University of Central Florida

Karlsruher Institut für Technologie

Computing Frontiers 2016

Area-Energy Tradeoffs of Logic Wear-Leveling for BTI-induced Aging

Rizwan A. Ashraf, Navid Khoshavi, Ahmad Alzahrani, Ronald F. DeMara
Department of Electrical and Computer Engineering
University of Central Florida

Saman Kiamehr, Mehdi B. Tahoori
Department of Computer Science
Karlsruhe Institute of Technology

May 2016

Logic Wear-Leveling for Anti-aging

Logic Wear-Leveling (LWL):

- a *post-fabrication self-adapting* circuit-level approach to mitigate timing degradations.

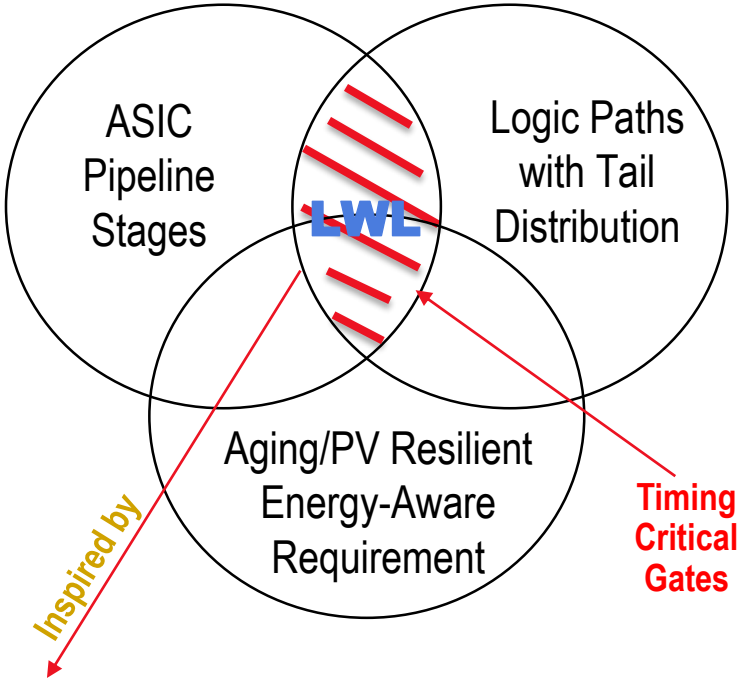
Research Objective:

- Mitigate *Transistor Aging* due to BTI and HCI thus reducing the energy wastage due to conservative selection of guardbands**.**

Targeting Aging-critical Elements:

- aging-critical logic portions of the circuit are targeted for protection → minimal overhead
- power-gating is effective in reducing BTI and HCI. Switching activity (p) effects the shift in V_{th}

$$\Delta V_{th}(t) \propto (pt)^n$$



How to leverage Dark Silicon???

area of chip which cannot be operated due to constrains of simultaneous operation of all transistors which are:

- cooling cost
- power cost

Potential Benefits of Logic Wear-Leveling

Need addressed	Approach	Benefit
Transistor Aging	Power-gating of critical elements	Reduced lifetime delay degradation through stress reduction
Energy Consumption	Reduced voltage guard bands	Low energy requirement with narrower margins for longer periods
	De-emphasized role of voltage regulators	Circumvent conversion inefficiencies and switching losses
	Selective Redundancy	Power-gating lowers the leakage energy overheads

