivers to allow 20 GSa/s DSP. Although Xilinx devices have been used in this comparison (and throughout the paper), FPGAs from other vendors show similar trends and performance.

Considerations for implementing FPGA-based DSP for optical transmission

Off-chip bandwidth

21.4 GSa/s DSP (for the generation of 10.7 Gb/s EPD signals for example) on a single FPGA using 4-bit DAC resolution requires a total output bit rate of 85.6 Gb/s. As shown in figure 1, this has been possible since 2004. Although as the bit rate of an individual transceiver is limited to 11.2 Gb/s or less, most applications require external TDM to access optical line rates. Greater overall bandwidths can be obtained by using multiple FPGAs, each outputting a proportion of the overall signal. However, the communications data rate between FPGAs is very limited, so they must operate independently. Most DSP algorithms will involve considerable duplication for implementation on multiple FPGAs. However, some algorithms may partition well, for example, the EPD implementation described below, which uses one FPGA to calculate the real part of the transmitted optical field and another FPGA to calculate the imaginary part. Each FPGA must receive the same clock and the bit sequence input and transceiver outputs must be synchronised.

FPGA clock rate and DSP parallelism

Given the difference between the FPGA maximum clock rate and the required sample rate, parallel processing must be employed. In addition, in order to meet timing constraints, most algorithms will need to be pipelined over several clock cycles. The higher the degree of parallelism (i.e. the difference between the optical line rate and FPGA clock rate) the more logic resources are required for a given functionality. However, the parallelism must be sufficiently high to allow place and route within timing constraints. The clock rate is further constrained by the width of the transceiver parallel input (for example, configurable to 8, 10, 16, 20, 32, 40, 64 or 80 bits wide on the Xilinx Virtex-4). The clock rate must be the transceiver bit rate divided the input width.

An FPGA design for 21.4 GSamples/s DSP

As an example of the selection of off-chip bandwidth and clock rate, see figure 2 showing a block diagram of an FPGA design for 21.4 GSa/s DSP (requiring an external 4-bit DAC) [5]. The design was intended for 10.7 Gb/s EPD with linear FIR filters to compensate for chromatic dispersion. However, by replacing the DSP, many applications can be implemented using this scheme and it formed the basis of all the examples presented below. The target Virtex-4 4VFX100 contained 20 transceivers each capable of operating at up to 6.2 Gb/s. Therefore to obtain the required 85.6 Gb/s off-chip bandwidth, 16 transceivers operating at 5.35 Gb/s were used. The FPGA clock rate must be 5.35 GHz divided by one of the configurable parallel transceiver input widths. The

design was specified in VHDL and by an iterative process using synthesis and place and route design tools, it was determined that the maximum clock rate was 167.2 MHz ($5.35 \text{ GHz} / 32$). This sets the number of bits to be processed in a single clock cycle to 64 and the number of parallel FIR filters required to be 128 ($64 \text{ bits} \times 2 \text{ Sa/bit}$).

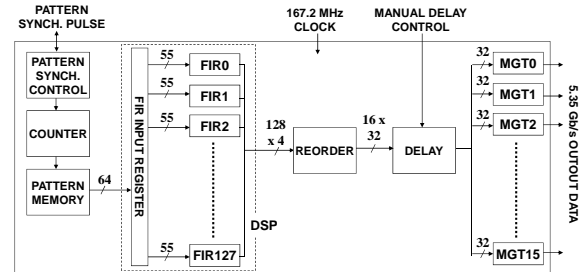


Fig. 2: An FPGA design for 21.4 GSa/s DSP with 4-bit resolution. A 55-tap parallel FIR filter is shown, but this can be replaced with many types of DSP [6].

Operation is as follows: The pattern to be transmitted is stored in a ROM on the FPGA. The ROM outputs 64-bits of the sequence in each clock cycle to the DSP, controlled by a counter. A synchronization circuit ensures that the counter is aligned to other FPGAs if required in the application. Each clock cycle, the DSP block inputs 64-bits of the pattern and outputs 128 ($64 \text{ bits} \times 2 \text{ Sa/b}$) 4-bit words each representing a DAC sample. These outputs are re-ordered into sixteen 32-bit words, appropriately delayed for alignment purposes and output by the transceivers.

