

Simulating Wear-out Effects of Asymmetric Multicores at the Architecture Level

Nikos Foutris Christos Kotselidis Mikel Luján

The University of Manchester, Department of Computer Science, United Kingdom
{first.surname}@manchester.ac.uk

Abstract - As the silicon industry moves into deep nanoscale technologies, preserving Mean Time to Failure at acceptable levels becomes a first-order challenge. The operational stress, along with the inefficient power dissipation and the unsustainable thermal thresholds increase the wear-induced failures. As a result, faster wear-out leads to earlier performance degradation with eventual device breakdown. Furthermore, the proliferation of asymmetric multicores is tightly coupled with an increasing susceptibility to variable wear-out rate within the components of processors. This paper investigates the reliability boundaries of asymmetric multicores, which span from embedded systems to high performance computing domains, by performing a continuous-operation reliability assessment. As our experimental analysis illustrates, the variation between the least and the most aged hardware resource equals to 2.6 years. Motivated by this finding, we show that an MTTF-aware, asymmetric configuration prolongs its lifetime by 21%.

I. INTRODUCTION

In recent years, semiconductor technology has been delivering continuous increases in performance and functionality. This reality, however, is predicted to undertake a radical shift as ITRS [16] indicates that a 10-fold decrease in wear-out rate will be required to maintain current design lifetimes in the forthcoming years. That trend can dramatically impact timing guard-bands, noting that current practices already lead to a loss of up to 20% in the maximum achievable frequency [10]. Thus, it is foreseen, that lifetime reliability will become a first-order challenge in the forthcoming and future technologies.

In parallel with the ever-growing reliability concerns, demands for higher energy efficiency and computational throughput have led to the proliferation of the asymmetric multicore architectures. Many commercial products, such as ARM's big.LITTLE chip and DynamIQ technology, integrate asymmetric processing units. Even Intel's multicore processors using TurboBoost technology could be also considered as dynamically asymmetric architectures. Such asymmetric designs, however, can affect in a non-uniform manner reliability-related parameters, such as the temperature, voltage and frequency, and unavoidably manifest into variable lifetimes [7][8][25][33], unlike the symmetric designs. As a result, some components might age faster, leading to earlier performance degradation with eventual device breakdown. Therefore, understanding the wear-out effects on asymmetric multicores is an important research question, as lifetime reliability concerns continue to grow. To guarantee correct operation throughout the lifetime of a processor, computer architects mainly provide worst-case timing guard-bands and runtime wear-out mitigation techniques. Among them, the most common state-of-the-art methods are wear-out monitoring through sensors [1][4][23], aging-aware scheduling [25][30][35], dynamic voltage and

frequency scaling (DVFS) [9][24], and spare structures [19][34] to replace malfunctioning components and mitigate aging [18]. In addition, previous reliability studies [5][13][26] introduced wear-out estimation methodologies and assessed the vulnerability of symmetric multicore hardware designs on device degradation phenomena. Another recent approach [32] has targeted lifetime reliability of heterogeneous processors and proposed a reliability model based on Amdahl's law. On the contrary, this paper studies the effects of wear-out on asymmetric multicore hardware designs ranging from mobile systems to low-power server deployments. Furthermore, as process variability and wear-out cause variations in threshold voltages, that decrease the range of voltage scaling [10][22][38], the effectiveness of DVFS may be limited in the future. Therefore, we perform a stress-case reliability assessment considering the processors under continuous operation and without DVFS or operating into low-power states.

In detail, the key design guidelines of this reliability assessment are the following:

- We simulate the effects of design technology, processor configuration, physical parameters and thermal setup, on wear-out rate, and measure the extend of lifetime variability. Our experimental results highlight that future lifetime reliability mechanisms have to be aware of the processors' asymmetry, since the average measured lifetime variability between the least and the most aged processor's components can be up to 2.6 years.
- We assess the impact of an MTTF-aware, asymmetric design on wear-out variability. As our simulation results show, the processor's lifetime is prolonged by up to 21.3% (i.e., additional 2.03 years) compared to the MTTF-oblivious design.

II. MODELING DEVICE DEGRADATION

The most widely used method to model the failure rates of device degradation mechanisms has been the Exponential distribution. The assumption of constant failure rate throughout a processor's lifetime simplifies the problem of lifetime estimation as it allows system-level reliability to be calculated by applying the Sum of Failures Rates (SOFR) model. However, this assumption is not realistic, since the wear-out phase follows complex non-exponential lifetime distributions while the failure rate increases over time [4][12][31][34]. To address this issue, more general lifetime distributions, such as the Weibull distribution (or lognormal), can be utilized [15][33]. The Weibull distribution consists of the shape (β) and the scale parameter (a). The scale parameter determines when, in time, a given portion of the population will fail, while the shape parameter enables the modelling of any phase in a chip's lifetime. For instance, the normal life

period where a constant failure rate exists is modelled when β equals to one. On the contrary, values greater than one represent an increasing failure rate, as in the wear-out phase.