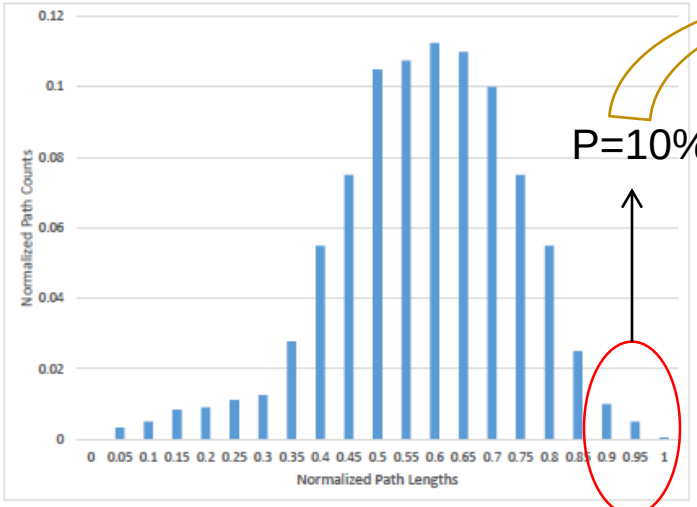
Related Works

Technique	Anti-Aging Strategy	Design Requirements/ Parameters	Adaptability Characteristics/ Degree	Overheads		
				Throughput	Power	Area
Worst-case Design						
VM, FM	Static Margin	MD-RoD/ ΔV_{DD}, $\Delta F_{nominal}$	None	FM: High	VM: High (Dynamic & Leakage)	None
Gate-Sizing	Static Margin	MD-RoD; Extended Std. Lib.; Multi-obj. Opt./ $\Delta\beta_i$, ∀ gates *i*	None	None	Medium (Dynamic & Leakage)	Low (Gate-level)
Re-Synthesis	Static Margin	MD-RoD annotated Std. Lib.; Aging-aware Synthesis/ $\Delta\beta i$, $\Delta V_{th,i}$ ∀ gates *i*	None	None	Low-Medium (Dynamic & Leakage)	Low (Gate-level)
Dynamic Operating Conditions						
DVFS	Dynamic Margin	Timing Sensors; Feedback Control/ $\Delta V_{DD}(t)$, $\Delta F(t)$, $\Delta V_{bb}(t)$	Yes/ Fully Autonomous	Low	Medium (Dynamic & Leakage)	Medium (On-chip VR & sensors)
SVS	Dynamic Margin	MD-RoD/ $\Delta V_{DD}(t + \Delta t_{step})$	Yes/ t_{step}	None	Medium (Dynamic & Leakage)	Medium (On-chip VR)
GNOMO	Static Margin + Power-Gating	MD-RoD/ ($V_{DD,g}$, t_{idle})	None	Medium (Workload Dependent)	Medium (Dynamic & Leakage)	None
Adaptive Resource Management						
SD	Proactive Mngt. + Power-Gating	Modular Redundancy/ Sleep Interval	Yes/ Sleep Interval	None	High (Leakage)	High (Module-level)
ITL schemes	Proactive Mngt. + Power-Gating	Exploit App. Redundancy/ Idle time	Yes/ Task Scheduling	Medium (Workload Dependent)	None	None
LWL	Proactive Fine-Grain Mngt. + Power-Gating	CPRT/ Sleep Interval	Yes/ Sleep Interval	None	Minimal (Leakage)	Low (Gate-level)
RR	Proactive Fine-Grain Mngt. + Power-Gating	Timing Sensors; Feedback Control; CPRT/ ERT%	Yes/ Fully Autonomous	None	Minimal (Leakage)	Low (Gate-level & sensors)

Proposed herein

Aging-Sensitive Logic Domains

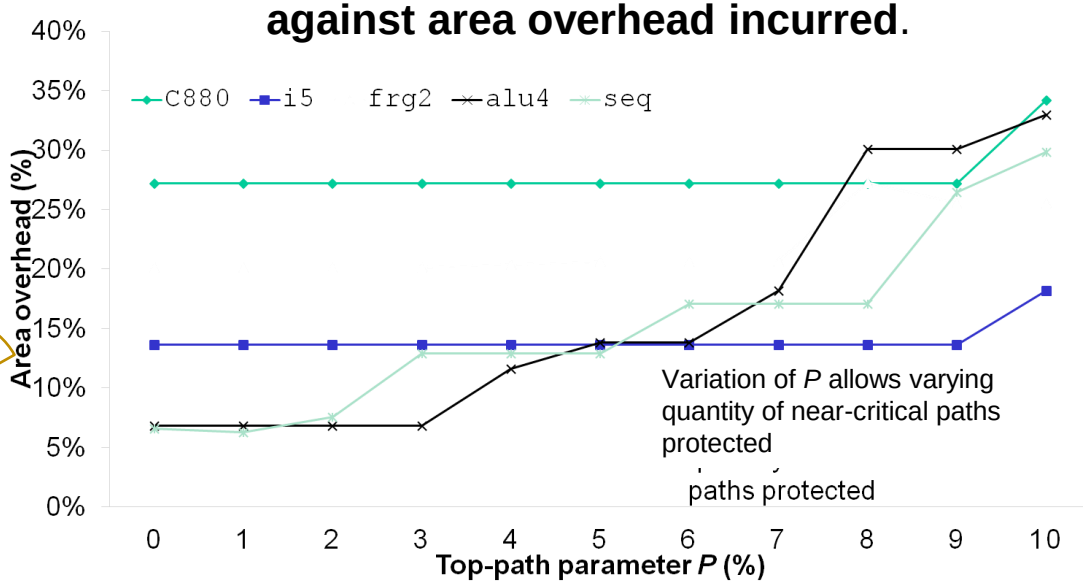
1) Path distribution of OpenSPARC ALU



P=10%

[G. Hoang et al."Exploring circuit timing-aware language and compilation," *SIGPLAN Not.,* vol. 47, no. 4, pp. 345–356, Mar. 2011.]

2) Parameter *P* can be traded-off against area overhead incurred.

