22.1 Razor II: In Situ Error Detection and Correction for PVT and SER Tolerance

David Blaauw[1], Sudherssen Kalaiselvan[2], Kevin Lai[1], Wei-Hsiang Ma[1], Sanjay Pant[1], Carlos Tokunaga[1], Shidhartha Das[3], David Bull[3]

[1]University of Michigan, Ann Arbor, MI, [2]AMD, Sunnyvale, CA
[3]ARM, Cambridge, United Kingdom

Performance variation due to PVT uncertainty has led to an increased interest in adaptive designs. Traditional methods for adaptive design have used so-called canary circuits that mimic the critical-path delay of the actual design [1-4]. However, canary circuits require safety margins for possible mis-tracking and intra-die PVT variations. They also have difficulty responding to rapidly changing conditions.

To eliminate these margins, the Razor approach [5] introduces in situ delay-error detection and correction where the supply voltage is tuned using the observed error rate. Razor provides energy reduction in two ways: 1) by lowering the supply voltage to the point of first failure (PoFF), all margins due to global and local PVT variations are eliminated. 2) by purposefully operating below the PoFF, the energy gain from a lower supply voltage is optimally traded-off with the overhead of higher error correction activity.

A key finding from our initial Razor (Razor I) measurements is that the error rate at the PoFF is extremely low, ~1 error in 10 million cycles, making the error-correction energy negligible at this operating point. However, it is also found that beyond the PoFF, the error rate increases exponentially at one decade per ~10mV supply voltage. Hence the energy gain from operating substantially below the PoFF is small (~10%) compared to the energy gain from eliminating the PVT margins (~35 to 45%) [5].

We take advantage of these findings and propose a Razor II approach that introduces two components. First, instead of performing both error detection and correction in the FF, Razor II performs only detection in the FF, while correction is performed through architectural replay. This allows significant reduction in the complexity and size of the Razor FF. Since Razor II is intended to operate near the PoFF, the increased overhead from using architectural correction has a negligible impact on the energy efficiency. Second, the Razor II FF naturally detects SER in the logic and registers without additional overhead. Hence, the processor provides both low-energy operation through dynamic supply adaptation as well as SER tolerance, as demonstrated by radiation tests.

The Razor II FF (Fig. 22.1.1) uses a single latch combined with a transition detector (TD) controlled by a detection clock (DC). While the implementation uses a latch, it operates as a positive–edge-triggered FF. If the data transitions before the rising clock edge, the short negative pulse on DC suppresses the TD and no error is registered (Fig. 22.1.2). However, if the input data transitions after the rising clock edge, during transparency, the transition of latch node N occurs when TD is enabled and results in assertion of the error signal and instruction roll-back. Hence, late arriving signals are flagged as an error which enforces FF based operation of the design. In contrast, the Razor I FF detects late-arriving data by comparing the FF state with that of a latch with a delayed clock. In total it consists of three latches, a comparator and a meta-stability detector. By using a latch instead of a FF, the Razor II FF has a slightly improved clk-to-q delay compared to Razor I and 0ps setup time at the positive edge. It uses 47 transistors (Razor I FF uses 76) if the DC is generated internally and 39 if the DC generation is shared between several FFs. The power overhead for a Razor II FF as compared to a conventional FF for a 10% activity factor is 28.5%. The total power consumption overhead due to inserting Razor II FFs in the processor is 1.2%.

The Razor II error-detection window lies between the rising edge of DC and the falling edge of CLK and is controlled with the duty cycle of CLK. Increasing the detection window increases the available timing speculation but also requires additional buffering to address hold time constraints. Timing-critical FFs have a clock with a 40% duty cycle, resulting in a 25-F04 detection window while non-critical FFs have a 13% clock duty cycle to reduce buffer insertion. A total of 1924 buffers are added to meet hold time constraints, which adds a 1.3% power consumption overhead. Since the supply voltage is never lowered to the point where the latch transitions at the falling clock edge, meta-stability of the latch is avoided. The possibility that the TD becomes meta-stable is mitigated by double-latching the error signal, which does not incur a performance penalty during normal operation.