This paper focuses on NBTI device degradation, since it is the most critical failure mechanism due to technology scaling [11][17][27]. This failure mechanism reflects long-term device behavior. However, it is infeasible to analyze it with RTL-level details due to the low simulation throughput. On the contrary, a microarchitecture-level abstraction is employed to study its effects. Therefore, by merging equations (1) and (2) (as defined in [33]) we estimate the $MTTF_{NBTI}$ using the Weibull distribution. Clearly, the operating ranges along with the temperature manifest in the scale parameter. Likewise, any other degradation mechanism (e.g., $MTTF_{HCI}$) can be applied to our modelling methodology.

$$MTTF_{weibull}(T) = a(T) \times \Gamma\left(1 + \frac{1}{\beta}\right), \quad (1)$$

$$a(T) = \frac{a \times V_{gs}^n \times e^{-E_a/kT}}{\Gamma\left(1 + \frac{1}{\beta}\right)}, \quad (2)$$

III. EXPERIMENTAL FINDINGS

A. Experimental Setup

We focus on emerging, low-power, ARM-based processors, developed with the big.LITTLE technology and we consider 8-core configurations. The big.LITTLE architecture, commercialized by ARM, although was originally conceived as an energy-efficient design for mobile phones, is also an interesting technology for low-power servers [36].

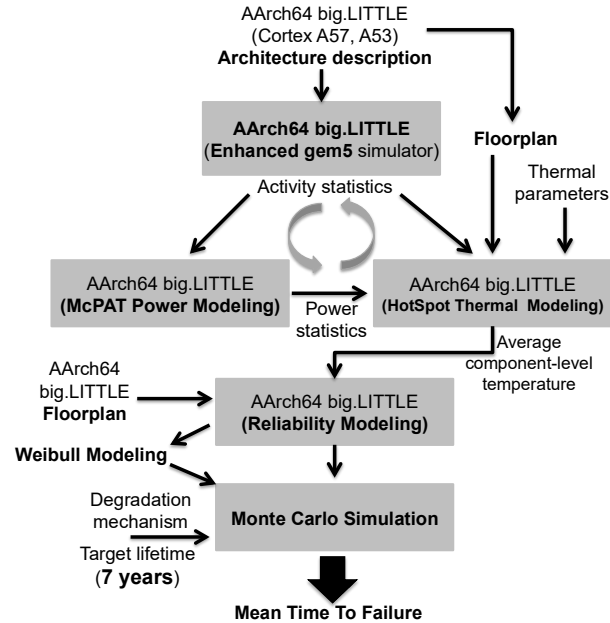


Fig. 1. The developed tool-chain to estimate Mean Time to Failure.

Figure 1 presents a schematic of the developed tool-chain. At first, the architecture descriptions of the Cortex A57 (big) and Cortex A53 (little) cores comprising an 8-core asymmetric processor were implemented on the gem5 full-system simulator [3]. Furthermore, gem5 was enhanced with an interface that handles, on the fly, the communication to the

McPAT power simulator [20] and the HotSpot temperature modelling tool [14]. To achieve that, a pipe is set to periodically transmit the activity statistics necessary to calculate the power and thermal profiles of the executed workloads. Additionally, McPAT feeds HotSpot with the power statistics to produce the temperature profile of the processor's components. Power and temperature trace files are generated at the granularity of 0.01 seconds. Consequently, the average temperatures of the Instruction Fetch Unit (IFU), Load/Store Unit (LSU), Memory Management Unit (MMU), Register File (RF), Rename Logic (RL) and the Execution Unit (EXEC), throughout the entire execution, are provided to the Monte Carlo simulation infrastructure. Additionally, the selected reliability model (NBTI) and an in-house developed floor-plan of the AArch64 big.LITTLE multicore processor is also needed to calculate the MTTF. Finally, throughout our calculations we set the **target lifetime equals to 7 years**, which is the most commonly used lifetime expectancy [29]. Note that having higher targeted lifetime, will proportionally affect the estimated lifetime. We employ Monte Carlo simulations to calculate lifetime, since a processor's MTTF based on the Weibull distribution is hard to be calculated analytically due to the variable failure rates of each hardware component [29]. The Monte-Carlo algorithm output is the minimum MTTF among the device degradation mechanisms (in our experiments only NBTI is considered). In the Monte-Carlo simulations, the accuracy of the analysis increases with the number of iterations performed. Therefore, the threshold was set to 10^7 since increasing the amount of iterations beyond that resulted in less than 0.003% variation on the estimated MTTF between subsequent runs of the algorithm. Furthermore, to verify that our Monte-Carlo algorithm converges to the analytical model, we run our tool-chain having a constant failure rate ($\beta = 1$); and compared it with the outcome of the analytical model when applying the same inputs. As expected, the estimated MTTFs were identical.

