

ENERGY-OPTIMIZED HIGH PERFORMANCE FFT PROCESSOR

Dongsuk Jeon¹, Mingoo Seok¹, Chaitali Chakrabarti², David Blaauw¹, Dennis Sylvester¹

¹Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI, USA

²School of Electrical, Computer and Energy Engineering
Arizona State University, Tempe, AZ, USA

ABSTRACT

This paper proposes an ultra low energy FFT processor suitable for sensor applications. The processor is based on R4MDC but achieves full utilization of computational elements. It has two parallel datapaths that increase throughput by a factor of 2 and also enable high memory utilization. The proposed design is implemented in 65nm CMOS technology and post-layout simulation including parasitic capacitances shows it achieves 9.25× higher energy efficiency than state-of-the-art FFT processors and high throughput relative to past subthreshold circuit implementations.

Index Terms— Fast Fourier Transform, Pipelined architecture, Parallel-pipelined architecture, Energy-optimal design

1. INTRODUCTION

FFT (Fast Fourier Transform) is one of the major digital signal processing algorithms that is widely used in digital communication, speech processing and image processing applications. There has been renewed interest in FFT because of its application in mid- or high-performance sensor nodes that require audio recognition or relatively advanced communication schemes [1][2]. Although an FFT can be simply realized on a general purpose processor, standalone FFT co-processors are well suited to sensor nodes due to their higher energy efficiency [3].

For battery-powered sensor nodes, energy per operation is the key constraint, rather than power, and voltage scaling is an efficient way to achieve low-energy computation. While switching energy is reduced as voltage scales down, leakage dominates below some voltage and it limits energy-efficient operation especially in the subthreshold regime (where the operating voltage is less than the transistor threshold voltage). Aggressive voltage scaling also degrades performance due to an exponential rise in delay in the subthreshold regime, making it difficult to meet throughput requirements. For example, the subthreshold FFT processor in [4] operates at 10kHz and dissipates 45% of its energy as leakage at its energy-optimal point of 350mV.

In this paper, we describe a FFT processor architecture that is highly optimized for energy-efficient operation in the ultra-low voltage regime. The R4MDC architecture is a straightforward way of implementing FFT algorithm with the major drawback of low utilization of computational elements [5]. Here the conventional R4MDC architecture is modified to allow full utilization of each

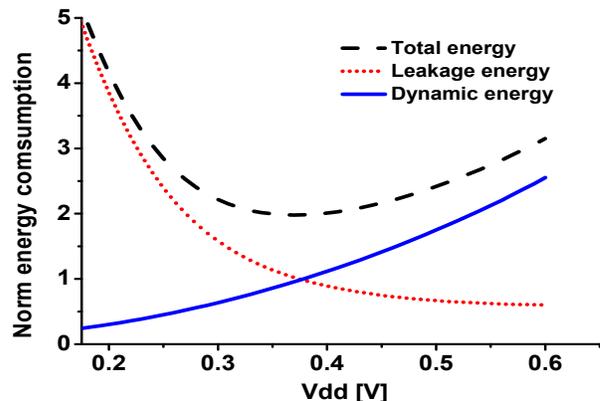


Fig. 1. Energy dissipation of an inverter chain of length 60 in 65nm CMOS technology, indicating a minima at 0.35V

computational element for energy-optimal computation as well as enhanced throughput. Also, parallel-pipelined architecture is used to further improve these benefits. We show that the proposed architecture has significant gains over other pipelined architecture such as R2²SDF. The 1024-point FFT processor based on this architecture was implemented using CMOS 65nm technology and post-routing simulation results show more than 9× higher energy efficiency than the previous best result published at 152MS/s, which is higher performance than conventional subthreshold designs with throughput on the order of 10kS/s [4].

2. BACKGROUND AND RELATED WORK

2.1. Aggressive voltage scaling

Since the dynamic energy consumption is proportional to the square of power supply voltage, we can save dynamic energy simply by reducing power supply voltage. However, delay increases rapidly at low voltages and logic gates spend more time leaking in each cycle. Therefore leakage energy consumption becomes comparable to, or even larger than, dynamic energy at very low voltages. This worsens the energy efficiency beyond a specific voltage and the most energy-optimal supply voltage (V_{min}) is defined as delivering the smallest energy-per-operation (E_{min}) as shown in Fig. 1 [6].