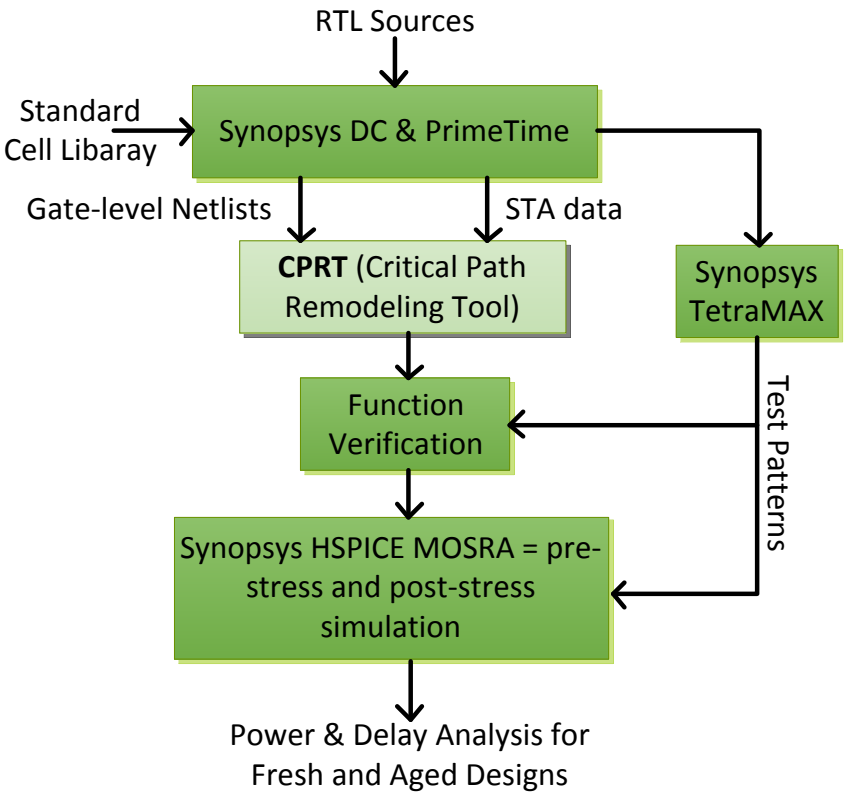


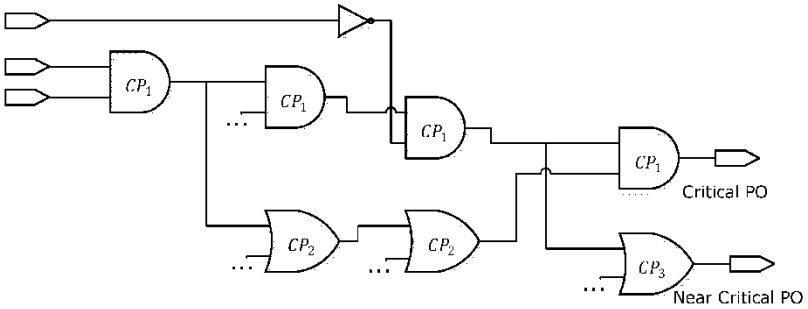
3) LWL covers logic paths having delays:

$$[D_{critical}(t) * (100\% - P), D_{critical}(t)]$$

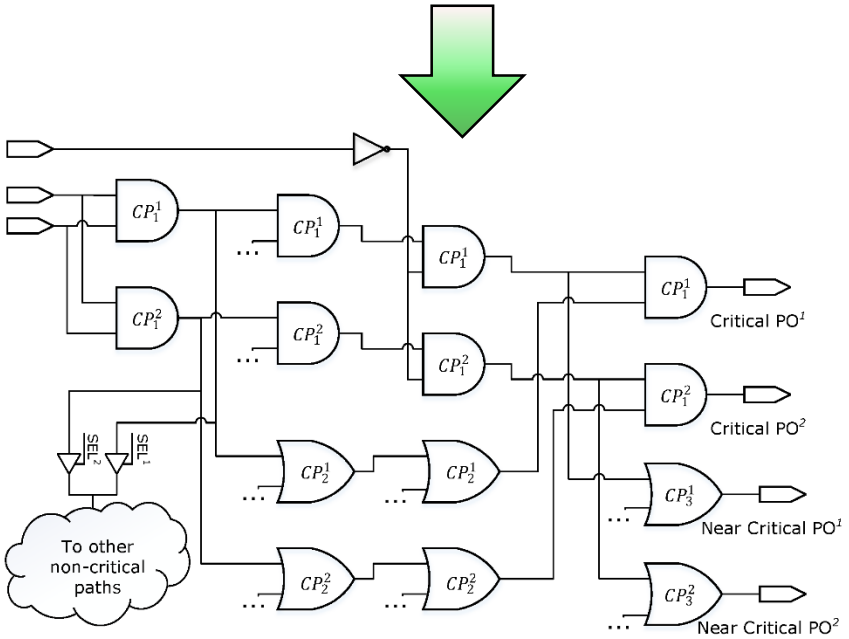
- P is top-path parameter
- based on near critical paths due to aging and/or PV effects
- *P=10%* used herein

Critical Path Remodeling Tool (CPRT)