TABLE 1: DESIGN, PHYSICAL AND THERMAL PARAMETERS.

Design Parameters						
Cores	4 Out-of-Order			4 In-Order		
DRAM	2GB					
L3	16MB, 16 ways					
L2	2MB, 16 ways			512KB, 8 ways		
L1-I	48KB, 3 ways			32KB, 2 ways		
L1-D	32KB, 2 ways			32KB, 2 ways		
Issue Width	8			1		
Physical Parameters [6][33]						
Area (Cluster)	13.58mm ²			4.8mm ²		
Power (per Core)	2.1Watts (@ 2.5GHz)			0.5Watts (@ 1.5GHz)		
Operating Point	Perf.	Bal.	Pow.	Perf.	Bal.	Pow.
Voltage (V)	1.2	1.0	0.92	0.92	0.84	0.80
Frequency (GHz)	2.5	2.0	1.5	1.5	1.0	0.73
Thermal Parameters (in Kelvin)						
Setup	Soft			Heavy		
Ambient Temp.	298			318		
Initial Temp.	303			333		
DTM Threshold	354					

The physical and thermal parameters were configured to model a wide range of operating modes, along with various phases during the lifetime of an asymmetric multicore processor. In particular, our experiments were performed on a diverse set of operating points (Table 1, physical parameters):

(1) *high performance*: sets the cores to the highest frequency-voltage point, (2) *Balanced*: this configuration functions on a performance-power balanced mode, and (3) *Power save*: adjusts the core to the lowest frequency-voltage point. In addition, we used the following thermal scenarios: (1) a *soft* thermal stress setup modelling a system with a nominal utilization of its resources, and (2) a *heavy* thermal stress setup with a high utilization profile during a long period of time. Finally, the same operating point was statically assigned to each core at the beginning of each simulation run. Note that any thermal throttling avoidance mechanism were deactivated from our design (e.g., DVFS), since we attempt to discover the reliability limits of asymmetric multicore designs by performing a stress-case reliability assessment.

Traditionally, reliability studies and new architectures have been evaluated using well-studied and broadly accepted sequential benchmarks, such as SPEC2006, focusing on the single-thread characteristics. Nevertheless, as this paper focuses on multicore asymmetric architectures, parallel workloads from diverse domains implemented in various programming models are used. As a result, in our experimental setup, we utilize PARSEC 3.0 [2] and OpenStream [28] workloads in order to get a deeper understanding on the reliability requirements of current and future asymmetric processors (Table 2). For each benchmark, eight software threads were assigned to the available AArch64, big.LITTLE cores (i.e., 4 software threads on cluster-level and 8 threads on processor-level). The majority of the benchmarks are simulated for over a billion instructions, which is an order of magnitude higher than what is typically used in the literature. Finally, we use a warm-up period of 10M cycles after booting the Linux kernel (the performance counters were nullified at the end of the warm-up period).

TABLE 2: THE LIST OF BENCHMARKS RUN ON THE TOOLCHAIN.

OpenStream
Jacobi2d, Bitonic, Cholesky, Kmeans, Bzip2, SparseLu, FFT, Seidel2d
PARSEC (simsmall dataset)
Blackscholes Facesim, Fluidanimate Freqmine, Swaptions

B. Hardware Impact on Wear-out Rate

The accurate estimation of the lifetime of a processor assists computer designers to carefully plan for reliability enhancements with low cost and energy efficiency. In this section, we perform a quantitative analysis of how the lifetime reliability is affected by: (a) the design technology, (b) the processor configuration, (c) the physical parameters, and (d) the thermal setup.

Design Technology

Fig. 2 shows the impact of the design technology on MTTF analysis. In particular, we present the MTTF across all benchmarks and operating points in the soft thermal setup on the cluster-level granularity. On that mode of operation, a single cluster of cores, either the big or little, was activated each time. As shown, the MTTF for the big cluster ranges from 7.5 to 9.6 years (high performance to power save operating point), while for the little cluster the MTTF is 8.7 years on the high performance, 9.6 years on the balanced, and 10.1 years on the power save operating point. Obviously, the core type (i.e., Out-of-Order vs. In-Order cores) affects the

lifetime as well as increases the MTTF variability. For instance, the maximum difference between the Out-of-Order and the In-Order cluster equals to 1.19 years. To justify the aforementioned findings, we measured the temperature of each cluster (Fig. 2 - Bottom). As expected, the higher the temperature is, the shorter the lifetime will be since the temperature is adversely related to MTTF for the NBTI. Moreover, the MTTF variation is also reflected on the temperature values of each cluster.

