

# Implementing Low-Power Configurable Processors — Practical Options and Tradeoffs

John Wei and Chris Rowen

Tensilica, Inc.  
3255-6 Scott Blvd.  
Santa Clara, CA 95054

jwei@tensilica.com  
rowen@tensilica.com

## Abstract

Configurable processors enable dramatic gains in energy efficiency, relative to traditional fixed instruction-set processors. This energy advantage comes from three improvements. First, configuration of the instruction set permits a much closer fit of the processor to the target applications, reducing the number of execution cycles required. Second, configuring the processor removes unneeded features, reducing power and area overhead. Third, automatic processor generation tools enable logic optimization, signal switching reductions, and seamless mapping into low-voltage circuits and processes, for very low-power operation. The first improvement has been well-studied. Analysis of the second and third improvements requires detailed circuit and layout experiments, which is the primary focus of this paper.

Starting from a range of existing available power saving options, this work explores the tradeoff and analyzes the results: the design priority tradeoff, the process technology impact, and implementing low-power configurable processor using commercial scaled-VDD cell libraries compatible with mainstream SOC practices. These real processor designs can achieve power dissipation approaching 20 $\mu$ W/MHz at 0.8V and close to 10 $\mu$ W/MHz at 0.6V, using production 0.13 $\mu$ m libraries. Finally, this work quantifies the dramatic process, voltage and temperature dependence in post-layout leakage power for small processor designs.

## Categories and Subject Descriptors

M2.5 [Design Methods]: Low-power design

**General Terms:** Algorithms, Performance, Design, Verification.

**Keywords:** Configurable embedded processor, SOC (system on chip), PVT (process, voltage, temperature), Low-power, Leakage Power, Dynamic Power, Dynamic power efficiency, Scaled VDD.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

DAC 2005, June 13–17, 2005, Anaheim, California, USA.

Copyright 2005 ACM 1-59593-058-2/05/0006...\$5.00.

## 1. Introduction

Processors are fundamental building blocks for designs of SOCs. General-purpose, fixed ISA processors dissipate significant switching and leakage power. Furthermore, functionality and feature additions in a general purpose processor to achieve higher performance leads to an increasing trend in the dissipated power, independent of process technology. This power dissipation trend can be mitigated by use of configurable processors, which are customized for a target application.

Given familiar abstraction provided by processors; the increase in the transistor count on a chip; and the shortening time-to-volume-market windows for power-aware consumer products, the increasingly relevant question to ask is what practical options exist to reduce the power and the energy dissipated by a processor-based solution. Based on the assumption that a configurable processor is able to retain application performance while enabling lower MHz performance, the paper quantifies practical power saving options for state-of-the-art ASIC design flows. Figure 1 summarizes the results of a detailed study of the energy efficiency optimization of four algorithms: dot-product of two 2048 element vectors, Advanced Encryption Standard (AES) security coding, Viterbi decoding for wireless communication, and 256-point complex Fast Fourier Transform (FFT). The energy savings relative to a power-efficient reference configuration of the same Xtensa architecture ranges from 2x to 82x. [1]

Configuration		Dot Product	AES	Viterbi	FFT
Reference Processor	Cycles (K)	12	283	280	326
	Energy ( $\mu$ J)	3.3	61.1	65.7	56.6
Optimized Processor	Cycles (K)	5.9	2.8	7.6	13.8
	Energy ( $\mu$ J)	1.6	0.7	2.0	2.5
<i>Energy Improvement</i>		<i>2x</i>	<i>82x</i>	<i>33x</i>	<i>22x</i>

Figure 1. Efficiency gains for application-tuned processors

The number of embedded processor cores in multi-core SOC designs is projected by the ITRS (International Technology Roadmap for Semiconductors) to double with each successive technology node [2]. The use of multiple embedded processors in SOC designs is increasingly gaining popularity. Future SOC design improvement will be based on large numbers of processors as the basic build block, according to the Processor Scaling Model [1].

