A Compact 446 Gbps/W AES accelerator for Mobile SoC and IoT in 40nm

Yiqun Zhang, Kaiyuan Yang, Mehdi Saligane, David Blaauw, Dennis Sylvester
University of Michigan, Ann Arbor, MI Email: zhyiqun@umich.edu

Abstract

An AES hardware accelerator targeting energy efficient, low cost mobile and IoT applications is fabricated in 40nm CMOS. The proposed design eliminates the ShiftRow stage in conventional AES implementations and replaces flip-flops in data and key storage with latches using re-timing, saving 25% area and 69% power. Along with a 2-stage Sbox in native $GF(2^4)^2$ composite-field computation and glitch reduction techniques, this results in a compact 2228 gate design achieving 446 Gbps/W and 46.2 Mbps throughput at 0.47V.

Introduction

Security is critical for modern electronic devices with internet connectivity. Advanced Encryption Standard (AES) is a widely-used block cipher algorithm for symmetric encryption in a large range of applications. For mobile devices, silicon area (i.e., cost), throughput, and energy efficiency are all key design constraints [4]. Recently, several energy efficient implementations were presented [1,2]. However, their kbps-range throughput cannot meet the demands of mobile devices with high-speed data streaming. Highly parallelized implementations [3] provide Gbps throughput, which is critical in server applications. However, their large silicon footprint is disadvantageous in cost-sensitive mobile SoCs. This paper presents a voltage-scalable AES accelerator targeting mobile SoCs and IoT devices with ~50−500Mbps throughput, while achieving best-in-class area and energy efficiency. The proposed accelerator is fully synthesizable and implements 128-bit AES using only 2228 logic gates. By eliminating the ShiftRow and MixColumn registers and replacing data and key storage with latches, area is reduced by 41%. This, along with retiming of a 2-stage Sbox design in native $GF(2^4)^2$ composite-field computation, leads to a 3.38× energy efficiency improvement over a baseline implementation at nomial voltage with four 128-bit registers and 1-cycle $GF(2^4)^2$ Sbox methods. The proposed design achieves 1.3GHz at 0.9V, peak throughput of 494 Mbps, and peak energy efficiency of 446Gbps/W. Implemented in 40nm CMOS, the accelerator area is only 0.00429mm^2, marking the smallest AES accelerator considering technology scaling.

Energy Efficient AES

Fig. 1 shows the standard implementation of AES encryption using an 8-bit datapath, which was implemented in the same 40nm test chip as a baseline. Simulated power breakdown of functional modules (Fig. 1) shows that the four 128-bit registers (DataReg, MixColReg, KeyReg, and ShiftReg) constitutes ~50% of total AES power. Our approach reduces this storage to only a 128-bit latch-based DataReg, a 48-bit latch based StorageReg, and a "one-hot" indexed 128-bit latch based KeyReg; these changes reduce total sequential power by 31%.

Fig. 2 describes several ways of storing data (both input and intermediate) within AES accelerators. DataReg first stores the initial plain text and is then updated with calculated cipher text at each iteration of the algorithm. ShiftRow and MixColumn blocks compute 32-bit outputs every 4 cycles that are stored in ShiftReg and MixColReg, respectively. The authors of [4] eliminate ShiftReg by loading plaintext into DataReg in the ShiftRow byte-order.

As shown in the byte-location index in the matrix of Table 1, the data of locations L2, L5, and L8 in the 4th cycle, L1 and L4 in the 8th cycle, and L0 in the 12th cycle cannot be stored back to DataReg immediately after they are computed by the MixColumn module (highlighted in Table 1). In the proposed design, these 6 bytes are stored in a 48-bit StorageReg using the decode logic in Fig. 4. The hardwired data transfer from MixColumn output to DataReg removes the 128b ShiftReg and MixColReg (each built of 128 flip-flops) and instead uses StorageReg, consisting of only 48 registers. As a result, the total register count for the datapath is reduced to 176, compared to 384 in a conventional design and 256 in [4], marking a 30% reduction.