In ultra low voltage FFT processors, generally low utilization of memory cells or computational elements exacerbates leakage

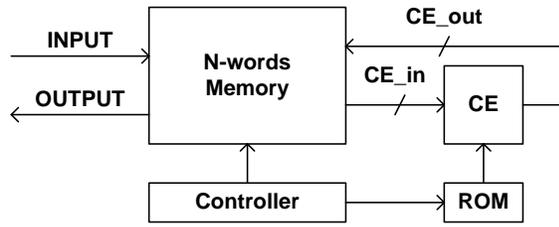


Fig. 2. Memory-based architecture.

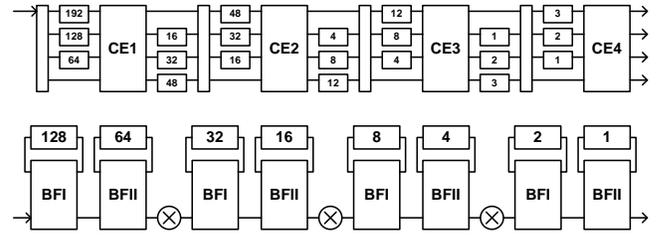


Fig. 3. R4MDC and R2²SDF FFT architecture (N=256).

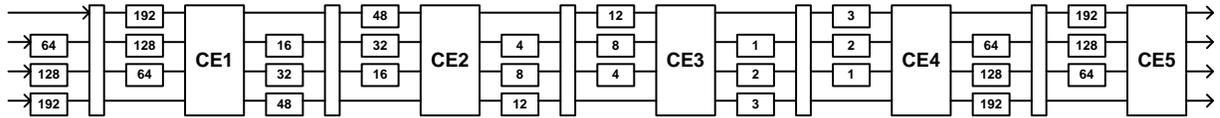


Fig. 4. Modified R4MDC processor (N=1024).

energy and causes higher V_{\min} and E_{\min} . Thus it is critical to limit idling cells for higher utilization and enable better overall energy efficiency.

2.2. Memory-based architecture

The memory-based architecture is the simplest way of implementing the FFT algorithm. As shown in Fig. 2, it consists of a large memory used for input/output buffer and scratchpad, a single CE (Computational Element), and a control module. The CE accesses a few words of the entire memory for each computation. As a result, while the CE processes one set of data, the remaining memory cells simply store intermediate results and consume leakage power. Although the CE works every cycle and hence has full utilization, the memory utilization is very low. From SPICE simulations, we observe that 85% of overall energy is dissipated in memory at 300mV supply for a radix-4 memory-based architecture.

Furthermore, the memory-based architecture is not appropriate for applications that require successive FFT of incoming data such as voice recognition. As the main memory has to store intermediate results during one set of FFT, another (ping-pong type) input buffer is needed to store input data temporarily, which significantly increases memory leakage power consumption.

2.3. Pipelined architecture

The pipelined architecture is composed of several stages, where each stage consists of a CE and input/output FIFO buffers. Although the pipelined architecture incurs more hardware cost and consumes relatively high power because all stages are switching every cycle, it can be more efficient in terms of energy per operation. The CEs access only the FIFOs of the previous stage and since each FIFO is much smaller than the single memory in the memory-based architecture, the average number of memory cells per CE is reduced significantly, lowering memory leakage compared to the memory-based architecture.

The most straightforward type of pipelined architecture is MDC (Multi-path Delay Commutator). Each pipeline stage corresponds to one segment of the signal flow graph and the CE computes the input data in exactly the same way. However, conventional one-input-per-cycle R4MDC architecture suffers from only 25% utilization of CE and requires $3(\log_4 N - 1)$ complex multipliers and $5N/2 - 4$ memory cells. Its drawbacks are low utilization and high hardware cost.

One of the popular pipelined architectures is the R2²SDF (Single-path Delay Feedback) architecture [7]. It has a feedback path beside every CE and stores the input data temporarily as shown in Fig. 3. It has only $\log_4 N - 1$ complex multipliers and $N - 1$ memory cells and can achieve 75% utilization for the complex multiplier. Therefore it is considered a low-power and low-cost architecture.

3. ENERGY-OPTIMAL ARCHITECTURE

3.1. Modified R4MDC architecture

The original R4MDC architecture accepts only one input and waits for all 4 input data required for the first stage radix-4 butterfly to be collected. Each CE is activated only once in a 4 cycle period, resulting in CE utilization as low as 25%. However, if 4 inputs from a single channel are fed at once, we can achieve perfect CE utilization and reduce memory requirements.