The reallocation of silicon real estate to favor multiple processors in SOC design warrants the need to implement low-power

processors, particularly for mobile, hand-held, battery-operated applications. Indeed at gate level, library vendors have pushed to deliver low-power design platforms such as the Artisan SAGE-X™ and Metro™ standard cell libraries and Virtual Silicon Mobilize™ standard cell libraries to enable low-power implementations [3] [4].

On the other hand, the quest for higher performance for general-purpose embedded processors has motivated the addition of more gates leading to ever-larger processor area and increased dynamic power (rising uW/MHz). This increase stems from increased pipeline depth to permit high clock frequency, addition of advanced architectural features for reduction of the resulting branch penalties, and inclusion of new instructions to support new application domains. To counter this trend, optimization of the processor’s instruction set and implementation for specific application domains is emerging as an important tool for energy minimization.

## 2. Power Saving Options

A matrix of existing, available power saving approaches is shown in Figure 2. Dynamic or switching power is expressed by

$$P_{dynamic} = k * C * V^2 * F * SA$$

leakage power

$$P_{leakage} = I_{leakage} * V * A$$

where, **k** = Constant (usually varies from 0 to 1), **C** represents capacitance, **V** is the operational voltage and **F** is the frequency for the design, **SA** is the switching activity, **I<sub>leakage</sub>** is the unity gate leakage current, and **A** is the total effective transistor width (usually proportional to gate area and gate count). The option of architecture configurability enables a system architect to use flexible application-specific processor configurations instead of using a general-purpose, fixed-feature processor.

Design Level	Dynamic Power	Leakage Power
Physical Synthesis	Push min area for min speed spec	Push min area for min speed spec
Choice of Cell Library and Process Technology	Scaled VDD Choice of process Back bias	Gate bias Scaled VDD Multi-Vt Choice of process
Configuration of Logic, Micro-architecture and Architecture	Clock gating Memory configuration/size, pipeline depth Instruction set	Functional clock gating+ gate bias Memory size, pipeline depth Instruction set
System Design	Processor configuration Sleep process Energy management, Dynamic frequency and voltage scaling	Processor configuration Sleep process Gate bias, substrate bias

Figure 2. Power reduction methods

While significantly saving the silicon real estate, the configurable processor attacks most of the variables in the power equations to achieve low-power designs, e.g., smaller microprocessor configuration without unused features which reduces **C** and **A**, application-tuned architectures allow more tasks get done within each clock cycle or at minimum **F**, and extensive clock gating makes possible the lowest **SA** [5]. This work focuses on the power saving options of design priority (**C** and **A** vs. **F**), process

technology (**C**), scaled VDD (**V**), and DVFS (**V** and **F**), by exemplifying the tradeoffs wherever applicable.

The paper deals with techniques to further reduce power for any given architecture. The same small configurable processor design, a small configuration of a Tensilica Xtensa LX processor, is used throughout this work to assure consistency in comparisons. This complete implementation of the 32-bit Xtensa architecture requires less than 0.2mm<sup>2</sup> for the logic. The best-case corner was used for hold-time fixing in Silicon Ensemble (SE). All clock timing reported is based on the target clock as limited by the processor’s worst negative slack for internal flop-to-flop paths under the worst-case process, voltage and temperature (PVT) conditions. The P&R (place and route) used 5 layers in a 6-layer system. Using 6 different test suites, post-route gate simulation was run to create SAIF files for measuring power in post-layout back-annotated netlist with Power Compiler [6].

In the scaled VDD approach to save power, special attention is warranted in the design’s final timing signoff, because of the inherited reduced noise margin coupled with IR drop, cross-talk-induced signal integrity, and on-chip variation effects.

## 3. Basic Tradeoffs among Speed, Area, and Power

Depending on applications, the embedded processor design may seek the highest clock speed at the expense of silicon area and power in desktop/server applications, but the design priority is reversed in battery-limited/form-factor-limited mobile designs.