To further reduce sequential power and area, the design is modified to accommodate latch-based registers instead of flip-flops. This is accomplished by adding an 8-bit AdderReg (Fig. 4) and 1 additional cycle of latency (a 0.3% increase to 337 total cycles of latency), since data is bypassed to skip MixColumn at the last iteration in the AES algorithm. This change does not impact the clock frequency since DataReg is not on the critical path. Fig. 2 includes the dynamic energy and area values for each implementation. The proposed approach has a 2.66×/2.9× energy/area improvement over a conventional design and a 1.78×/1.94× energy/area improvement over [4] for these three registers. Finally, dynamic glitch power is a significant concern in AES hardware accelerators. Hence, clock gating is used in dataReg to reduce glitch power for this part by 2.74× and total power by 30%.

In addition to plain/cipher text processing, the key is also updated in each iteration of AES. The 128-bit input key is stored in the KeyReg and one byte is updated by KeyGen in each iteration. Using address generation to access the correct byte in each iteration results in a large gate count and area. Using a basic shift register reduces area but increases power. Instead, in the proposed design KeyReg is changed from a 128-bit flip-flop register to a 128-bit latch register using one-hot shift-based addressing. This design uses a cyclic address generator with a single chain of 16 single-bit registers (Fig. 4), similar to [6]. This requires 1-bit shifting rather than 128-bit shifting, reducing area by 23% and improving power by 18% (Fig. 5) compared with the conventional register implementation with decoder.

The final block to be optimized is the Sbox stage, which contributes 12% of total power (Fig. 1) and contains the accelerator's critical path. Sbox implementation choices include SRAM-based, logic based look-up table, and native composite-field $GF(2^4)^2$; these are analyzed via simulation in Table 2. A conventional single-cycle $GF(2^4)^2$ offers compact area at the expense of power and performance. This higher power is due in part to the difference in signal arrival times of fast and slow paths in the Sbox (Fig. 3), resulting in glitch power [5]. To address this, we re-time the Sbox datapath by adding 12 flip-flops before the path converges, equalizing path delays. This incurs a modest 4.3% area overhead while providing 37% power savings at the same frequency as the 1-cycle $GF(2^4)^2$ implementation. Also, splitting Sbox into two cycles shortens the critical path, decreasing clock cycle time by 28%; through voltage scaling this improves accelerator energy efficiency by 3.38× energy efficiency at iso-frequency.

Measurements & Conclusion

The proposed AES accelerator was implemented in 40nm CMOS along with a separate baseline implementation. Fig. 5 shows the simulated power breakdown of baseline and proposed designs. At 0.9V and 25°C the proposed design has a measured Fmax of 1.3GHz while consuming 4.39mW (Table. 3). The proposed design is fully synthesized, enabling operation across a wide voltage range. Fig. 6 shows the measured clock frequency, throughput, energy efficiency, and power across Vdd. At 1V, performance of 1.47GHz is obtained while peak energy efficiency of 446 Gbps/W is achieved at Vdd = 0.47V. Compared with [4], the proposed design is 41% smaller considering technology scaling and 3.1× more energy efficient at 432Mbps throughput. Overall the power consumption of the proposed design is compatible with mobile SoCs at its highest performance point (4.39mW) and offers a compelling option for IoT applications as it consumes only 100μW with 46.2Mbps throughput at sub-0.5V. Fig. 7 shows the die photo.

Acknowledgement

We thank the TSMC University Shuttle Program for chip fabrication.

References

[1]J. Myers, et al., ISSCC, 2015.
[2]W. Zhao, et al., TVLSI, 2015.
[3]S. Mathew, et al., JSSC, 2011.
[4]S. Mathew, et al., JSSC, 2015.
[5]S. Morioka, et al., CHES, 2002.
[6]D. Jeon, et al., JSSC, 2012.
[7]P. Hamalainen, et al., EUROMICRO, 2006.
[8]T. Good, et al., TVLSI, 2010