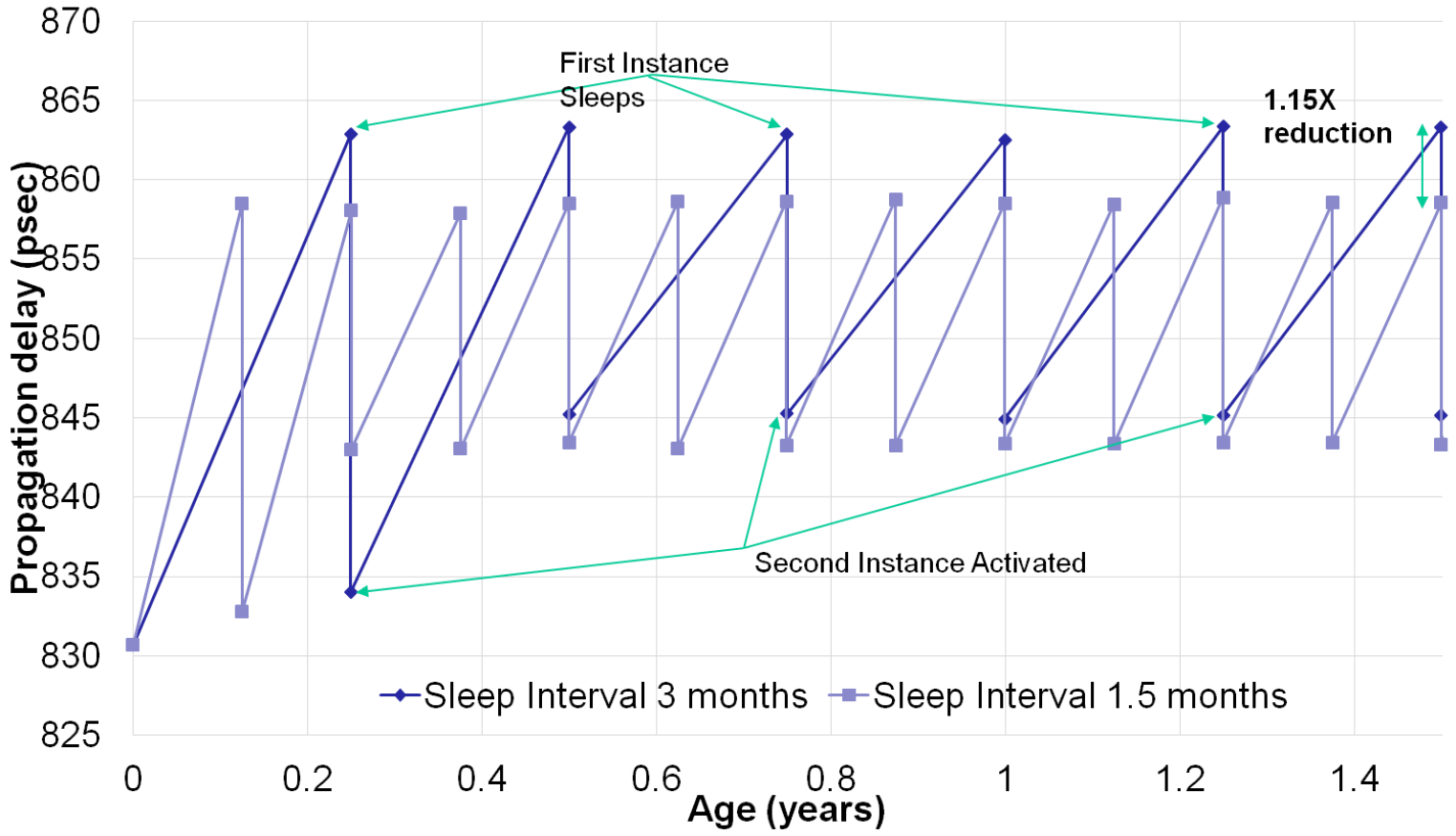
Circuit Before Critical Path Replication