Although accepting multiple inputs per cycle can cause performance mismatch with other modules of the system, this can be accommodated by applying $4\times$ slower clock domain and attaching a small 4 word buffer to store incoming data for 4 cycles with negligible overhead. For example, in battery-powered or energy-harvesting sensor node applications, the entire system typically includes several voltage domains based on the power/performance requirements of each module. A multiple-input-per-cycle type FFT can enable high energy efficiency while leveraging the voltage and clock domains commonly found in the targeted systems. The multi-channel FFT to achieve higher utilization of conventional pipeline architectures in [8] is no longer as attractive.

Fig. 4 shows the architecture of a 1024-point modified R4MDC FFT processor. It is similar to the original R4MDC architecture with the key distinctions being: 1) it reads 4 input samples at a time, 2) the schedule is different, and 3) the configurations of each FIFO are different. While the original R4MDC employs a set of large buffers at the input stage to convert a serial data stream to a parallel data stream that is fed into the first CE, the modified R4MDC requires only half as many buffers for input re-ordering. The switching network configuration in the commutator remains the same as the original one. The twiddle-factor ROM and controller are embedded in each CE.

An important advantage of this architecture is that it requires fewer memory cells. While the conventional R4MDC requires

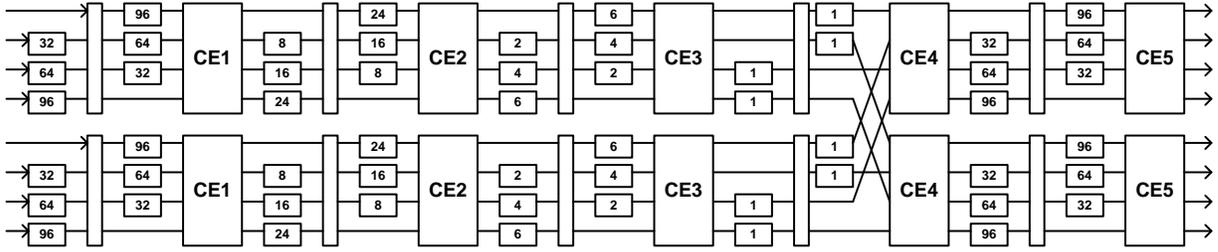


Fig. 5. Modified R4MDC processor with 2 processing paths (N=1024).

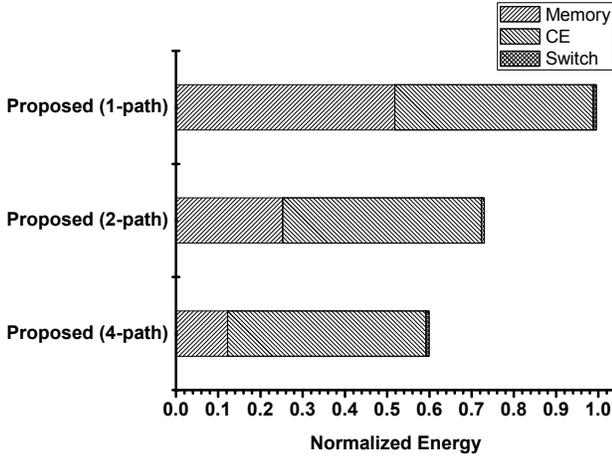


Fig. 6. Energy breakdown of modified R4MDC architecture.

5N/2-4 memory cells, the new architecture contains only 7N/4-4 cells (30% fewer for 1024-pt FFT). In comparison, the 4-channel R4MDC architecture in [8] requires 4N-4 memory cells. Since it does not need the large buffer before the first stage to store incoming data and change serial input to parallel data for the CE, the hardware cost and energy consumption are reduced significantly. Overall, memory utilization is improved and memory leakage energy is reduced.

Nearly all conventional architectures access the scratchpad memory every cycle and the memory is therefore regarded as perfectly utilized. However, this does not reflect the actual proportion of accessed cells in memory. The average number of memory cells per complex multiplier can be considered a good alternate metric in our energy-constrained applications. For a 1024-pt FFT, the R²SDF architecture has 255.75 cells per multiplier. On the other hand, the modified R4MDC has 149 cells per multiplier, or 41% fewer. This indicates that even though the R²SDF has smaller hardware cost, modified R4MDC achieves 4× more throughput and saves energy per FFT by reducing the number of underutilized memory cells. Simulations show that memory consumes 52% and 68% of total energy at 250mV in modified R4MDC and R²SDF, respectively.

