Recent Advances in Low-Power Digital Signal Processing Technologies for Data Center Applications

Radhakrishnan Nagarajan⁽¹⁾, Lenin Patra⁽¹⁾, Agustin Martino⁽¹⁾, Christian Lutkemeyer⁽¹⁾, Damian Morero⁽¹⁾

⁽¹⁾ Marvell, 5488 Marvell Ln., Santa Clara, CA 95054, radha@marvell.com

Abstract Increasingly, low power optical interconnects are critical in data centers implementations. In this paper, we will discuss power-performance optimization strategies for DSP's that are widely used in these optical interconnects. ©2023 Marvell Technology

Introduction

Front face pluggable modules have been the mainstay of the optical transceiver market for the past 20 years. With very few exceptions, these modules are used as data center interconnects.

The evolution of aggregate data rates and power consumption in pJ/bit, (normalized to the data rate) of the pluggable modules for the past 20 years is shown in Fig. 1. This is an amazing trend where data rates have increased by 3 orders of magnitude and the energy consumption per bit has dropped by 2 orders of magnitude!



Fig. 1: Data rate and energy consumption for optical transceivers for the past two decades.

In the early 2000's, starting with the NRZ modulation thru the deployment of modules with the PAM4 format (at 200Gbit/s and higher), this downward trend has held.

Fig. 2 shows a nominal breakdown of power consumption for a DSP based pluggable module [2]. Although the actual numbers may vary between different vendors, we find this is to be generally true for a variety of DSP based modules (for both IMMDD as well as coherent). The power breakdown is also in line with other published data [3]. The DSP power is nominally about 50% of the module. Since the optical modules typically operate from a 3.3V supply, and most high-end ASIC's operate at supply voltages fraction of 1V, there is a power conversion loss (labeled "power overhead") that is between 10% and 15%.

As the data rates have increased, both the

analog (including optical) and DSP powers have scaled to keep the fractional power allocation in the pluggable modules at about the same levels. In this paper, we will focus on the contribution of the DSP to the power reduction trend.





The downward DSP power trend is due to several factors. First, is the continuous power reduction in the ever-shrinking CMOS nodes. There are also several other contributory factors. In this paper, we will highlight the role of lower complexity DSP designs, power efficient forward error correction (FEC) codes and adaptive power management techniques towards the DSP power reduction in data center optical interconnects.

Low Power Coherent DSP Designs

Equalizer and timing recovery are two critical blocks in DSP receivers in terms of power consumption. For coherent applications, an adaptive, fractionally spaced feed forward equalizer is required to compensate channel distortions, track dynamic effects and demultiplex the two orthogonal polarizations. Timing recovery is required to estimate and correct the sampling phase in the receiver.

One of the main design parameters impacting power is the oversampling factor. Power dissipation in the receiver front end is proportional to the sampling rate so an oversampling factor close to one is desired. However, there is one drawback, the fractional delay filters used in timing recovery become difficult to implement. This limitation can be eliminated by moving the timing recovery to the frequency domain.

With the increase in symbol rate some dynamic optical effects like polarization mode dispersion and residual chromatic dispersion have an impact on the number of taps required for proper equalization. To reduce the computational complexity of such large filters, a frequency domain approach is used to perform the filter convolution and the gradient correlation required to achieve adaptive filtering as shown in Fig. 3. Overlap and save or overlap and add can be used on the input samples to get linear instead of circular convolution. The ratio of overlap samples to FFT size must be carefully selected for the optimal power consumption.



Fig. 3: Frequency domain feed forward equalizer for coherent receivers. It is a fractionally spaced equalizer, so it needs more than one sample per symbol to work.

Traditionally for the timing recovery, a time domain fractional delay filter is used to resample the signal and align samples and symbols. After the timing correction, the signal is down sampled to one sample per symbol (Fig. 4).



Fig. 4: Interpolator used for timing recovery, the variable delay is used to align samples and symbols. After the filter, one sample per symbol is obtained.

As mentioned earlier, this operation, with oversampling factors close to one, can be efficiently handled in the frequency domain. We can take advantage of the already existing FFT and IFFT used for equalization, adjusting the ratio of FFT to IFFT size to match the change in sampling rate (Fig. 5). The required time delay corresponds to applying a linearly increasing phase shift as a function of frequency. The timing estimate can be obtained from the time domain, after the IFFT, or from the frequency domain bins. The number of overlap samples for the combined filter must accommodate the time domain response length of both filters.



Fig. 5: Combined frequency domain filter to perform equalization and timing recovery. A fold and add operation is required to change sampling rate in the frequency domain.

Low Power FEC Designs

Power consumption and latency of error correction codes are critical as well. For up to 10km (LR) the state-of-the-art 800G solution employs concatenated codes to combine the advantages of different code families. The concatenation of an inner soft-decision code and an outer hard-decision code with a total overhead of 21.23% and a BER threshold of 1.15e-2 (Fig. 6) was selected in the OIF 800LR standard since it can achieve efficient error correction while minimizing power consumption and latency [4]. The inner soft-decision code is a Bose-Chaudhuri–Hocquenghem (BCH) code of length 126 and dimension 110 bits, known for its robust error correction capabilities. By employing a power and latency optimized Chase decoding algorithm, it is possible to achieve reliable and efficient error correction. However, this relies on the stochastic nature of the noise. The outer hard-decision code based on a Reed-Solomon (RS) code of length 544 and dimension 514 10bits symbols, which is good at correcting burst errors, is used to eliminate the need for long interleaving schemes that degrades latency. Thus, the decoding process is simplified, reducing the computational complexity and latency compared to non-concatenated iterative soft-decision codes like iterative braided codes [5].