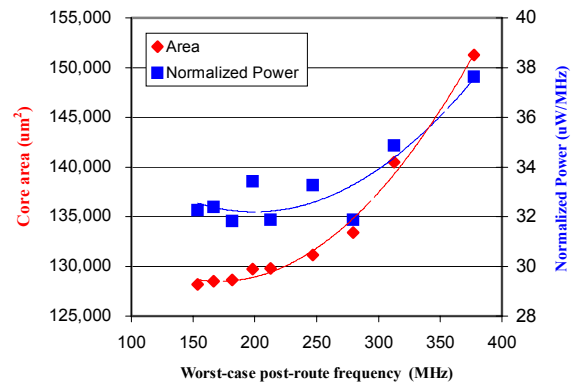
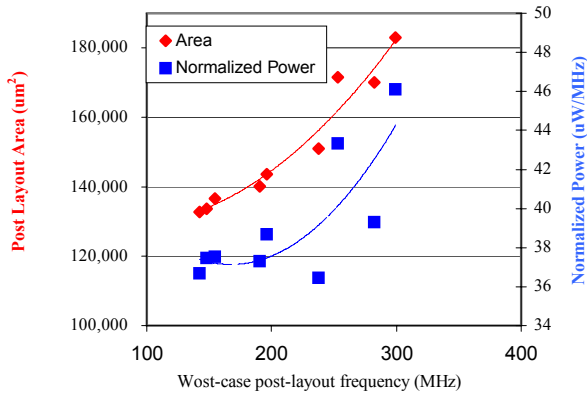


Figure 3. Speed to area and power tradeoff (0.13um LVLK-OD)

Nine different synthesized versions of the same embedded processor, targeting different clock frequencies, are implemented using the Artisan SAGE-X standard cell library in TSMC 0.13um LVLK-OD process. The Cadence Silicon Ensemble P&R utilization is about 97% for all the design cases. There is significant tradeoff between speed, area, and power as shown in Figure 3. The post-layout area differs 18% between the fastest processor of 377MHz and the slowest processor of 150MHz using 0.13um LVLK-OD. These nine processors are all fairly small, so the correlation between synthesis results and layout results for worst-case clock frequency and power are fairly close – within about 2% for frequency and 10% for power.

Results for the nine processors on the 0.13um GFSG process (Artisan SAGE-X libraries) are shown in Figure 4. The Silicon Ensemble P&R utilization ranges from 90% to 97% for all the

design cases. In Figure 4, the post-SE area differs 37% between the fast 299MHz processor and the slower 150MHz processor using 0.13um TSMC GFSG. The GFSG process still shows tight correlation between post-layout and post- synthesis area



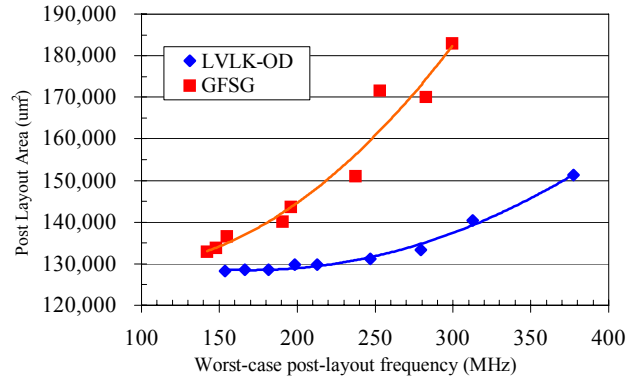
**Figure 4.** Speed to area and power tradeoff (TSMC 0.13um GFSG process)

(within 5%) but looser correlation for power (around 20%). The looser power correlation for the GFSG process appears to stem from differences in interconnect parasitic capacitance, as the difference in gate count between post-synthesis and post-layout circuits is only a couple of percent. The much tighter correlation in LVLK-OD process is due to the smaller interconnect parasitics in LK (inter-metal dielectric constant is ~2.7 in the LK process while 3.7 in FSG). For the LVLK-OD process, mW/MHz remains fairly constant for all cores, except for the two with increased area needed to achieve speed above 300MHz. As the target frequency is pushed higher for a specific process technology, the mW/MHz increases as higher drive strength cells are used. For low-power designs, it is important to determine if the target frequency for a given process is beyond the knee of the curve. In GFSG process, more of the cores require greater area to achieve their target frequency, increasing the power proportionally.