Circuit After Critical Path Replication

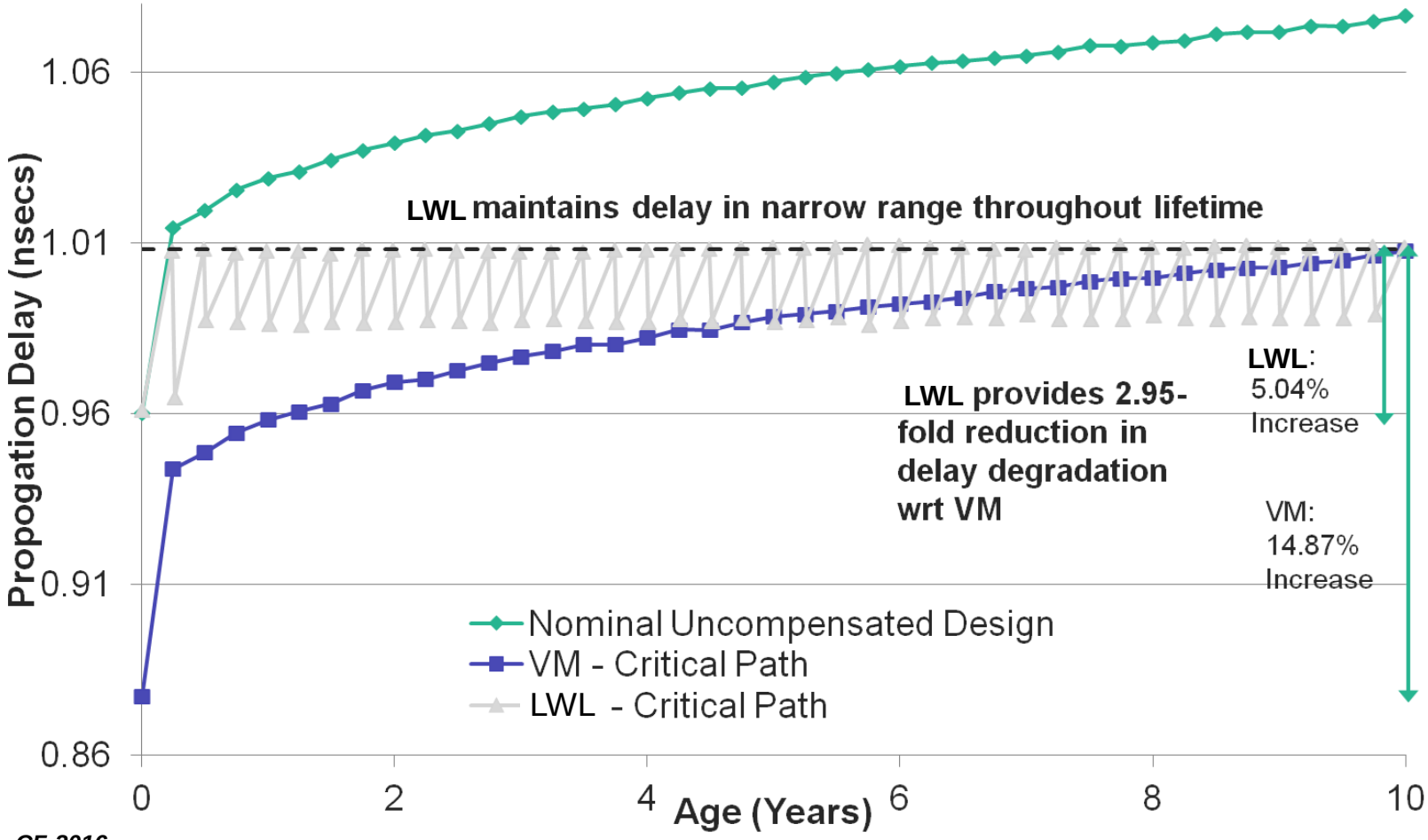
Effect of Switching Interval

LWL: Proactive resource switching to balance stress among replicated components.



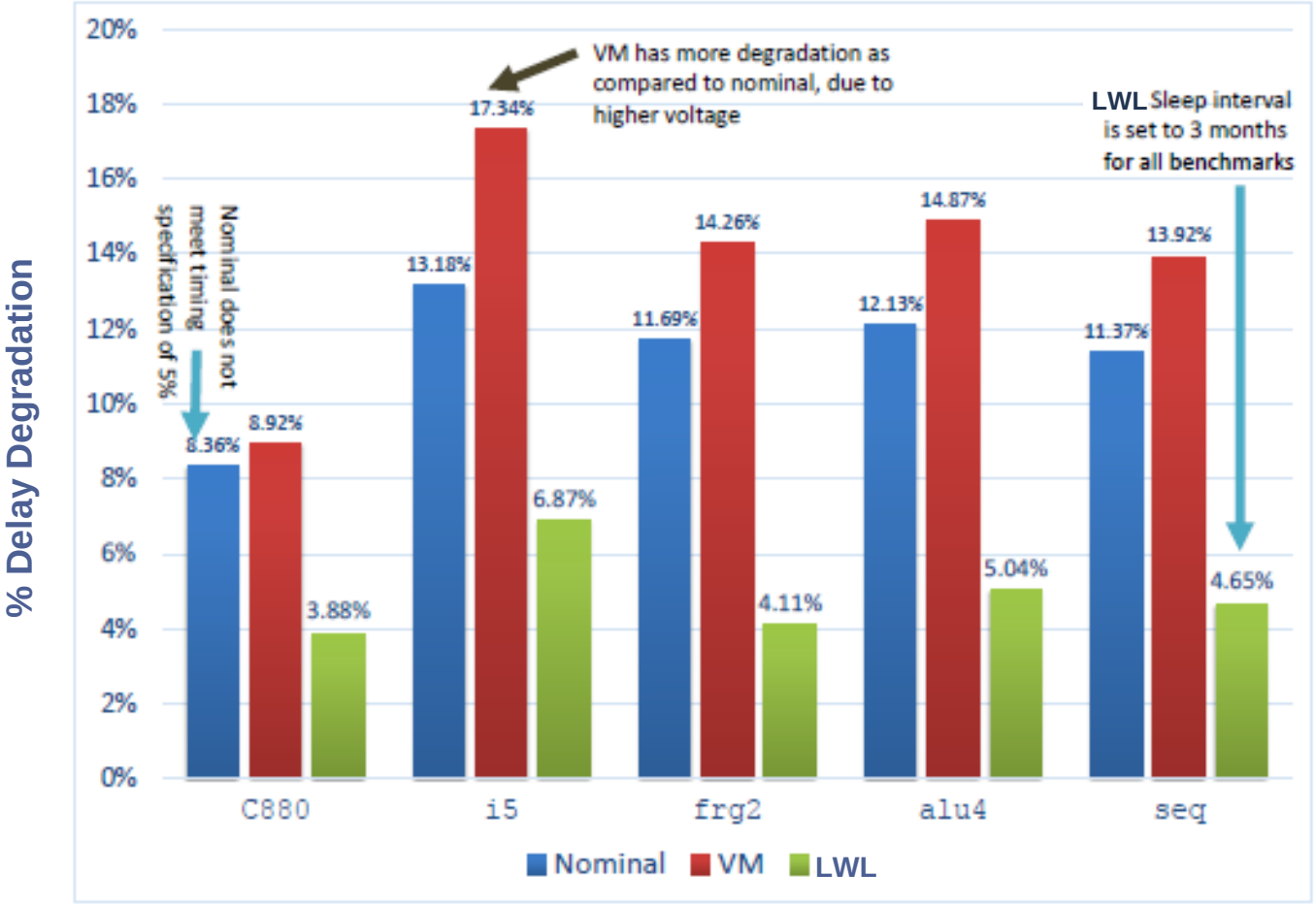
LWL Reduction of Delay Degradation [ALU4]