3.2. Parallel-R4MDC architecture

Despite the improvements shown above, even the modified R4MDC consumes 52% of overall energy in its memory. Next we show how by parallelizing the modified R4MDC architecture, we not only increase the throughput but further reduce the number of

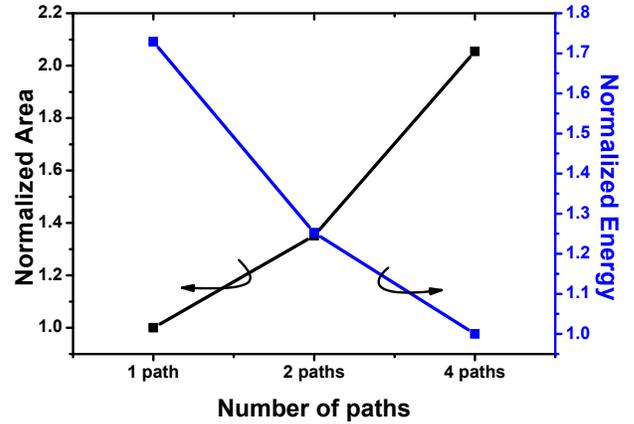


Fig. 7. Area-energy plot for parallel-pipelined architecture (N=1024).

memory cells per complex multiplier, translating to lower leakage energy.

Fig. 5 shows the modified R4MDC architecture with 2 paths in parallel. Each path processes incoming data within their own input set until the 2nd stage from the last. After that, the intermediate data is exchanged as shown in Fig. 5. The hardware requirements of the conventional and two proposed architectures are shown in Table 1. This 2-path version requires 4N/7-8 memory cells. While it appears that the memory size is nearly the same as the 1-path version, the average number of memory cells per multiplier decreases by more than half since the number of CEs doubles and throughput doubles simultaneously. The 2-path version has only 34.9% total energy consumed in memory and the energy dissipated in the CEs per FFT does not change.

The method can be extended to 4 or more paths. The energy breakdown of the 1-path, 2-path and 4-path versions is shown in Fig. 6. We see that as the number of processing paths grows, memory energy is further reduced. However, with more paths the area increases linearly while the energy starts to saturate as the CEs become dominant and the contribution of memory is negligible in terms of both energy and area.

4. EXPERIMENTAL RESULTS

The proposed architecture was simulated in a commercial 65nm CMOS technology. Dynamic and leakage energy of each module including memory and CE were obtained by simulations that include parasitic RC elements extracted post-layout. Based on the result in Fig. 7, the 2-path version is chosen as the optimal design

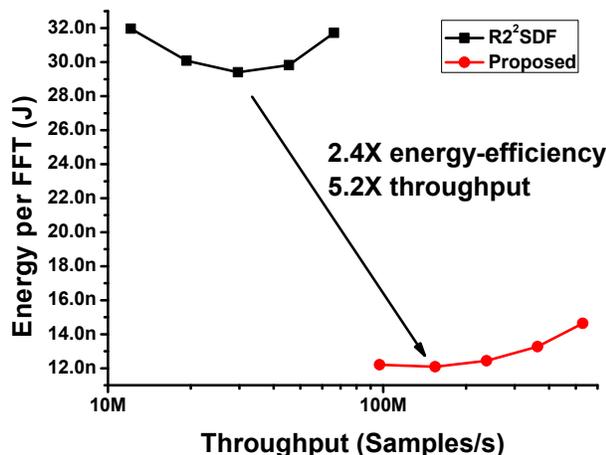


Fig. 8. Energy per FFT and throughput of R²SDF and proposed architecture at 200-300mV.

point. Beyond this point, energy savings plateau with an undesirable linear growth in area.

Simulations shows that a 1-path modified R4MDC architecture consumes 43.2% less energy with 2.6× better performance than R²SDF architecture at their respective energy-optimal points V_{min} . The V_{min} decreases from 250mV to 225mV in the 1-path modified R4MDC case due to the reduction of memory leakage. In addition, the 2-path version consumes 27.5% less energy and shows 2× more throughput than the 1-path version with the same V_{min} . In sum, the proposed 2-path version has 2.4× better energy efficiency and 5.2× more throughput compared to conventional R²SDF architecture at their energy-optimal points.

The characteristics of the implemented 2-path version FFT processor and comparisons with other processors published are shown in Table 2. Our design shows 9.25× better energy efficiency than current state-of-the-art while providing high throughput compared to other subthreshold based designs. In addition, this concept was successfully proven with fabricated chip [11].

5. CONCLUSIONS

In this paper, we proposed a highly energy-optimized FFT architecture for smart sensor applications. We modified the conventional R4MDC topology to achieve full CE utilization and applied parallel-pipelined concepts for further reducing memory energy. This approach can be extended to mixed-radix systems. Finally, the proposed 2-path FFT processor was implemented in 65nm technology and post-layout simulation results show record energy efficiency.

