International Journal of Engineering Research and General Science Volume 3, Issue 3, May-June, 2015
ISSN 2091-2730

Asynchronous Technology For Neural Signal Processor with Leakage Suppression

Varpe Prashant Bhikaji [1], Prof. S.M.Turkane[2]

[1]Department of E&TC Savitribai Phule Pune University, PREC Loni , INDIA

[2]Head of Department of E&TC Savitribai Phule Pune University, PREC Loni , INDIA

Email- prashvarpe@gmail.com

Abstract— Further power and energy reductions via technology and voltage scaling have become extremely difficult due to leakage and variability issues. In this paper , we present a robust and energy-efficient computation architecture by employing an asynchronous self-timed design methodology. We establish that asynchronous self-timed design methodologies inherently address the majority of concerns raised—avoiding the need for complicated, demanding and ultimately limited Band-Aid's to the synchronous design philosophy. By automatic transition to leakage-reduction, self-timed design aggressively suppresses standby power in the presence of unpredictable variations and robustness concerns. As such, it aids to effectively reduce the minimum-energy operational point of a digital circuit.

Spike sorting is an important processing step in various neuro scientific and clinical studies. Energy-efficient spike-sorting ASICs are necessary to allow real-time processing of multi-channel, wireless neural recordings. Spike-sorting ASICs have to meet stringent power-density constraints and must provide significant data-rate reduction for wireless transmission.

Keywords— Application-specific integrated circuits (ASIC), Process, voltage, and temperature (PVT), Return-to-zero (RTZ), Spike alignment and feature extraction (SAFE), Slow fast (SF), Fast slow (FS), Complementary metal oxide semiconductor (CMOS).

INTRODUCTION

Emerging biomedical and wireless applications would benefit from the availability of digital processors with substantially improved energy-efficiency. One approach to realize ultra-low energy processors is to scale the supply voltage aggressively to below the transistor threshold, yet the major increase in delay variability under process, voltage, and temperature (PVT) variations combined with the dominance of leakage power makes robust sub-threshold computations and further voltage scaling extremely challenging. Traditional synchronous design methodology deals with delay variability by first estimating the performance of a digital system at the worst-case process corners, and subsequently allocating extra timing margin to guarantee circuit functionality for the worst-case scenario.

This project is introduces an asynchronous self-timed design methodology as an attractive alternative for the realization of robust and energy-efficient computation. Asynchronous self-timed schemes allow for an adaptive adjustment to latency variations and support for an inherent leakage minimization under both static and dynamic variations, leading to a robust and energy-efficient sub-threshold computation architecture.

The major implementation challenge of asynchronous methodology is the additional handshaking protocol circuitry overhead required for self-timed operation. While the latter increases the dynamic energy consumption, it effectively enables a reduction of the overall minimum operational energy of the increasing number of applications in which leaka6ge plays a dominant role.

METHODOLOGY

We have decided to develop the system **"Design & Implementation of ultralow power Asynchronous Neural Signal Processor for brain machine interface With Leakage Suppression"** because my aim of the project is employing optimum circuit architecture and self-timed scheme to minimize the overall circuit leakage.

In our paper, we are basically concentrating on following applications such as

- *Introduces an ultra-low power asynchronous self-timed design methodology for brain-machine interface systems as an attractive alternative for the realization of robust and energy-efficient computation.*

- *The design methodology and circuit implementation of asynchronous processor are described in detail.*

- *Measure the results of proposed asynchronous and a traditional synchronous neural signal processor using simulation tool Hspice.*

Traditional synchronous design methodology deals with delay variability by first estimating the performance of a digital system at the worst-case process corners, and subsequently allocating extra timing margin to guarantee circuit functionality for the worst-case scenario. This methodology is seriously challenged in ultra-low voltage designs. Furthermore, the delay variability not only slows down the speed, but also degrades energy efficiency. Since the worst-case event happens very infrequently, a traditional synchronous digital circuit would stay idle most of time and consume extra leakage energy

In this project, we establish that asynchronous self-timed design methodologies inherently address the majority of concerns raised avoiding the need for complicated, demanding and ultimately limited Band-Aid's to the synchronous design philosophy. By automatic transition to leakage-reduction, self-timed design aggressively suppresses standby power in the presence of unpredictable variations and robustness concerns. As such, it aids to effectively reduce the minimum-energy operational point of a digital circuit. While the protocols and circuits introduced in this project may not be entirely new, its major contribution is establishing that self-timed methodologies can play a major role in a design world where aggressive voltage scaling is a must.