NanGate Library based on 45nm Predictive Technology Model is used.
Built-in models for BTI and HCI are utilized for HSPICE simulations



LWL Reduction of Delay Degradation

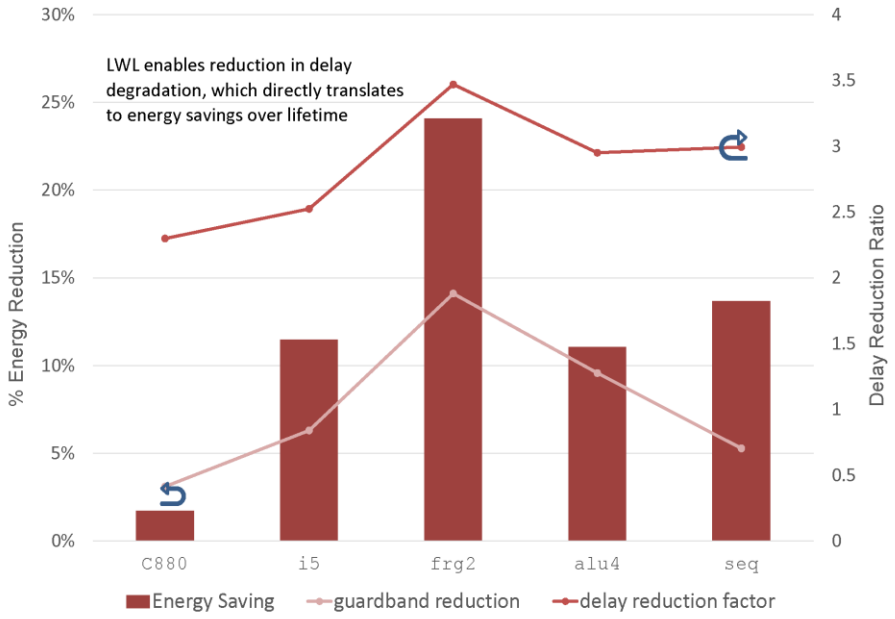
Delay degradation (% inc in delay from initial time) over a lifetime of 10 years.

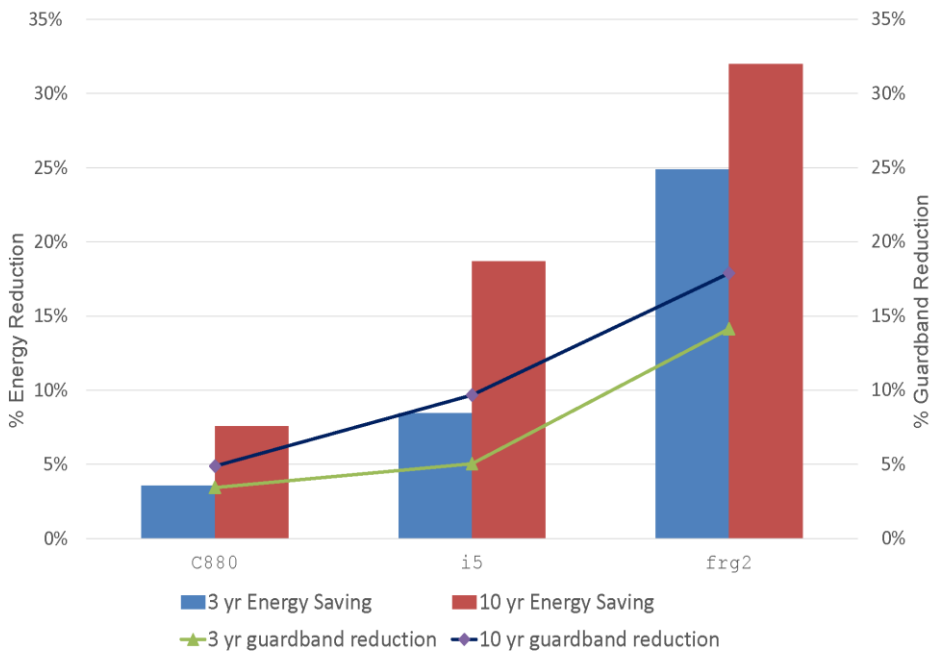


LWL Energy Savings

Delay reduction factors are calculated by taking the ratio of VM's degradation and LWL's degradation.