Because the correlations of power and area between post-synthesis and post-layout varies substantially across different cell library/process technologies even for the same processor design, only the results at the post-layout level representation are reported throughout the rest of this work.

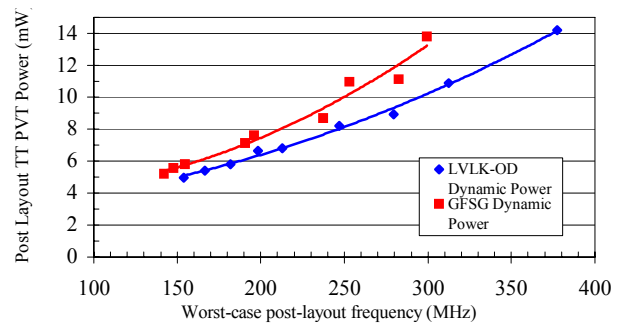
#### 4. Impact of Process Technology

Figure 5 summarizes the impact of process technology showing processor area vs. speed for 0.13um LVLK-OD and GFSG processes, based on Artisan SAGE-X libraries. For the same speed performance, LVLK-OD saves 10-20% processor area compared with GFSG: the higher the speed, the larger the area saving – the saving diminishes at lower speed.

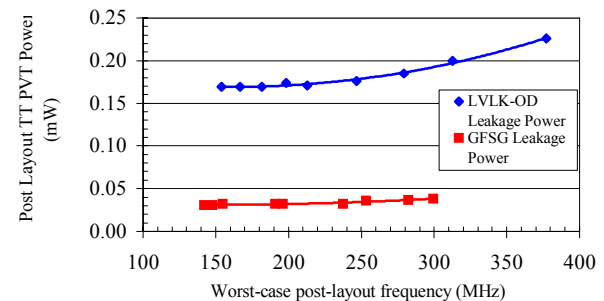


**Figure 5.** Impact of process technology on area

The impact of process on dynamic power (TT process corner) vs. speed is shown in Figure 6. For the same speed performance, the technology of 0.13um LVLK-OD saves 20%+ dynamic power compared with the 0.13um GFSG. Notice both the 0.13um LVLK-OD and 0.13um GFSG processes run at the same 1.2V core VDD.



**Figure 6.** The impact of process technology on dynamic power



**Figure 7.** Impact of process technology on leakage power

One clear disadvantage of 0.13um LVLK-OD vs. 0.13um GFSG is some 440% increase in leakage power consumption as shown in Figure 7. However, the total power saving using 0.13um LVLK-OD is still at least 10% from 150MHz through 300MHz, with the savings diminishing at lower speed.

#### 5. The Scaled VDD Approach

Because of the  $V^2$  effect in the  $P_{dynamic}$  discussed in Section 2, the scaled VDD approach is very effective to reduce dynamic power. Leakage power  $P_{leakage}$  is also reduced since it is proportional to  $V$ . IP vendors have redesigned and fine crafted the building block

circuitries so that they all work well under the reduced VDD using the TSMC 0.13um G (generic) process [3][4]. We look separately at the two leading low-voltage libraries, Virtual Silicon's Mobilize™ and Artisan's Metro libraries

### 5.1 Virtual Silicon Mobilize Results

The speed-to-area tradeoff has been characterized for the processor at the nominal VDD of 1.2V, 1.0V, and 0.8V as shown in Figure 8.