Logic resource use

Determining the algorithm size (such as the number of FIR taps) which can be implemented on a given FPGA is largely an iterative process using the synthesis and place/route design tools. Tradeoffs between parallelism, pipelining and resource use should be investigated. As 100 % resource use is approached, achieving successful timing becomes increasingly difficult. Using additional FPGAs to increase available logic resources may be useful if each FPGA can carry out an independent function, otherwise logic duplication will yield diminishing returns.

For algorithms requiring large memories with a memory access per DAC output word (e.g. look-up tables (LUT)), the design is effectively limited by the on-chip RAM provided on the FPGA. Much larger DRAM memories could be employed off-chip. However, off-chip bandwidth restrictions and long/uncertain access times make this highly challenging.

The latest high-end FPGA also feature hard-wired DSP MAC circuits which have the potential to increase performance and reduce the requirements for configurable logic. The usefulness of these circuits depends on the nature of the application. The

DSP circuits were successfully used in our implementation of OFDM signal generation, but for EPD [6], configurable logic was used for all functions as routing signals to the fixed location DSP circuits impacted timing performance in this application. It is likely that the performance and quantity of these circuits on high-end FPGAs will increase with future generations.

Digital-to-analog converters (DAC)

External TDM and DAC are required to interface the DSP to the optical terminals. Depending on the application, 3-6 bits of effective resolution are required for low penalty operation which can be determined by simulation [7]. Arbitrary waveform generators are now available with sample rates of 24 GSa/s and resolution of 10 bits. DACs with sample rates of over 20 GSa/s suitable for 10 Gb/s EPD applications have been demonstrated in recent years [8] but these have been developed for proprietary applications and availability has been an issue for researchers. The work at UCL initially used a 4-bit TDM/DAC constructed from discrete components which was successfully used to demonstrate 10.7 Gb/s EPD and OFDM generation (described in the following sections). However, it suffers from a number of limitations [6, 9] and is not scalable to higher resolutions. More recently, integrated DACs have become available specifically for interfacing to FPGAs, for example the 25 GSa/s, 6-bit nominal resolution device from Micram [10]. This achieves an effective resolution of 5.5 bits for frequencies up to 6.25 GHz and has been used in a demonstration of EPD transmission at 10.7 Gb/s [11].

One issue with interfacing FPGAs to DACs is time alignment of the transceiver outputs for input to the DAC. Current FPGA do not have guaranteed skew-free transceiver outputs and additional fixed misalignment can be introduced by routing on the PCB. In addition, some families of FPGA (e.g. Virtex-4) have a random variable delay at start-up. In the UCL work with the discrete component DAC, digital delay circuits (to remove misalignments greater than 1 bit) and microwave phase shifters (for less than one bit) were added to every transceiver output. Manual calibration procedures were then used to eliminate misalignment [6, 12]. The Micram DAC allows an independent sampling phase to be set for each input, removing the need for phase shifters. In other respects the automated alignment of transceiver outputs to the Micram DAC inputs is achieved in a similar way to the UCL procedures [11]. An alternative automated technique was described in [13].

Linear compensation for chromatic dispersion using FIR filters

Chromatic dispersion compensation can be implemented using EPD with linear finite impulse response (FIR) filters. Simulation results have shown

that 3×10^{-3} taps/ps/nm, corresponding to 5.3 taps/100km of standard single mode fibre (SSMF) is optimum for EPD with 10.7 Gb/s NRZ-OOK signals [9]. One advantage of performing linear compensation at the transmitter is that the FIR inputs are 1-bit wide (compared to 3+ bits wide for receiver-based DSP) considerably reducing resource use. Another implication of this is that with transmitter-based compensation, FIR filters of up to 200 taps (sufficient to compensate for the dispersion of 4000 km SSMF at 10.7 Gb/s) are more efficiently implemented in the time domain than the frequency domain [8]. A further improvement in efficiency (reducing resource use by roughly half) can be obtained by exploiting the symmetry of the chromatic dispersion transfer function. In the case where cascaded FIR filters are used, for example, for compensating chromatic dispersion and the frequency response of the DAC [14], it is more efficient to combine the two impulse responses into a single response and use a single FIR to keep the advantage of 1-bit wide inputs.