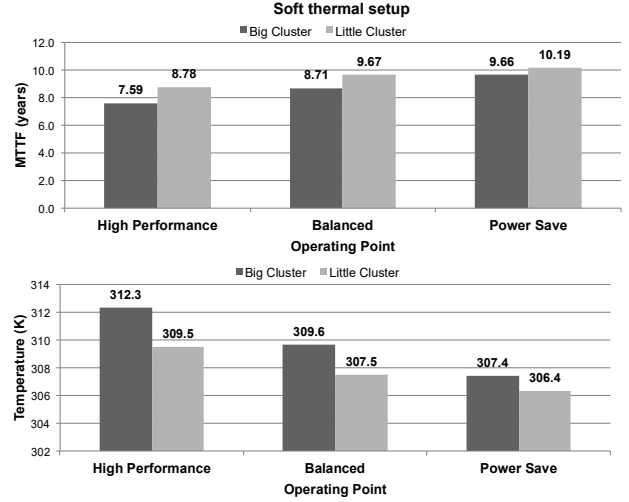


Fig. 2. The MTTF (Top) and the measured temperature (Bottom) of big and little cores. Note that small deviation, between processor- and cluster-level, MTTF are due to OS activity.

We also noticed that the cluster with the highest throughput has the shortest MTTF since the more operations it executes, the hotter it becomes. The throughput of the big cluster ranges from 4.06 to 3.35 operations per cycle (from the high performance to the power save operating point), while for the little cluster the throughput equals to 2.88 on the high performance point, 2.6 on the balanced, and 2.56 on the power save operating point. As a result, the big cluster is stressed more than the little one; and when its frequency is reduced, its capacity is also decreased.

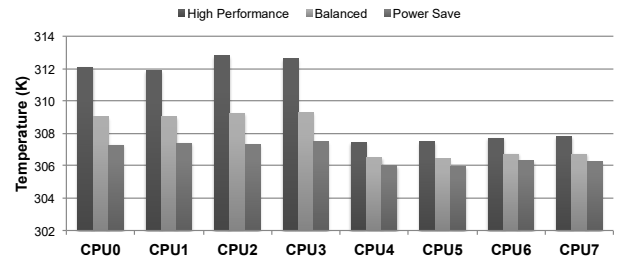


Fig. 3. The temperature of the big.LITTLE cores. Big cluster is comprised of cores 0 to 3, while little cluster of cores 4 to 7.

Apart from studying the correlation of lifetime with the performance characteristics, we performed a top-down analysis on the temperature profiles of the components comprising the big and the little clusters. Fig. 3 depicts the temperature of each core of the processor. As expected, the Out-of-Order cores exhibit higher temperature than the In-Order cores, due to their higher complexity and throughput.

Moving deeper into the micro-architecture of each core type, the temperature trend remains unaltered. Fig. 4 presents the temperature profiles of the Instruction Fetch Unit (IFU),

the Rename Logic (only on the Big core), the Load/Store Unit (LSU), the Memory Management Unit (MMU – including TLBs), the Register File, and the Execution Unit. A first finding is that the components in the little core have constantly lower temperature than their counterparts on the big core. Moreover, the temperature imbalance is increased on higher operating points due to the exploitation of the ILP capabilities of the big core. Another observation is that the heat is uniformly distributed among the micro-architectural components. Overall, the design diversity of the micro-architectural components is reflected on their temperatures. As a result, the temperature variation increases the lifetime variability at the component-, core-, cluster- and, ultimately, at the processor-level.

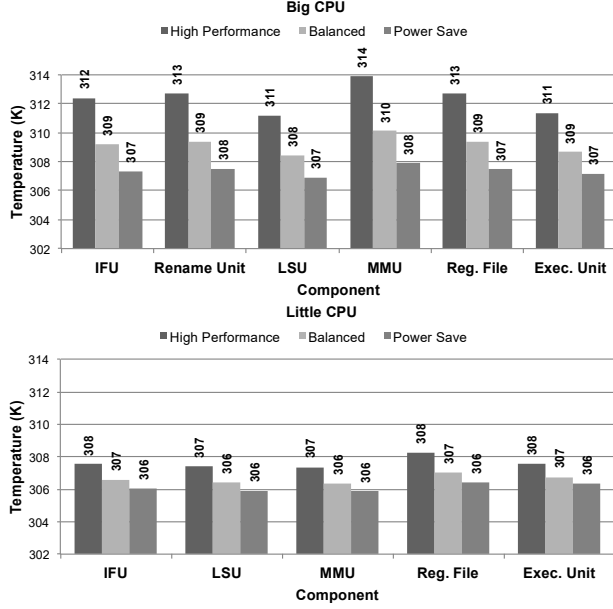


Fig. 4. The temperature profile of the micro-architectural components of a big and little core for the soft thermal setup.