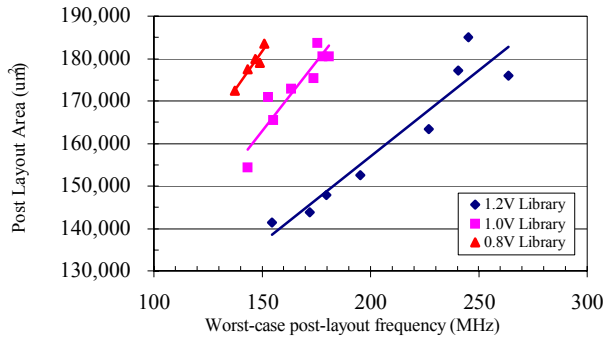


Figure 8. Speed to area tradeoff using Mobilize

There is an increasing tradeoff between speed and area at lower VDD, i.e., the cost in area is increasingly higher at lower VDD to progressively acquire each additional MHz in clock rate, as it takes more scaled VDD (hence weak cells) to gain additional speed. Despite the larger area for fast, low-voltage designs, power dissipation improves. For the same 150MHz post-layout speed, the scaled VDD shows massive power saving, due largely to the  $C \cdot V^2$  effect: compared to 1.2V, Virtual Silicon's 1.0V library achieved a 17% typical power reduction and the 0.8V library achieved a 53% power reduction, as shown in Figure 9. Designs were not targeted below 150MHz for the Mobilize experiments.

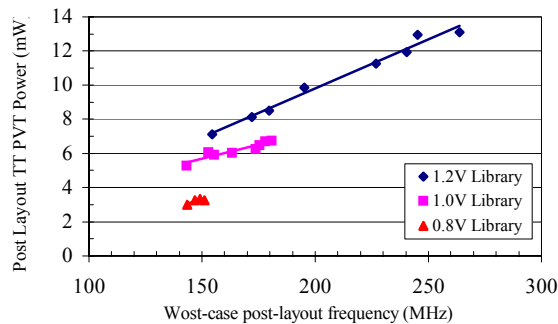


Figure 9. Speed to power tradeoff using Mobilize™

Power efficiency improves dramatically at lower VDD, with a typical Normalized power of 21.5uW/MHz achieved at 0.8V VDD, as shown in Figure 10. The total leakage power is shown in Figure 11 for the three embedded processors of 150MHz post-layout: the leakage does not decline as much in scaled VDD, because of the area increase to keep up the target speed. However the leakage power density (per unit cell area) scales with VDD, indicating the majority of the modeled leakage current is from sub-threshold diffusion current rather than drift current. Otherwise the scaling factor would have been  $VDD^2$  instead of VDD.

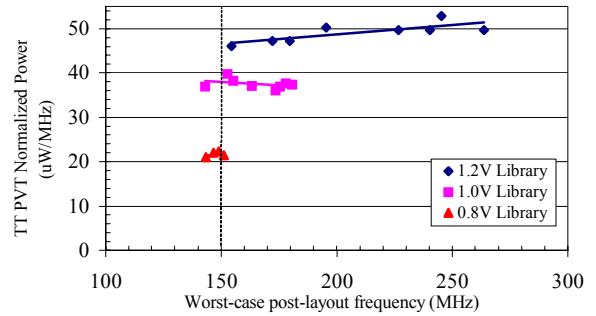


Figure 10. Speed to power efficiency tradeoff using Mobilize

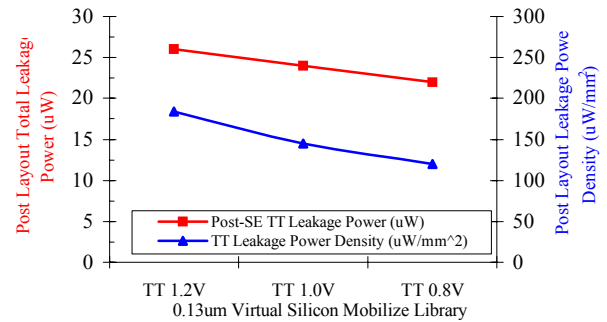


Figure 11. Leakage @ 150MHz WC using Mobilize

### 5.2 Artisan Metro Results

A somewhat different set of experiments were run using the popular Artisan Metro low-voltage libraries, with particular focus on low-voltage and low-frequency operation, including use of dynamic frequency and voltage scaling power management schemes. The Artisan libraries include fully-characterized leakage models, allowing detailed analysis of processor logic leakage. One processor configuration is optimized in WC 1.08V Metro library to create two logically-equivalent versions – high-speed and low-speed. Both the timing and WC power are then re-characterized for these two processors by switching to the other four lower VDD Metro libraries, without changing the overall layout, so all the high-speed versions are 148,000  $\mu m^2$  and all low-speed versions are 121,000  $\mu m^2$ . The voltage-to-speed tradeoff is shown in Figure 12.

