Power Minimization of Pipeline Architecture through 1-Cycle Error Correction and Voltage Scaling

Insup Shin¹, Jae-Joon Kim² and Youngsoo Shin¹

¹Dept. of Electrical Engineering, KAIST, KOREA ²Dept. of Creative IT Engineering, POSTECH, KOREA

Outline

- Introduction
- Motivation
- 1-cycle error correction method
 - Proposed architecture
 - Clock gating signal propagation
 - Extension to general pipeline architectures
- Experimental results
- Conclusion

Timing Speculation Based Voltage Scaling

- Voltage scaling is best way to reduce power
 - Switching power $\propto V_{dd}^2$, subthreshold leakage $\propto V_{dd}^3$, gate leakage $\propto V_{dd}^4$
 - Reducing V_{dd} comes at cost of reduced circuit speed
- Timing speculation allows deeper V_{dd} reduction by eliminating timing margin
 - Required timing margin becomes substantial in today's nanometer design
 - Error correction incurs timing penalty

Motivation

- Minimizing the number of error correction cycles is important to achieve deeper voltage scaling
 - Previous methods focus on reducing timing penalty per error correction
- Voltage reduction below critical operating point causes massive errors
 - There is no error correction method considering multiple error correction

Our method

- Minimize timing penalty per error correction
- Correct multiple errors simultaneously

Previous Works

Instruction replay

• 3N-cycle penalty (N: number of pipeline stages)

Counterflow pipelining

- 2k-cycle penalty (k: order of stage which detects error)
- Bubble Razor^[ISSCC 2012]
 - 1-cycle penalty
 - Only applicable to two-phase transparent latch based designs

I-CTEC^[ISLPED 2013]

- 1-cycle penalty
- Limitation in handling massive errors

Proposed Architecture

 To alter clock toward shadow latch in such a way that shadow latch opens after main latch closes



Conceptual schematic of Razor latch





Proposed Architecture

 Shadow latch can send previous correct data to main latch and also capture new input data during restore cycle



Circuit level schematic

Clock Gating Signal Propagation

- Clock gating (CG) signal is propagated to output stages from stage where error occurred
 - Stall signal is issued to pipeline when CG signal reaches last stage





Error-Free Mode

No late timing error occurs in this mode

- Stage gets into this mode once error occurs at the stage
- Example: stage B operates in error-free mode from cycle 2 to cycle 4



Multiple Timing Errors

 Multiple errors at same stage are corrected with only 1-cycle





1-CTEC method (2 cycle penalty)



Our method (1 cycle penalty)

Multiple Timing Errors

 Multiple errors at different stages can be corrected simultaneously





General Pipeline Architecture

Multiple fan-in/fan-out structure

Problem occurs when not all input stages sent CG signal

Loop structure

• Key challenge: to prevent indefinite looping of CG propagation



Multiple Fan-In/Fan-Out Case

Problem

- Data loss at a multiple fan-in stage when not all input stages sent CG signal
- Example: instruction i2 is lost at stage D in cycle 2



Multiple Fan-In/Fan-Out Case

Solution

 Generate virtual errors at all the stages that did not send CG signals to multiple fan-in stages



Virtual Error (VE) Signal

Modified propagation algorithm

 If stage receives CG signal from any of its input stages, send VE to all of its input stages in the same cycle



Loop Case

VE signal prevents infinite looping of CG propagation

- VE is generated regardless of location where error happens
- Propagation of CG stops at stage where virtual error occurred

Three examples for verification

- 1) Error occurs before loop
- 2) Error occurs in loop
- 3) Error occurs after loop

Loop Case: Examples

E

Error occurs before loop

Error occurs in loop

Error occurs after loop

Experimental Results

Setting

- Six pipelined circuits with 45-nm open cell library
 - Two different number of pipeline stages (5 and 10)
 - c1908, c3540, and c6288 from ISCAS'85 were assumed for each pipeline stage
- Pulse width of latch: 105 ps (main), 400 ps (shadow)
- Extra delay buffers were inserted to fix hold violations
- Applied 100 random vectors to each circuits to determine its throughput and energy dissipation using fast SPICE simulation

When Target Throughput = 0.9

# Stages	Base	Count	erflow	1-CTEC		Ours	
	circuit	Voltage [V]	Energy [pJ]	Voltage [V]	Energy [pJ]	Voltage [V]	Energy [pJ]
5	c1908	0.92	783	0.84	707	0.84	716
	c3540	0.94	2107	0.88	1816	0.86	1751
	c6288	0.98	5108	0.90	4307	0.90	4221
	Average		1.16		1.00		0.98
10	c1908	0.94	1591	0.88	1362	0.86	1316
	c3540	0.96	4931	0.90	3991	0.88	3692
	c6288	0.98	10449	0.90	8489	0.88	7596
	Average		1.21		1.00		0.93

- Normalized energy dissipation of counterflow pipelining increases with more pipeline stages
 - Timing penalty per error correction depends on # of pipeline stages
 - Multiple errors cannot be corrected simultaneously

When Target Throughput = 0.7

# Stages	Base	Counterflow		1-CTEC		Ours	
	circuit	Voltage [V]	Energy [pJ]	Voltage [V]	Energy [pJ]	Voltage [V]	Energy [pJ]
5	c1908	0.90	751	0.78	614	0.76	576
	c3540	0.92	1912	0.82	1594	0.80	1515
	c6288	0.98	5108	0.86	3994	0.84	3706
	Average		1.23		1.00		0.94
10	c1908	0.92	1540	0.88	1254	0.80	1153
	c3540	0.96	4931	0.90	3411	0.80	2906
	c6288	0.98	10449	0.90	7457	0.84	6625
	Average		1.36		1.00		0.89

Energy reduction (compared to 1-CTEC)

- 5-stage: 2% (@ 0.9) and 6% (@ 0.7)
- 10-stage: 7% (@ 0.9) and 11% (@ 0.7)
- # of cycles that each stage runs in error-free mode increases with # of pipeline stages

Conclusion

 Presented 1-cycle error correction method that can handle massive errors

• Multiple errors can be corrected simultaneously

Experiments (compared to 1-CTEC)

 2~6% energy reduction for 5-stage pipeline and 7~11% energy reduction for 10-stage pipeline

Q & A

Thank you for your attention

Design Technology Lab., KAIST

Insup Shin (isshin@dtlab.kaist.ac.kr)