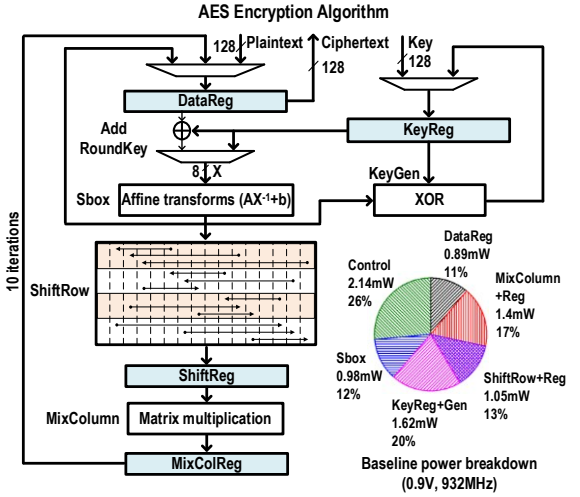


Fig.1. Standard implementation of AES encryption circuit. Total Baseline power is based upon measurement results.

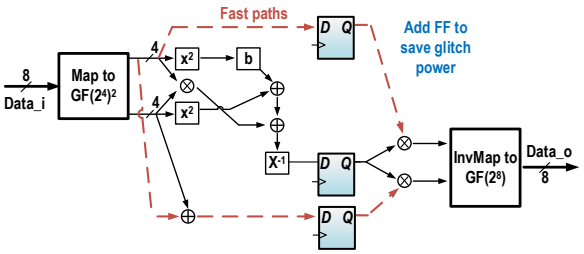


Fig.3. Sbox logic path in native composite-field (GF).

Table 2. Comparison table of Sbox implementations (based on simulation results)

S-box Architecture	Conventional			Proposed
	Look-up table (SRAM)	Look-up table (logic)	$GF(2^4)^2$ in 1 cycle	$GF(2^4)^2$ in 2 cycle
Area (μm^2)	2175	816	558	582
Power (mW)	1.7	1.15	1.42	0.9
Cycle Time (ns)	0.55	0.4	0.64	0.46

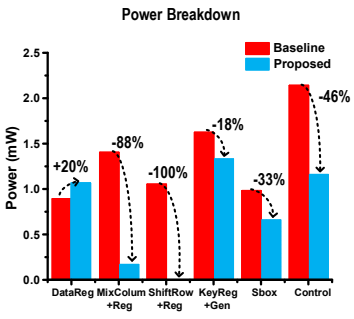


Fig.5. Simulation based power breakdown.

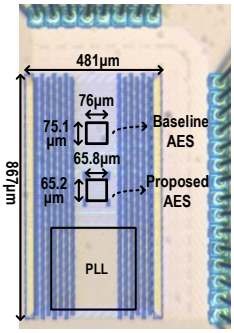


Fig.7. Die Photo

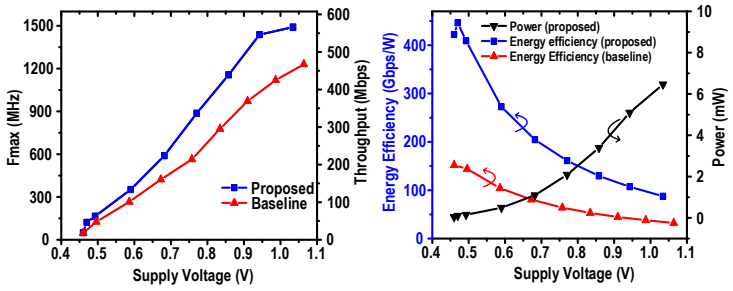
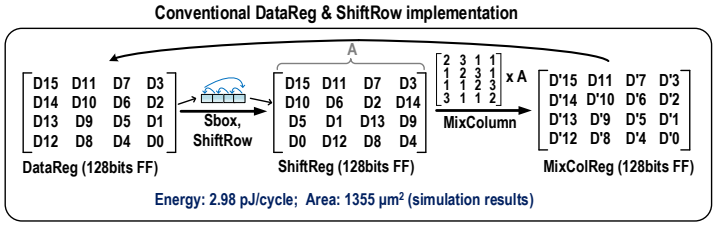
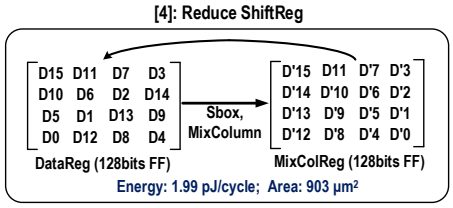
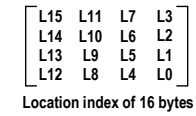


Fig.6. Frequency, throughput, power and energy-efficiency measurements.