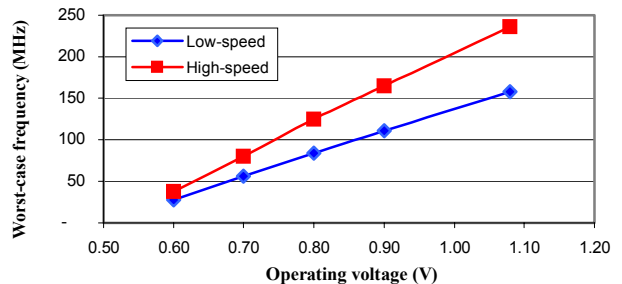


Figure 12. Voltage to speed tradeoff using Metro

Useful performance (30-40 MHz) is achieved even at 0.6V. The speed vs. power is shown in Figure 13: there is a tremendous power saving when the same processor design downshifts to lower VDD while also operating at lower clock rate.

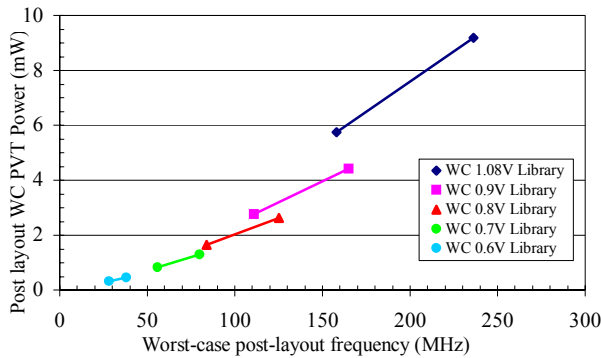


Figure 13. Speed to power tradeoff using Metro

The normalized power (uW/MHz) for both the high- and low-speed processor designs is plotted in Figure 14. The power efficiency improves dramatically: more than 3x from 1.08V to 0.6V, with a WC normalized power of 11.3uW/MHz achieved at 0.6V VDD.

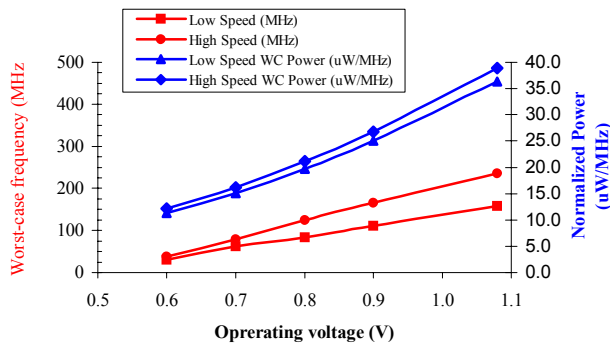


Figure 14. Speed to power efficiency tradeoff using Metro

### 5.3 Artisan Metro Dynamic Voltage/Frequency Scaling Results

Dynamic voltage and frequency scaling (DVFS) reduces power for low-performance task by allowing both the voltage and clock frequency to be scaled down when the immediate task requires less than full processor performance. The complementary benefit of application-specific extensibility is clearly exhibited in DVFS. By extending the processor, the operating frequency needed to reach a particular performance level is reduced, so clock and voltage can both be reduced - more than compensating for the increment in power dissipation per cycle from the extended processor logic.