Physical Parameters

Based on the insights gained from Figures 5 to 7, the lower the voltage/frequency point is, the longer the lifetime of the processor will be. This is justified since on lower operating points (e.g., on the power save point, the cores constantly operate on 50% of their maximum performance), the processor is less stressed, dissipates heat more efficiently and therefore prevents the formation of thermal hotspots, which are the main cause of a chip’s deterioration. Furthermore, the variant operating ranges, between the big and little cluster, affect the aging rates and, in turn, increase the MTTF variability. Thus, the MTTF difference can be up to 2.6 years, when clusters function on different operating points (High performance vs Power save).

Thermal Parameters

In this subsection, we examine the impact of the thermal configuration on a processor’s lifetime. To that end, we setup a stress case scenario, such as the heavy thermal setup, and rerun the experiments. Fig. 5 shows that the cluster-level MTTF trend is significantly different to that of the soft thermal setup (Fig. 2). In particular, the big cluster’s MTTF ranges from 4.4 to 6.1 years, while the little cluster’s MTTF equals to 4.2 years on the high performance, 5.2 years on the balanced,

and 5.7 years on the power save operating point. Strangely, the little cores age faster than the big cores. Even though this behavior was unexpected, it is fully justified by further analyzing the characteristics of the execution profiles. Carefully analyzing this finding, we noticed that the inability to dynamically lower the frequency, along with the high initial and ambient temperature on the little cluster, result in thermal throttling very quickly (little cluster has higher temperature). Therefore, the little cluster is incapable to efficiently dissipate the excessive heat. Furthermore, the MTTF difference between the big and little clusters is also depicted on the measured power density since the former has only about 1.5x higher throughput for 2.8x larger area. Greater power density means higher operating temperatures and thus lower reliability per area.

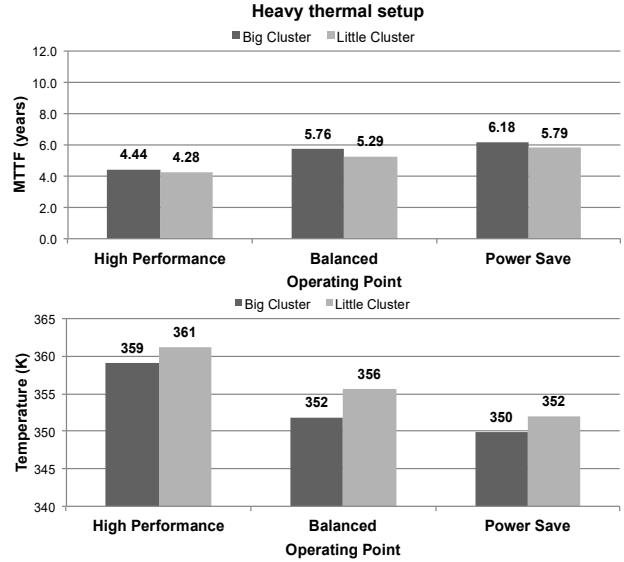


Fig. 5. The estimated MTTF (Top) and the measured temperature (Bottom) of big and little cluster for the heavy thermal setup.

Another finding from our experimental results is that the MTTF in the heavy thermal stress setup is always lower than the soft thermal setup (Fig. 2). This is attributed to the huge difference between the thermal setup temperatures: on the heavy thermal scenario the temperature ranges from 353.0K to 362.1K, while on the soft setup it ranges from 306.0K to 312.0K. Finally, the lifetime of the big and little clusters differs up to 0.5 years meaning that the big cluster will continue to be fully operational after the breakdown of the little cluster.

Design Configuration

To further analyze the impact of processors’ asymmetry on lifetime reliability, we perform a quantitative comparison with different processor configurations. In particular, we compare a symmetric processor consisting of eight Out-of-Order cores. Fig. 6 shows the MTTF for all operating points and thermal setups for the symmetric and asymmetric processors across all benchmarks. As shown, the asymmetric processor has up to 2 years shorter MTTF for the same operating point and setup. This translates to 40% difference against the symmetric processor. The AArch64, big.LITTLE processor, highly increases the cluster-level lifetime variability due to its asymmetry. Finally, on the symmetric processor the maximum MTTF difference between the least

and the most worn-out core is 0.7 years, while on the asymmetric design is 2.6 years. This is justified since the system utilization on the symmetric processor is uniformly distributed across cores (i.e., throughput STDEV across all cores and operating points equals to 0.02), and the temperature variations are minimal.