Since the latch node is monitored by the TD during both clock phases, SER strikes at the latch node or propagated from the logic to the latch are automatically detected without additional overhead (Fig. 22.1.2). A strike during the low phase of DC when the TD is disabled is either benign, if the signal returns to its valid value before TD is re-enabled, or is detected as an error. Hence, the transparency window must extend beyond the low phase of DC to avoid latching of SER strikes before they are detected by the re-enabled TD.

Razor II is incorporated in a 64-bit, 7-stage Alpha processor in 0.13μm CMOS. The architecture (Fig. 22.1.3) is divided into a pipeline with speculative state protected using Razor II FFs, and a non-speculative memory and register file protected by ECC or triple-module redundancy (TMR). The error signals of all Razor II FFs in each pipeline stage are ORed together and the result is propagated and ORed with that of the next stage. The 7th stage is designed to be non-timing critical to stabilize the pipeline state. It also encodes the speculative state before it is passed to the RF or SRAM. In the event of an error, the pipeline is flushed and the failing instruction is re-executed. In case of repeatedly failing instructions, the error controller reduces the clock frequency by half for 8 cycles. The error rate is kept at 0.04% using an off-chip controller. However, since failing instructions are guaranteed to complete, control can also be performed in software on the chip itself.

Figure 22.1.4 shows the measured energy dissipation for 3 die when operating at 0.04% error rate. Gains are 33.1 to 37.5% compared to the energy when the supply voltage is elevated to ensure correct operation for all 31 fabricated die at 85°C with 10% margin for wearout, supply fluctuation and safety. Figure 22.1.5 shows the measured failure rate and energy efficiency as the supply voltage is lowered past the PoFF. The energy optimal point is reached at 0.04% error rate and incurs an IPC penalty of 0.2%. As expected, the gain from operating below the PoFF is small compared to that from eliminating margins and is reduced compared to Razor I due to the higher correction overhead.

Figure 22.1.6 shows the radiation setup and the different test cases. In Test 1, the test-chip is exposed to SER with error detection disabled and as expected the final program result is incorrect. When error detection is enabled (Test 2) the processor is able to detect and correct the SER-induced errors. This is verified for different operating voltages (0.8 to 1.0V). The soft-error rate is recorded for memory and pipeline elements. The test-chip continues to operate correctly when the frequency of operation is increased beyond PoFF causing delay errors in addition to SER (Test 3).

Acknowledgements:
We are thankful to Prof. Vijaykrishnan Narayanan and Ramakrishnan Krishnan at Penn. State University for their help with the radiation tests.

References:
[1] A. Drake, R. Senger, H. Deogun et al., "A Distributed Critical-Path Timing Monitor for a 65nm High-Performance Microprocessor," *ISSCC Dig. Tech. Papers*, Feb. 2007.
[2] T. Burd, T. Pering, A. Stratakos, R. Brodersen, "A Dynamic Voltage Scaled Microprocessor System," *IEEE J. Solid-State Circuits*, vol. 35, no. 11, Nov. 2000.
[3] M. Nakai, S. Akui, K. Seno et al., "Dynamic Voltage and Frequency Management for a Low-Power Embedded Microprocessor," *IEEE J. Solid-State Circuits*, vol. 40, no. 1, Jan. 2005.
[4] K. Nowka, G. Carpenter, E. MacDonald et al., "A 32-bit PowerPC System-on-a-chip With Support for Dynamic Voltage Scaling and Dynamic Frequency Scaling," *IEEE J. Solid-State Circuits*, vol. 37, no. 11, Nov. 2002.
[5] S. Das, D. Roberts, S. Lee, S. Pant et al., "A Self-Tuning DVS Processor Using Delay-Error Detection and Correction," *IEEE J. Solid-State Circuits*, vol. 41, no. 4, Apr. 2006.