Since low-power design platforms Metro and Mobilize both support DVFS features, the critical processor-circuit question is this: what maximum clock frequency and power is achieved at each voltage set-point? The high-speed version of the processor is used to extract the DVFS power reduction using data of Figure 14. The DVFS operational voltages are WC 0.6V, 0.7V, 0.8V, 0.9V, and 1.08V. The corresponding 19x dynamic power reduction can be achieved, during processor operational frequency downshift as shown in Figure 15. Note DFS (dynamic frequency scaling) alone would have only reduced the dynamic power by 6x as the processor downshifts from 240MHz to 40MHz.

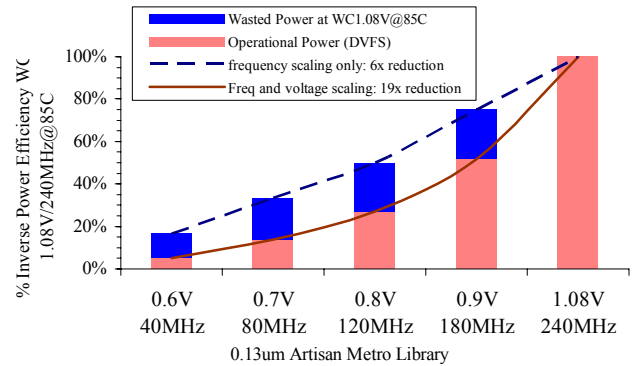


Figure 15. DVFS operational power using Metro

## 6. Leakage Dependence on Process, Voltage and Temperature

Dynamic power and energy efficiency are finally starting to be widely characterized and understood. The benchmarking consortium EEMBC is just starting to develop standardized methodology to measure energy consumption with processor benchmark tests [7]. Even today however, processor leakage power is rarely reported in the literature, let alone its process, voltage and temperature dependence.

This section evaluates leakage behavior for a processor design, using available leakage modeling in standard cell libraries. Using the high-speed processor version from Section 5.2 and Artisan SAGE-X standard cell libraries for LV and LVOD of the same standard Vth (threshold voltage), the process-dependent leakage power is shown in Figure 16. Originally, the leakage power in the libraries was not characterized at TT/1.1V/125°C. In order to assess the effect of process at the same set of voltage and temperature, the modeled TT/1.1V/125°C data point in Figure 16 is an average of 4 other TT data points (originally characterized in LVOD/LV) each being scaled by  $VDD \cdot \exp(-qV_{th}/KT)$ . Leakage power has exponential process dependence: over 16x leakage power difference is observed between the FF process corner and SS corner for the same Vth technology, at the same VDD and temperature.

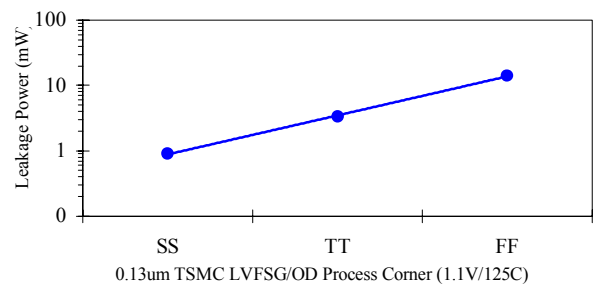


Figure 16. Leakage power process dependence

It is worth noting that leakage varies across the mix of code sequences used for power analysis by only a couple of per cent. Using both the high- and low-speed processor versions from Section 5.2, the voltage-dependent leakage power for the SS process corner at 85°C (WC) is shown in Figure 17. The leakage power density also linearly scales with VDD as previously seen in Figure 11. Though closely matched, the high-speed processor

tends to pick up more leak cells as reflected in the higher processor leakage density. Using the high-speed version processor and Artisan SAGE-X standard cell libraries for LV and LVOD processes, the leakage is shown to have exponential temperature dependence in Figure 18. TT corner at 25°C has the lowest leakage. Over 70x leakage power difference is observed between the PVT of FF/110%VDD/125°C, and the PVT of TT/VDD/25°C.