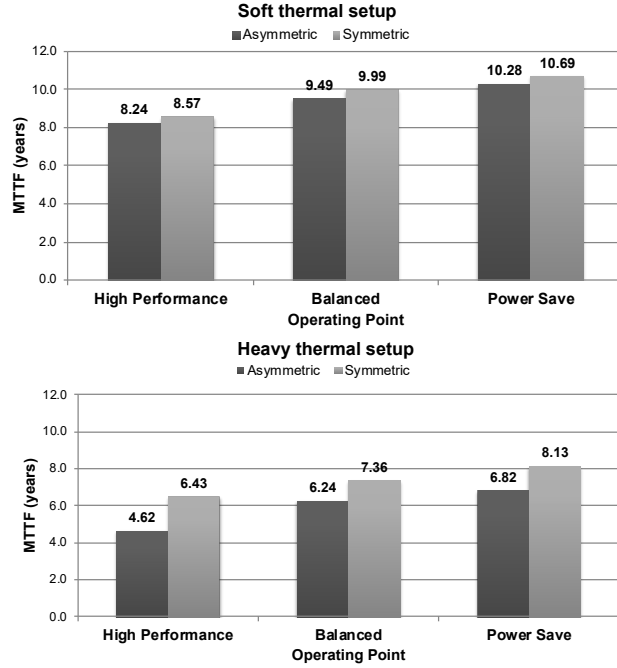


Fig. 6. The average MTTF of the homogeneous versus the big.LITTLE processor on the soft and heavy thermal stress setup.

Technical Guidelines Summary

Our experimental findings lead to the following guidelines:

1. Wear-out preventing techniques must limit the increased lifetime variation to avoid reliability bottlenecks.
2. Workload schedulers should be thermal-aware, and, thus, MTTF-aware.
3. Operating conditions have the highest impact on MTTF variability, followed by design technology and software structure. Computer designers should be aware of the aforementioned priority, when designing a system.

IV. MTTF-AWARE DESIGN CONFIGURATION

As shown in Fig. 2, the MTTF of the processor is bound by the big cluster, although the little cluster has 13.5% longer MTTF. It is evident though that the higher the wear-out rate variability, the shorter the lifetime of an asymmetric system since the weakest component characterizes the lifetime of the whole processor. Therefore, having a more balanced operational stress among the clusters, will diminish the cluster-level, wear-out rate variability and, in turn, will positively affect the overall lifetime. To that end, we assess the impact of an MTTF-aware, asymmetric design configuration on wear-out variability. In particular, the employed design is inspired by [37]. Note that a set objective while modelling this configuration was to avoid increasing the die area compared to the MTTF-oblivious design.

The MTTF-aware design configuration consists of three clusters of cores each containing the following resources: (1) *big*: two AArch64 Out-of-Order cores, (2) *Little*: four AArch64 In-order cores; and (3) *Little-faster*: two AArch64

with identical design parameters with the little cluster, but with different physical parameters (Table 3).

TABLE 3: DESIGN CONFIGURATION OF THE LITTLE-FASTER CLUSTER.

Little-faster Physical Parameters			
Area (Cluster)	4.8mm ²		
Power (per Core)	0.5Watts (@ 1.5GHz)		
Operating Point	Perf.	Bal.	Pow.
Voltage (V)	1.0	0.92	0.84
Frequency (GHz)	2.0	1.5	1.0

Fig. 7 presents the MTTF of the MTTF-oblivious (i.e., the regular big.LITTLE processor) and the MTTF-aware processor design configuration. As shown, the MTTF-aware asymmetric configuration decelerates the wear-out rate, due to cluster-level variability reduction, and thus prolongs the lifetime of the system by up to 21.3% (2.03 years additional lifetime).

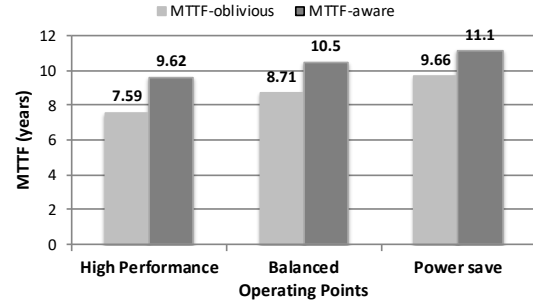


Fig. 7. MTTF-oblivious vs. MTTF-aware estimated lifetime.

Although the MTTF-aware architecture manages to prolong the lifetime reliability, another interesting study is how it affects other key design parameters, that indirectly correlate to the lifetime reliability, such as the power consumption. Fig. 8 shows that the average power of the whole system is reduced by 17% due to the MTTF-aware design. At first, the removal of the two big cores from the design results in less static power consumption. Furthermore, the little-faster cores are more power-efficient (than the big cores) due to their simplest design complexity and lower operating points. Finally, the addition of the little-faster cluster generates a more balanced resource utilization profile (with lower power density), which in turn positively affects the overall power consumption of the processor (and its lifetime). Regarding performance, the proposed architecture, naturally, falls behind the baseline design. This is justified as the peak performance of the little-faster cluster is lower than that of the replaced one. However, in the Pareto-optimal space of reliability versus power, the MTTF-aware design is more efficient.