For more than 10km and up to 80km (ZR) applications the latency is less critical, and the focus is on performance and power. Here the state-of-the-art solution is based on an iterative braided code denoted as Open FEC (OFEC) proposed in the OIF 800ZR [5]. This code has an overhead of 15.32% and a threshold of 2e-2 as shown in Fig. 6. Several innovations were required in the decoder algorithm to reduce power, the most important is to divide the algorithm into several stages, typically two. The first stage takes advantage of the soft information to reduce the number of errors as much as

possible while keeping the number of iterations low, (typically <=3). The second stage is an iterative hard decoder algorithm focused on removing the residual errors of the first stage.

For the next generation of 1.6T transceivers it may be necessary to further reduce power consumption, particularly for ZR applications. A good candidate would be a combination of the current LR FEC previously described and the CFEC code proposed for 400G ZR standard [6]. The combination of the inner soft BCH of the LR FEC with the outer hard Staircase of the CFEC. This code here denoted CFEC-HP has an overhead of 22.21% and a threshold of 2e-2 as shown in Fig. 6. Due to the low power consumption of the Staircase, this provides similar power advantages as the concatenated scheme used in LR at the expense of a higher latency. Also note that same threshold of the OFEC is achieved at the expense of a higher This type of tradeoff between overhead. performance, complexity/power and code overhead is analyzed in more detail in [7] and it is exemplified through the relation between the CFEC (low complexity, low overhead, low performance), CFEC-HP (low complexity, high overhead, high performance) and OFEC (high complexity, low overhead, high performance).



Fig. 6: BER performance of the various FEC codes.

Adaptive Voltage Scaling

Manufactured CMOS chips exhibit a significant variation in their dynamic performance and power dissipation due to variations of physical device parameters like channel length, fin height, gate oxide thickness, and interconnect wiring parasitic capacitance. In a traditional fixed voltage design approach designers assume a digital supply voltage at a nominal level that is provided by a high efficiency switching regulator with a specific tolerance. In such a design, the worst-case dynamic performance of a device occurs when a "slow" chip is combined with a regulator at the low end of the tolerance window. The worst power occurs for a "fast" chip combined with a regulator at the upper end of the tolerance band.





The concept of Adaptive Voltage Scaling (AVS) has been used widely in chip design to reduce the worst-case power of digital circuits by about 20%, by adapting the supply voltage in a closed loop system so that the total power is minimized while the expected dynamic performance is still maintained. Fig. 7 shows an example of the AVS convergence voltage of a coherent DSP product over a wafer.

We observe significant spatial performance gradients over the wafer, up to 50mV between dies that are touching at their corners. An AVS system measures the dynamic performance in several locations to ensure a robust choice of the operating voltage even in the presence of substantial intra-die performance gradients.

Digital power of switching CMOS circuits is proportional to the square of the supply voltage. As such, to deliver smallest possible power, digital designers must aggressively reduce the operating voltage of the digital subsystems.

Aggressive voltage reduction creates significant challenges in the design process of high-performance digital circuits. The use of an AVS system can help to overcome these challenges and enables to simplify the timing closure of such designs by replacing a circuit optimization at cold temperatures with an increase of the supply voltage depending on the junction temperature of the device. Such a strategy results in an overall more favorable chip.

Conclusions

A holistic approach to architecture, circuit, and process optimization, together with adaptive operating techniques, is needed to minimize power consumption in DSP's. In this paper we have explored a few techniques that we will discuss in detail during the presentation.

Acknowledgements

The authors would like to thank the exceptional engineering team at Marvell who contributed enormously to this work.

References

- [1] R. Nagarajan, L. Ding, R. Coccioli, M. Kato, R. Tan, P. Tumne, M. Patterson, and L. Liu, "2.5D Heterogeneous Integration for Silicon Photonics Engines in Optical Transceivers," *IEEE J. of Selected Topics in Quantum Electronics*, vol. 29, no. 3, pp. 1-10, 2023, DOI: 10.1109/JSTQE.2022.3214418.
- [2] R. Nagarajan, I. Lyubomirsky, and O. Agazzi, "Low Power DSP-based Transceivers for Data Center Optical Fiber Communications," *J. of Lightwave Technology*, vol. 39, no. 16, pp. 5221-5231, 2021, DOI: 10.1109/JLT.2021.3089901.
- [3] C. Fludger, "Performance orientated DSP design for Flexible Coherent Transmission (Tutorial)," in *Proc. OFC*, San Diego, CA, USA, 2020, pp. Th3E.1.

- [4] Implementation Agreement 800LR (*baseline text*), Optical Internetworking Forum, Fremont, CA USA, oif2023.070.01.
- [5] Implementation Agreement 800ZR (*draft*), Optical Internetworking Forum, Fremont, CA USA, oif2021.144.15.
- [6] Implementation Agreement 400ZR, Optical Internetworking Forum, Fremont, CA USA, OIF-400ZR-01.0, Mar. 2020.
- [7] D. Morero, M. Castrillón, A. Aguirre, M. Hueda and O. Agazzi, "Design Tradeoffs and Challenges in Practical Coherent Optical Transceiver Implementations," in *J. of Lightwave Technology*, vol. 34, no. 1, pp. 121-136, 1 Jan.1, 2016, DOI: 10.1109/JLT.2015.2470114.