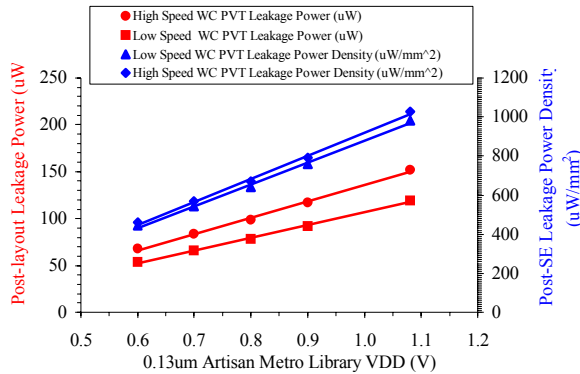


Figure 17. Leakage Power Saving and its Voltage Dependence

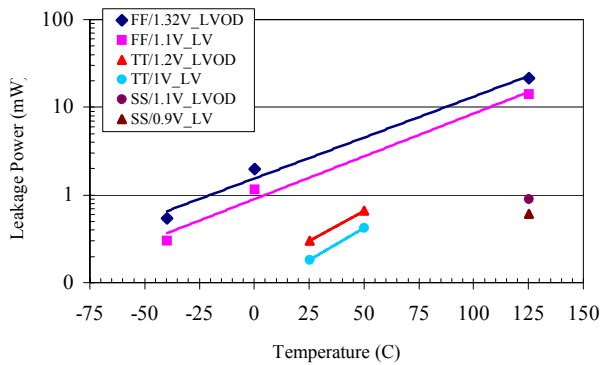


Figure 18. Exponential temperature dependence of leakage

## 7. Conclusions

Detailed analysis of a series of very-low-power processor implementations reveals significant tradeoffs between the design priority of speed and area: an increasingly larger gate count is needed to sustain a progressively faster processor design. Faster process technology benefits processor implementation in terms of less area and less power dissipation for the same speed

performance, at the expense of leakage power and its weight in the total power consumption depending on the application speed.

Tremendous power saving has been demonstrated in embedded processor implementation using the scaled VDD approach. More than 3x improvement has been observed in power efficiency at the lowest VDD, despite an increasing tradeoff between speed and area at lower VDD. DVFS enables 19x power reduction. The leakage has exponential dependences on temperatures and process corners for these processors. The results for nine variations of one processor configuration show significant dynamic and leakage power reduction through use of low-voltage libraries. These circuit-level energy savings complement the energy savings from architectural-level processor configuration. The absolute results for these cores will vary depending on the design, flows and methodology, among other factors. The data is only necessarily useful for comparing the tradeoffs and techniques discussed in this paper.

## 8. Acknowledgements

The authors are grateful for the helpful technical discussions with Jagesh Sanghavi, Eliot Gerstner, Eileen Peters Long, and Grant Martin of Tensilica, Dhruvil Gandhi of ARM and Dan Hillman of Virtual Silicon Technology, Inc.

## 9. References

- [1] Rowen, C. *Engineering the Complex SOC – Fast, Flexible Design with Configurable Processors*. Prentice Hall, Upper Saddle River, NJ, 2004, Chapters 7&8.
- [2] *System Drivers*, International Technology Roadmap for Semiconductors, 2003, 10-11. <http://public.itrs.net/>.
- [3] Artisan Components, Inc. <http://www.artisan.com/promo/metro.html>
- [4] Virtual Silicon Technology, Inc. [http://www.virtual-silicon.com/view\\_press\\_release.cfm?prid=83](http://www.virtual-silicon.com/view_press_release.cfm?prid=83)
- [5] Peters, E., Taglieri, G., and Vemury, L. *Low Power Synthesis Flow For a Configurable Core*. SNUG (Synopsys Users Group) Boston 2000.
- [6] Hillman, D., and Wei, J. *Implementing Power Management IP*. SNUG (Synopsys Users Group) Boston 2004.
- [7] EEMBC press release. *EEMBC Developing Standardized Methodology to Measure Energy Consumption*. Nov.9, 2004. <http://www.eembc.hotdesk.com>