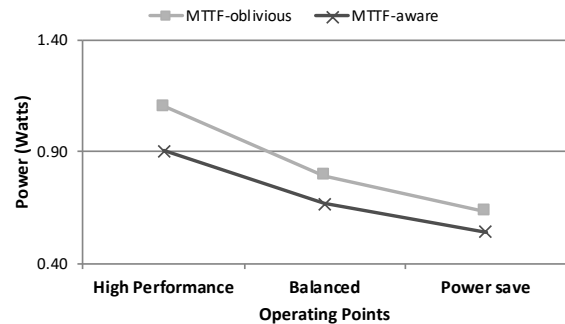


Fig. 8. Average power consumption of the MTTF-oblivious vs. MTTF-aware configuration.

V. CONCLUSIONS

Device miniaturization along with the inefficient power dissipation, and the unsustainable thermal thresholds accelerate the wear-out rates. Furthermore, hardware asymmetry has entered mainstream computing. This trend not only creates new opportunities for driving forwards the performance and energy-efficiency boundaries, but also poses new challenges in reliability. In this paper, we simulate the effects of wear-out on asymmetric processors and measure the MTTF variability within their components. As the experimental results highlight, the maximum difference between the least and the most aged components is 2.6 years. Motivated by this finding, we show that an MTTF-aware asymmetric configuration using three clusters prolongs its lifetime by 21%.

ACKNOWLEDGMENT

This work is supported by the EU Horizon 2020 EuroEXA 754337, E2Data 780245 and ACTiCLOUD 732366 grants. Mikel Luján is supported by an Arm/RAEng Research Chair.