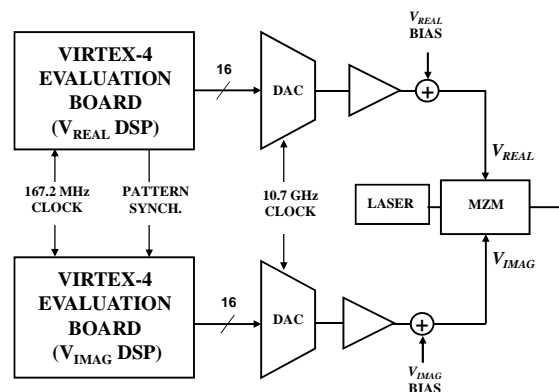


Fig. 3: 10.7 Gb/s FPGA-based EPD transmitter

A 10.7 Gb/s FPGA-based EPD transmitter was demonstrated using the design shown in figure 2 with the DSP block containing 128 x 55-tap parallel and pipelined FIR filters [6]. Figure 3 shows the top-level hardware used. 71% of the logic resources were used on each FPGA. In a recirculating loop experiment without optical compensation using 2^7 DeBruijn sequences, penalties of 1.4, 1.4 and 2.5 dB were obtained after 400, 800 and 1200 km of SSMF respectively compared with back-to-back for a BER of 10^{-3} . These penalties (and the larger penalties found for longer sequences) were found to be due to clock recovery issues (in recirculating loop operation) and imperfections in the discrete component DACs.

Nonlinear compensation using look-up tables

For overcoming non-linear impairments such as self-phase modulation (SPM), a non-linear DSP technique such as RAM-based look-up tables (LUT) is required [7]. A LUT allows any arbitrary waveform to be generated, however for accurate compensation, the LUT address width, n , must be chosen to be equal to the channel memory (pulse spreading due to

chromatic dispersion) in bits, and the required RAM size scales as 2^n . The RAM requirement is further increased by parallel DSP implementation as each parallel path requires a separate LUT. As only on-chip block RAM can be used (as explained in the 'resource use' section above), it is this which ultimately limits the LUT technique. To achieve compensation for 1200 km SSMF (as in the linear FIR compensation described above), 1 Gbits RAM would be required compared with 22 Mb on the largest current FPGAs [9].

In [11], the FPGA design of figure 2 and EPD hardware setup of figure 3 were upgraded to 6-bit output words in order to use the MICRAM DAC. The largest Virtex-4 FX series device (XC4VFX140) was used both to increase the off-chip bandwidth to accommodate 6-bit operation and maximise the block RAM available (9.9Mb). It was possible to use 11-bit LUTs when processing 64 bits in parallel at 10.7 Gb/s and 2 Sa/bit oversampling. Less than 1 dB penalty for distances up to 450 km of SSMF was measured with 0dBm launch power. However, when the launch power was increased to +4 dBm, less than 0.5 dB additional penalty was measured, compared with 2 dB penalty seen from the additional SPM effects introduced when using a 0 dBm LUT at +4 dBm.

OFDM Generation

Recent work has involved modifying the FPGA design of figure 2 to successfully demonstrate real-time generation of 8.3 Gb/s OFDM (the fastest implementation to date) using an optimised inverse FFT algorithm provided by Carnegie Mellon University through the SPIRAL project [15]. A single FPGA and DAC were used with an optical sideband filter following the MZM. Further details and results will be published shortly and presented at the conference.