6. ACKNOWLEDGEMENT

This research was supported by the U.S. Army Research Laboratory under contract W911NF and prepared through collaborative participation in the Microelectronics Center of Micro Autonomous Systems and Technology (MAST) Collaborative Technology Alliance (CTA). Authors also acknowledge the IC fabrication support of STMicroelectronics and the National Science Foundation through grants CNS-0910851 and CSR-0910699 for their support in this work.

Table 1. Comparison with other pipelined FFT architectures.

	Complex multipliers	Complex adders	Memory space	Utilization adder	Utilization mult	Throughput (sample/cycle)
R4MDC	$3(\log_2 N - 1)$	$8\log_2 N$	$5N/2 - 4$	25%	25%	1
R ² SDF	$\log_2 N - 1$	$4\log_2 N$	$N - 1$	75%	75%	1
Proposed (1-path)	$3(\log_2 N - 1)$	$8\log_2 N$	$7N/4 - 4$	100%	100%	4
Proposed (2-path)	$6(\log_2 N - 1)$	$16\log_2 N$	$7N/4 - 8$	100%	100%	8

Table 2. Characteristic comparison with other fabricated FFT processors. The energy is normalized based on the eqn. in [9] for technology and word length. The energy of real-valued FFT is converted to complex-valued FFT by doubling.

	This work	[4]	[9]	[10]
Technology	65nm	180nm	130nm	600nm
FFT mode	1024-point complex-valued	128-1024-point real-valued	64-16384-point complex-valued	1024-point complex-valued
word width	16 bit	16 bit	16 bit	20 bit
Vdd	0.2-0.3 V	0.18-0.9 V	1.1-1.5V	1.1-3.3 V
area	$2.71 \times 3.15 \text{ mm}^2$	$2.6 \times 2.1 \text{ mm}^2$	N/A	$5.9 \times 8.2 \text{ mm}^2$
design point	1024-point CV 0.225V, 19MHz 152MS/s	1024-point RV 0.35V, 10KHz	1024-point CV 1.1V, 140MHz	1024-point CV 1.1V, 16MHz
Energy/FFT	12.1 nJ	155 nJ	1098.9 nJ	3135 nJ
Normalized Energy/FFT	12.1 nJ	111.9 nJ	549.5 nJ	250.8 nJ

7. REFERENCES

- [1] W. Tang and L. Wang, "Cooperative OFDM for energy-efficient wireless sensor networks," *Proc. IEEE Workshop on Signal Processing Systems*, Washington, pp. 77-82, 2008.
- [2] D. Zhao, H. Ma and L. Liu, "Event classification for living environment surveillance using audio sensor networks," *Proc. IEEE International Conference on Multimedia and Expo*, Singapore, pp. 528-533, 2010.
- [3] S. Sridhara, et al., "Microwatt Embedded Processor Platform for Medical System-on-Chip Applications," *Proc. IEEE Symposium on VLSI Circuits*, Hawaii, pp. 15-16, 2010.
- [4] A. Wang and A. Chandrakasan, "A 180-mV subthreshold FFT processor using a minimum energy design methodology," *IEEE Journal of Solid-State Circuits*, pp. 310-319, 2005.
- [5] E. E. Swartzlander, et al., "A radix 4 delay commutator for fast Fourier transform processor implementation," *IEEE Journal of Solid-State Circuits*, pp. 702-709, 1984.
- [6] B. Zhai, et al., "Energy-Efficient Subthreshold Processor Design," *IEEE Transactions on VLSI Systems*, pp. 1127-1137, 2009.
- [7] S. He and M. Torkelson, "A new approach to pipeline FFT processor," *Proc. International Parallel Processing Symposium*, Hawaii, pp. 766-770, 1996.
- [8] S. Lee, et al., "Low complexity pipeline FFT processor for MIMO-OFDM systems," *IEICE Electronics Express*, pp. 750-754, 2007.
- [9] T. Pitkanen and J. Takala, "Low-power application-specific processor for FFT computations," *Proc. International Conference on Acoustics, Speech and Signal Processing*, Taiwan, pp.593-596, 2009.
- [10] B. M. Baas, "A low-power, high-performance, 1024-point FFT processor," *IEEE Journal of Solid-State Circuits*, pp.380-387, 1999.
- [11] M. Seok, et al., "A 0.27V 30MHz 17.7nJ/transform 1024-pt Complex FFT Core with Super-Pipelining," *Proc. IEEE International Solid-State Circuits Conference*, in press, 2011.