REFERENCES

- [1] J. Abella, X. Vera and A. Gonzalez, "Penelope: The NBTI-Aware Processor", In *40th Annual IEEE/ACM Intl. Symposium on Microarchitecture*, Chicago, IL, 2007, pp. 85-96.
- [2] C. Bienia, "Benchmarking Modern Multiprocessors", Princeton University, 2011.
- [3] N. Binkert et al., "The gem5 simulator", *SIGARCH Comput. Archit. News* 39, 2, (August 2011), 1-7.
- [4] J. Blome, S. Feng, S. Gupta, and S. Mahlke, "Self-calibrating Online Wearout Detection", In *40th Annual IEEE/ACM Intl Symposium on Microarchitecture*, Chicago, IL, 2007, pp. 109-122.
- [5] C. Bolchini, L. Cassano, A. Miele, "Lifetime-aware load distribution policies in multi-core systems: An in-depth analysis", In *Design, Automation & Test in Europe Conference*, 2016, pp. 804-809.
- [6] Y. Cakmakci et al., "Cyclic Power-Gating as an Alternative to Voltage and Frequency Scaling", In *IEEE Computer Architecture Letters*, vol. 15, no. 2, July-December 2016.
- [7] A. Deb, B. Vermeulen, L. Van Dijk, "Overview of Health Monitoring Techniques for Reliability", In *Workshop on Early Reliability Modeling for Aging and Variability in Silicon Systems*, 2016.
- [8] S. Feng, S. Gupta, A. Ansari and S. Mahlke, "Maestro: Orchestrating Lifetime Reliability in Chip Multiprocessors", In *Proceedings of the 5th Intl. conference on High Performance Embedded Architectures and Compilers*, 2010, 186-200.
- [9] S. Feng, S. Gupta, S. Mahlke, "Olay: Combat the signs of aging with introspective reliability management", In *Workshop on Quality-Aware Design (W-QUAD)*, 2008.
- [10] D. Gnad et al., "Hayat: harnessing dark silicon and variability for aging deceleration and balancing", In *ACM/EDAC/IEEE Design Automation Conference*, San Francisco, CA, 2015, pp. 1-6.
- [11] H. Hong, J. Lim, H. Lim, and S. Kang, "Lifetime Reliability Enhancement of Microprocessors: Mitigating the Impact of Negative Bias Temperature Instability ACM Comput. Surv. 48, 1, Article 9 (September 2015), 25 pages.
- [12] L. Huang and Q. Xu, "AgeSim: A simulation framework for evaluating the lifetime reliability of processor-based SoCs", In *Design, Automation & Test in Europe Conference*, Dresden, 2010, pp. 51-56.
- [13] L. Huang and Q. Xu, "On modeling the lifetime reliability of homogeneous manycore systems", In *IEEE Pacific Rim Intl. Symposium on Dependable Computing*, Taipei, 2008, pp. 87-94.
- [14] W. Huang et al., "Accurate, Pre-RTL Temperature-Aware Processor Design Using a Parameterized, Geometric Thermal Model", *IEEE Transactions on Computers*, 57(9):1277-88, September 2008.
- [15] L. Huang and Q. Xu, "Characterizing the lifetime reliability of manycore processors with core-level redundancy", In *IEEE/ACM Intl. Conference on Computer-Aided Design*, San Jose, CA, 2010, pp. 680-685.
- [16] ITRS International Technology Roadmap for Semiconductors. *Process integration, devices, and structures*, 2013.
- [17] S. Khan et al., "Bias Temperature Instability analysis of FinFET based SRAM cells", In *Design, Automation & Test in Europe Conference & Exhibition*, Dresden, 2014, pp. 1-6.
- [18] H. Kim, A. Vitkovskiy, P. V. Gratz, and V. Soteriou, "Use it or lose it: wear-out and lifetime in future chip multiprocessors", In *Annual IEEE/ACM Intl. Symposium on Microarchitecture*, 2013, pp. 136-147.
- [19] T. Kim, X. Huang, H. B. Chen, V. Sukharev and S. X. D. Tan, "Learning-based dynamic reliability management for dark silicon processor considering EM effects", In *Design, Automation & Test in Europe Conference & Exhibition*, Dresden, 2016, pp. 463-468.
- [20] S. Li et al., "McPAT: An integrated power, area, and timing modeling framework for multicore and manycore architectures", In *IEEE/ACM Intl. Symposium on Microarchitecture*, 2009, pp. 469-480.
- [21] J. I. McCool, "Using the Weibull Distribution: Reliability, Modeling and Inference", Wiley Publications, 2012.
- [22] E. Mintamo et al., "Self-tuning for maximized lifetime energy efficiency in the presence of circuit aging", In *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 30(5):760-773, 2011.
- [23] P. Mercati, A. Bartolini, F. Paterna, T. S. Rosing, and L. Benini, "Workload and user experience-aware dynamic reliability management in multicore processors", In *CM/EDAC/IEEE Design Automation Conference*, Austin, TX, 2013, pp. 1-6.
- [24] M. G. Moghaddam, A. Yamamoto and C. Ababei, "Investigation of DVFS based dynamic reliability management for chip multiprocessors", In *Intl. Conference on High Performance Computing & Simulation*, Amsterdam, 2015, pp. 563-568.
- [25] T. R. Mück, Z. Ghaderi, N. D. Dutt and E. Bozorgzadeh, "Exploiting Heterogeneity for Aging-Aware Load Balancing in Mobile Platforms". In *IEEE Transactions on Multi-Scale Computing Systems*, vol. 3, no. 1, pp. 25-35, Jan.-March 1 2017.
- [26] F. Oboril and M. B. Tahoori, "ExtraTime: Modeling and analysis of wearout due to transistor aging at microarchitecture-level", In *Intl. Conference on Dependable Systems and Networks*, Boston, MA, 2012, pp. 1-12.
- [27] M. Ottavi et al., "Dependable Multicore Architectures at Nanoscale: The View From Europe", In *IEEE Design & Test*, vol. 32, no. 2, pp. 17-28, 2015.
- [28] A. Pop and A. Cohen, "OpenStream: Expressiveness and data-flow compilation of OpenMP streaming programs", In *ACM Trans. Archit. Code Optim.* 9, 4, Article 53, January 2013.
- [29] P. Ramachandran, S. V. Adve, P. Bose and J. A. Rivers, "Metrics for Architecture-Level Lifetime Reliability Analysis", In *IEEE Intl. Symposium on Performance Analysis of Systems and Software*, Austin, TX, 2008, pp. 202-212.
- [30] S. Rehman, F. Kriebel, M. Shafique and J. Henkel, "Compiler-driven dynamic reliability management for on-chip systems under variabilities", In *Design, Automation & Test in Europe Conference & Exhibition*, Dresden, 2014, pp. 1-4.
- [31] M. J. Schulte et al., "Achieving Exascale Capabilities through Heterogeneous Computing", In *IEEE Micro*, vol. 35, no. 4, pp. 26-36, July-Aug. 2015.
- [32] W. J. Song, S. Mukhopadhyay and S. Yalamanchili, "Amdahl's law for lifetime reliability scaling in heterogeneous multicore processors", In *Intl. Symposium on High Performance Computer Architecture*, Barcelona, 2016, pp. 594-605.
- [33] W. Song et al., "Architectural Reliability: Lifetime Reliability Characterization and Management of Many-Core Processors", In *IEEE Computer Architecture Letters*, vol. 14, no. 2, pp. 103-106, July-Dec., 2015.
- [34] J. Srinivasan, S. V. Adve, P. Bose, and J. A. Rivers, "Exploiting Structural Duplication for Lifetime Reliability Enhancement", In *Intl. Symposium on Computer Architecture*, Madison, WI, USA, 2005, pp. 520-531.
- [35] A. Tiwari and J. Torrellas, "Facelift: Hiding and slowing down aging in multicores", In *IEEE/ACM Intl. Symposium on Microarchitecture*, Lake Como, 2008, pp. 129-140.
- [36] <http://www.kaleao.com/Products/kmax/kmax-specifications>
- [37] <https://www.mediatek.com/products/smartphones/mt6797-helio-x20>
- [38] X. Wang, A. Brown, B. Cheng, and A. Asenov, "Statistical variability and reliability in nanoscale finfets", in *Proc. IEEE Int. Electron. Dev. Meeting*, 2011, pp. 5.4.1-5.4.