 978-1-4244-2011-7/08/$25.00

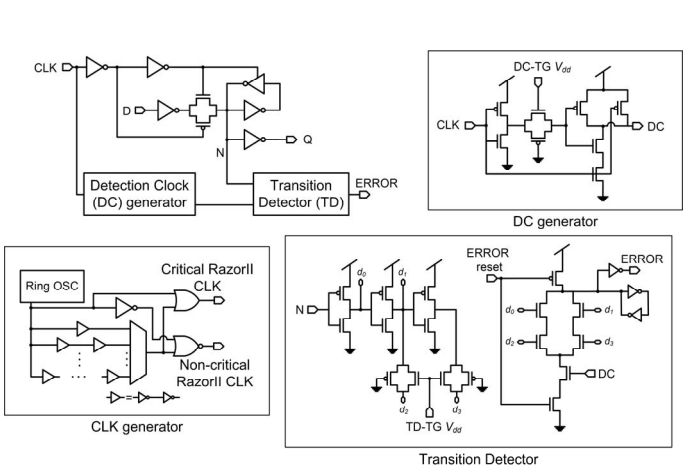


Figure 22.1.1: Schematic of the Razor II FF latch, clock generator, transition detector (TD) and detection clock (DC) generator.

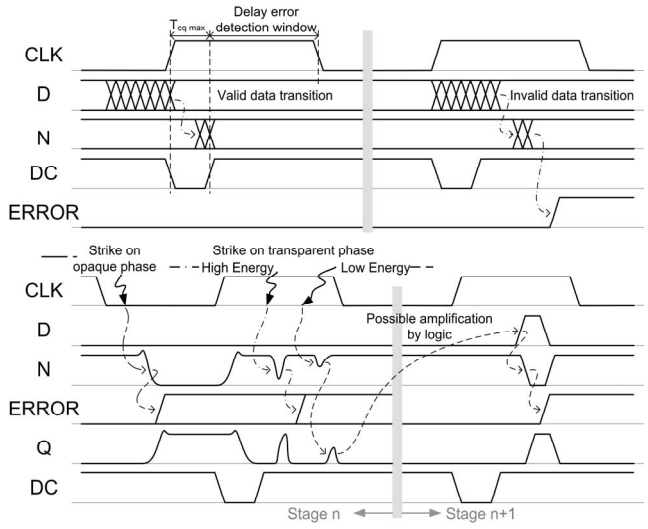


Figure 22.1.2: Razor II FF timing and error generation for PVT variation induced delay error and SER detection.

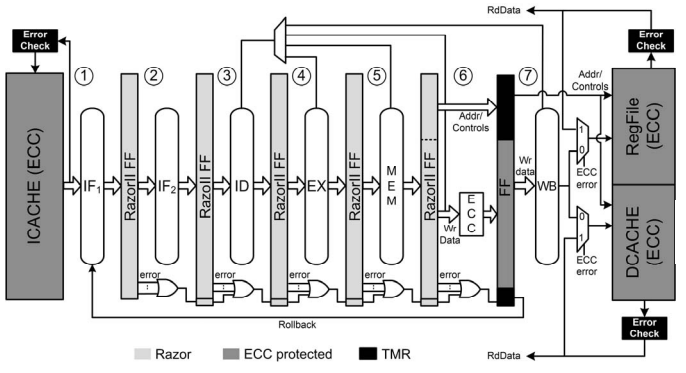


Figure 22.1.3: Razor II enabled 64-bit Alpha processor architecture.

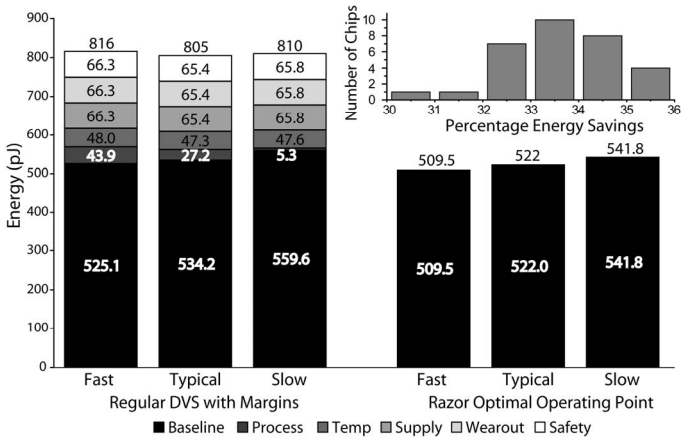


Figure 22.1.4: Measured Razor II energy consumption and distribution of energy savings.