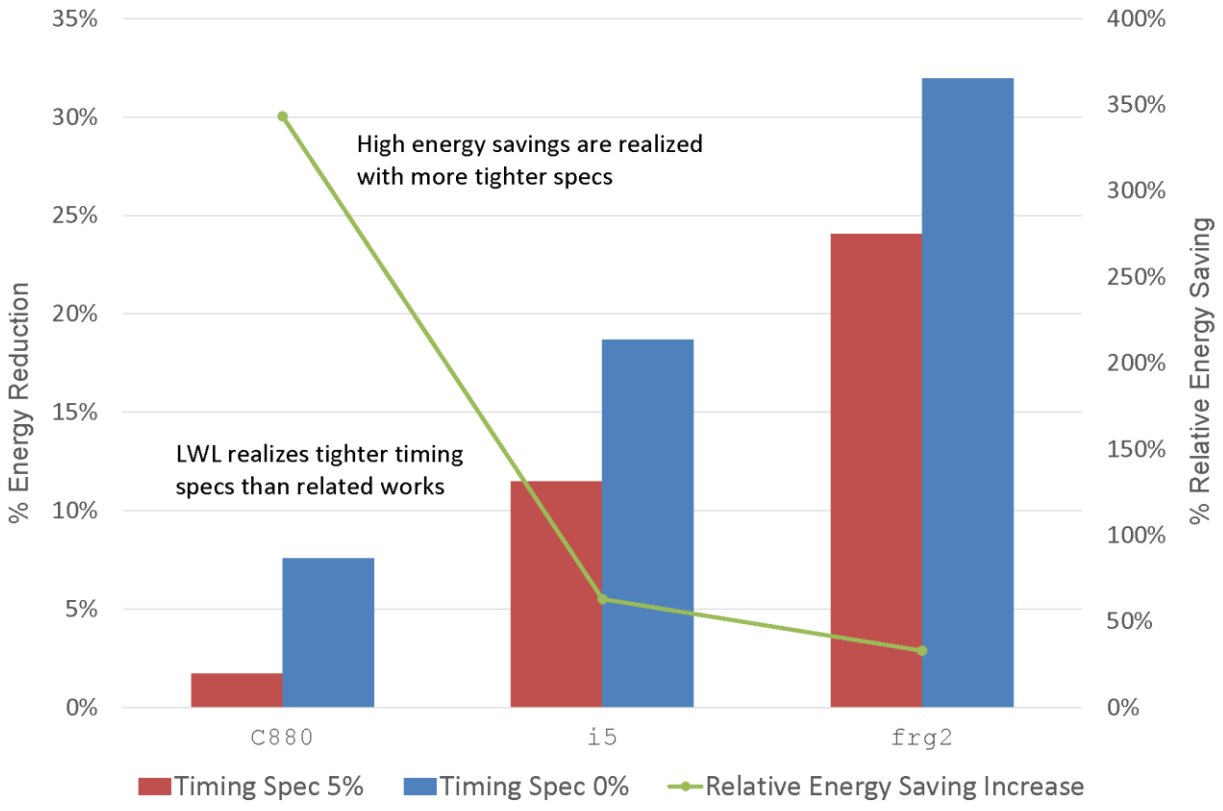
The reduction factors are seen to correlate with the energy savings achieved.





Implications of Tighter Timing Specifications

Implications of reducing delay headroom towards that of a baseline circuit operating at nominal voltage



Area Overhead w.r.t. Related Works

Case-study:

- 45nm-based Intel Penryn multicore processor
- based on [9], 46.1% of total die area is considered as Core-area

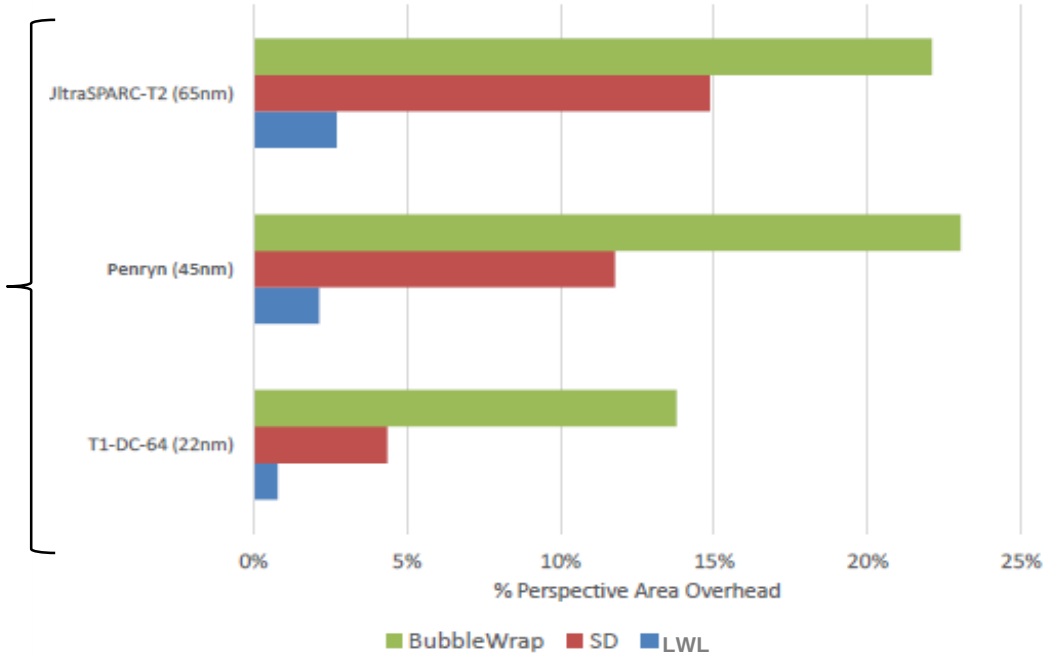
Execution unit:

- only 39.03% of a single core area is occupied by execution unit
- 65.5% of execution unit can be considered as aging-critical portion
  → aging-critical portion = 11.7% of die