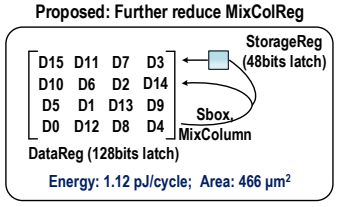


Fig.2. DataReg & ShiftRow implementation comparison.

Table 1. Proposed MixColumn output data & location

4th cycle	Data	D'15	D'14	D'13	D'12
	Location	L15	L2	L5	L8
8th cycle	Data	D'11	D'10	D'9	D'8
	Location	L11	L14	L1	L4
12th cycle	Data	D'7	D'6	D'5	D'4
	Location	L7	L10	L13	L0
16th cycle	Data	D'3	D'2	D'1	D'0
	Location	L3	L6	L9	L12

Data in the byte locations of L0/1/2/4/5/8 need to be delayed and put in StorageReg

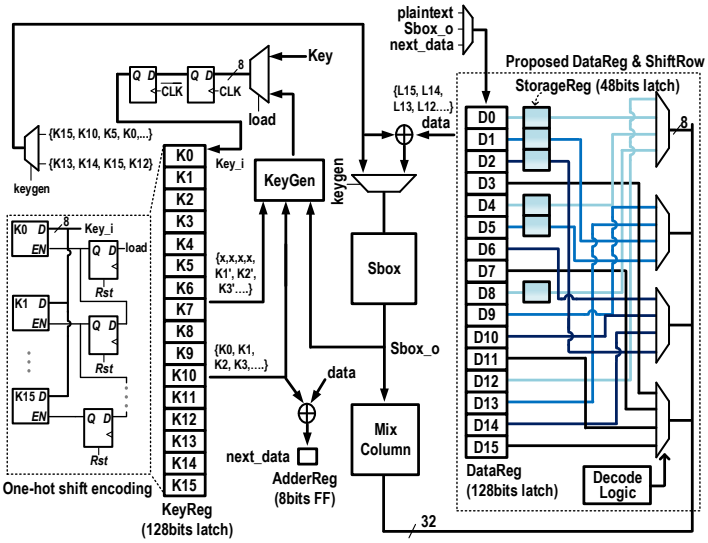


Fig.4. Proposed 8-bit datapath AES accelerator architecture

Table 3. Chip measurement summary and comparison table of AES designs

	EuroMicro'06 [7]	TVLSI'10 [8]	JSSC'11 [3]	JSSC'15 [4]	Proposed
Technology	130nm	130nm	45nm	22nm	40nm
Voltage (V)	Not Reported	0.8 0.75	1.1 0.32	0.9 0.43	0.9 0.47
Power (mW)	17.98 3.9	0.099 0.000692	125 0.409	13 0.45	4.39 0.10
Frequency	290 MHz 130 MHz	12 MHz 100 KHz	2.1 GHz 32 MHz	1.1 GHz 220 MHz	1.3 GHz 122 MHz
Throughput	232 Mbps 104 Mbps	34 Kbps 280 bps	53 Gbps 800 Mbps	432 Mbps 83.6 Mbps	494 Mbps 46.2 Mbps
Energy Efficiency (Gbps/W)	12.9 26.7	0.343 0.405	424 1955	33 186	113 446
Energy/bit (pJ/b)	77.5 37.5	2915 2469	2.36 0.512	31 5.38	8.85 2.24
Number of Gates	3200 3900	5500	Not Reported	1947	2228
Area $(mm)^2$ (Norm to 40nm)	Not Reported	< 1	0.15 (0.119)	0.0022 (0.0073)	0.00429