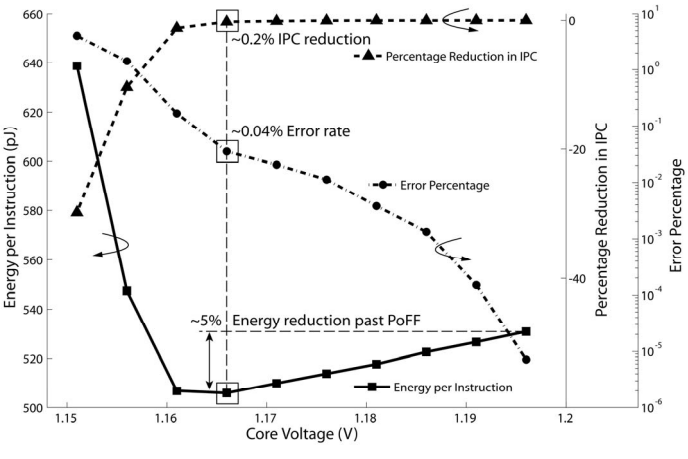
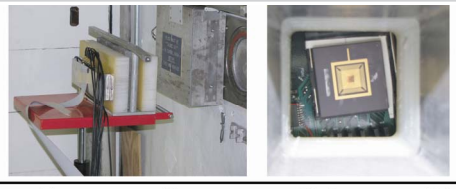


Figure 22.1.5: Measured energy per instruction and error rate for the Razor II processor.

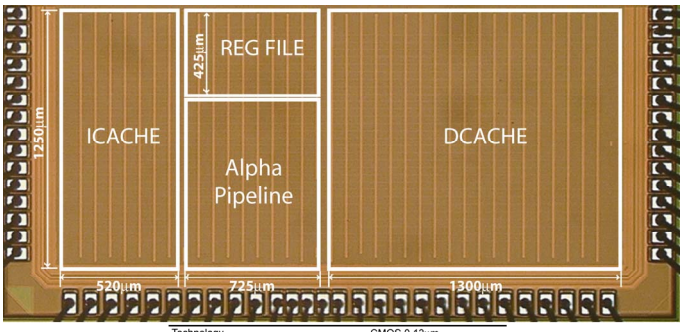


Radiation Test Description	Error detection (TD enabled)	Vdd (V)	Memory errors (ECC errors/State element/1000hr)	Pipeline errors (Errors/State element/1000hr)	Delay Error rate > 0	Correct program execution
Test 1: No timing failures, no error detection	Off	0.8	NA*	NA*	No	No
Test 2: No timing failures, with error detection	On	1.0	1.84	2.91	No	Yes
		0.9	0.91	3.23		
		0.8	2.04	3.87		
Test 3: Both SER and timing failures, with error detection	On	0.8	33M errors**		Yes	Yes

* When TD is disabled the counter of errors in the ECC and RazorII FFs is disabled
** Beyond PoFF (Delay error rate > 0), timing errors overwhelm SER, but they both are detected and corrected by the RazorII mechanisms.

Figure 22.1.6: Soft error radiation setup at Penn. State U. and measured radiation induced errors. Correct operation was verified with simultaneous delay errors and SER upsets.

Continued on Page 622



Technology	CMOS 0.13µm
Dimensions	2700µm x 1250µm
ICACHE	4096 x [32+17(ECC)] bits
DCACHE	2048 x [64+22(ECC)] bits
No. of critical RazorII FF	121 of 826
No. of non-critical RazorII FF	705 of 826
No. of buffers added for hold time fixing	1924
Buffer insertion power overhead	1.3%
RazorII FF power overhead	1.2%
Total no. of transistors	260k
Operation voltage	0.8-1.2V
Frequency	185MHz @ 1.2V
Power @ 185MHz	94.3mW

Figure 22.1.7: Die micrograph and performance summary.

 978-1-4244-2011-7/08/$25.00