Area cost:

- SD and BubbleWrap: 4.36% to 23.03%
- LWL: 0.79% to 2.7%

- **For BubbleWrap:** half of the cores are designated as *expandable*
- **For SD:** aging-sensitive logic is replicated
- **For LWL**: utilization model with P = 7% of paths in arithmetic unit



Summary of Approach

- LWL provides an **adaptive** technique for anti-aging using a **spatial redundancy** and **power-gating** to enable BTI recovery
- **Accurate aging modeling → unnecessary**

  as circuit degradation is determined using operational conditions
- LWL is shown to successfully **reduce the guardband** with delay reductions ranging from 1.92-fold to 2.84-fold over nominal values with 5% timing margin
- Favorable **energy savings** as high as 31.98% with 0% timing margin are obtained due to further reduction of operating voltage
- Area cost is traded for energy reduction by minimizing ΔV_{th}
- **Dark silicon** effect has inspired us to **allocate unused space to reduce aging impact** in critical region of circuit
  - reduce energy consumption,
  - avoid need for precise aging models, and
  - intrinsically accommodate the process variation within the actual as-built circuits

References

[1] X. Bai et al., “Uncertainty-aware circuit optimization,” in 39th Design Automation Conference, 2002.

[2] A. Calimera et al., “NBTI-aware power gating for concurrent leakage and aging optimization,” Low Power Electronics and Design, International Symposium on, pp. 127–132, 2009.

[3] T.-B. Chan et al., “On the efficacy of NBTI mitigation techniques,” in Design, Automation Test in Europe Conference Exhibition (DATE), 2011.

[4] J. Chen et al., “Efficient selection and analysis of critical-reliability paths and gates,” in Proceedings of the Great Lakes Symposium on VLSI, ser. GLSVLSI ’12, 2012, pp. 45–50.

[5] M. Ebrahimi et al., “Aging-aware logic synthesis,” in Proceedings of the International Conference on Computer-Aided Design, ser. ICCAD ’13, 2013, pp. 61–68.

[6] F. Firouzi et al., “Representative critical-path selection for aging-induced delay monitoring,” in Test Conference (ITC), 2013 IEEE International, Sept 2013, pp. 1–10.

[7] G. Hoang et al., “Exploring circuit timing-aware language and compilation,” SIGPLAN Not., vol. 47, no. 4, pp. 345–356, 2011.

[8] S. Kumar et al., “Adaptive techniques for overcoming performance degradation due to aging in CMOS circuits,” Very Large Scale Integration (VLSI) Systems, IEEE Transactions on, vol. 19, no. 4, pp. 603–614, 2011.

[9] S. Li et al., “McPAT: An integrated power, area, and timing modeling framework for multicore and manycore architectures,” in 42nd IEEE/ACM International Symposium on Microarchitecture, Dec 2009.

[10] F. Oboril et al., “Extratime: Modeling and analysis of wearout due to transistor aging at microarchitecture-level,” in 42nd IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), June 2012.

[11] J. Shin et al., “A proactive wearout recovery approach for exploiting microarchitectural redundancy to extend cache sram lifetime,” in 35th International Symposium on Computer Architecture, ISCA ’08, June 2008, pp. 353–362.

[12] Y. Shin et al., “Power gating: Circuits, design methodologies, and best practice for standard-cell VLSI designs,” ACM Trans. Des. Autom. Electron. Syst., vol. 15, no. 4, 2010.

[13] J. Srinivasan et al., “Exploiting structural duplication for lifetime reliability enhancement,” in 32nd International Symposium on Computer Architecture, ISCA ’05, June 2005.

[14] D. Sylvester et al., “Computer-aided design for low-power robust computing in nanoscale CMOS,” Proceedings of the IEEE, vol. 95, no. 3, pp. 507–529, March 2007.

[15] M. Taylor, “A landscape of the new dark silicon design regime,” Micro, IEEE, vol. 33, no. 5, pp. 8–19, 2013.

[16] B. Tudor et al., “MOS Device Aging Analysis with HSPICE and CustomSim,” Synopsys, Tech. Rep., 08 2011.

References

[17] B. Vaidyanathan et al., “Technology scaling effect on the relative impact of NBTI and process variation on the reliability of digital circuits,” Device and Materials Reliability, IEEE Transactions on, vol. 12, no. 2, pp. 428–436, 2012.

[18] J. Velamala et al., “Physics matters: Statistical aging prediction under trapping/detrapping,” in 49th ACM/EDAC/IEEE Design Automation Conference, 2012.

[19] W. Wang et al., “The impact of NBTI effect on combinational circuit: Modeling, simulation, and analysis,” Very Large Scale Integration Systems, IEEE Trans. on, vol. 18, no. 2, 2010.

[20] K.-C. Wu et al., “Analysis and mitigation of NBTI-induced performance degradation for power-gated circuits,” in Intl. Sym. on Low Power Electronics and Design, 2011.

[21] X. Yang et al., “Combating NBTI degradation via gate sizing,” in 8th Intl. Sym. on Quality Electronic Design, 2007.

[22] W. Zhao et al., “New generation of predictive technology model for sub-45nm design exploration,” in Proceedings of the 7th International Symposium on Quality Electronic Design, 2006.