Prospects for increased compensation and higher bit rates using FPGA-based compensation

Several approaches can be taken to increasing chromatic dispersion compensation at 10 Gb/s line rates, including using a higher proportion of the logic resources or DSP circuits to increase the number of FIR taps, use of alternative modulation formats with greater chromatic dispersion tolerance or adding LUTs. Perhaps the most effective method is to rely on CMOS scaling and use the latest FPGA technology. As shown in figure 1, the largest current FPGA contains 3.8 times more logic blocks than the 2004 (Virtex-4) equivalent used in the work reported here. As FIR logic resources scale linearly with taps, this should allow around 210 taps (equivalent to over 4000 km SSMF at 10.7 Gb/s) on a single FPGA while maintaining logic cell usage at 71%. For non-linear compensation however new approaches are required, as on-chip RAM is not increasing fast enough to significantly increase transmission distances using

the basic LUT technique. Significantly increased launch powers have been shown to be possible in transmission over 1200 km of SSMF using relatively modest LUTs (9-15 bit address width) with post filtering [16]. Another approach is to apply the output of a LUT as a perturbation to the FIR output [14]. Solving the nonlinear Schrödinger equation in real time [17] is very computationally intensive.

Increased bit rates lead to requirements for increased FPGA off-chip bandwidth, DAC conversion rate and DSP throughput (all scaling linearly with bit rate) and processor memory (scaling quadratically). The 400 Gb/s maximum total bandwidth of current FPGAs (figure 1) permits 40 Gb/s OOK-NRZ with 5-bit DAC resolution (assuming the availability of 80 GSa/s DACs). However, it is highly unlikely that the FPGA clock rate can be scaled up by a factor of four, leading to a more highly parallel DSP and greater resources required for a given number of taps. If the clock rate can be doubled to 334.5 MHz, 105 taps on a single FPGA should be possible leading to transmission distances of around 140 km of SSMF for 42.8 Gb/s OOK-NRZ. Advanced modulation formats such as QPSK and duobinary can reduce processor memory, off-chip bandwidth and DAC speed requirements [18] at the expense of greater logic complexity.

Acknowledgements

The authors acknowledge funding and support from EPSRC, the Royal Society, Intel Corporation, Huawei Technologies and Ericsson GmbH. Philip Watts would like to acknowledge funding from the Royal Commission for the Exhibition of 1851. The authors are grateful for the contributions to the work from Dr Madeleine Glick (Intel Research), Dr Stefan Herbst and Dr Cornelius Fürst (Ericsson GmbH), Prof. Markus Püschel, Prof. James Hoe, Peter Milder and Robert Koputsoyannis (Carnegie Mellon University).

References

- 1 J. McNicol et al., Proc OFC'05, OThJ3 (2005).
- 2 A. Leven et al., Proc OFC'08, OTuG3 (2008).
- 3 Q. Yang et al., Opt. Express **17**, 7985 (2009).
- 4 H. Song et al, OFC'08, OTuG3 (2008).
- 5 www.xilinx.com
- 6 P.M.Watts et al., Opt Express **16**, 12171 (2008).
- 7 R.I.Killey et al., PTL **17**, 714 (2005).
- 8 P.Schvan et al., Proc ISSCC'05, Paper 6.7 (2005).
- 9 P.M.Watts, PhD thesis, UCL (2008).
- 10 www.micram.com/index.php/products/vega
- 11 R.Waegemans et al, Opt Express **17**, 8630 (2009).
- 12 P.M.Watts et al., JLT **25**, 3089 (2007).
- 13 P.J.Winzer et al., Electronics Letters **44** (2008).
- 14 D.McGhan, Proc OFC'06, OWK1 (2006).
- 15 www.spiral.net
- 16 R.I.Killey et al., Proc OFC'06, OWB3 (2006).
- 17 X.Li et al., Opt Express **16**, 880 (2008).
- 18 P.M.Watts et al., Proc ECOC'07, 3.1.6 (2007)