ASYNCHRONOUS DESIGN

Biomedical sensor is one of the emerging applications that demand for ultra-low power and energy performance. Because of safety regulations and miniature size constraints, the electrical integrated system embedded in a biomedical sensor must be extremely energy-efficient and reliable. In this paper, a neural signal processor performing neural spike-sorting function for a brain-machine interface system is implemented. Fig. 1 shows the block diagram and circuit schematics of the prototype asynchronous neural signal processor. The processor implements a neural spike-sorting function in three steps: spike detection, spike alignment and feature extraction. The power and area constraints of overall brain-machine interface system prohibit the implementation of additional spike-sorting functions such as clustering. We use the spike-sorting algorithms that exhibit the power-density characteristics best suitable for real-time wireless neural signal processing as derived in [12]. The processor receives the 8-b digitized data from a neural signal acquisition. front-end running at 20 kHz [13].

A. Asynchronous Handshaking Protocol:

Since the front-end ADC operates at a fixed sampling rate, the input synchronous bit streams are first converted into dual-rail return-to-zero (RTZ) data with the necessary request-acknowledge handshaking control signals by a synchronous-to-asynchronous interface composed of differential pre-charged dynamic logic buffers. The data processing thereafter is driven only by local handshaking events without a global synchronous clock, eliminating any timing uncertainty such as skew and jitter associated with clock distribution. The communication between each module is reliably governed by a self-timed 4-phase dual-rail handshaking protocol [9], as shown in Fig. 2. Each self-timed module has a 2-phase operation cycle, with a data evaluation phase followed by a data reset phase.

The request and acknowledge handshaking control signals are generated only after the data are successfully processed and stored, avoiding any setup- and hold-time violations. As a result, each self-timed block operates at its own speed reliably and achieves best-effort performance that dynamically adapts to the changing operating condition. Fig. 3 shows the asynchronous control flow and handshaking communication between different self-timed modules. The state transition diagram of neural signal processor is shown in Fig. 4.

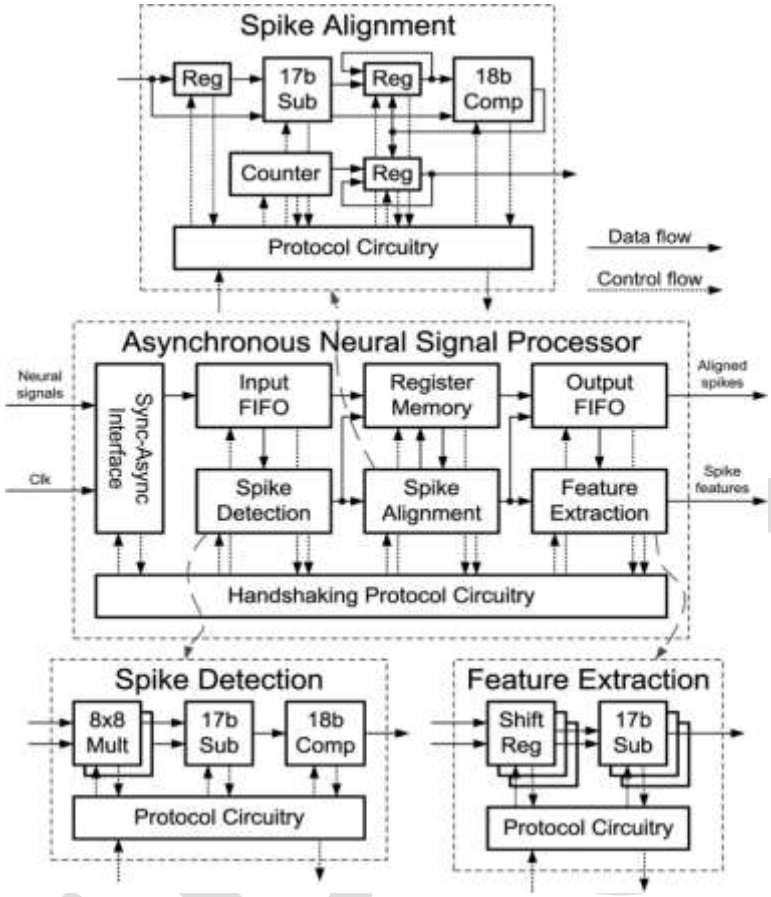

Fig. 1. Block diagram and circuit schematics of asynchronous neural signal processor.

B. Spike detector design :

The spike detector shown in Fig. 4 identifies the spike events by comparing the signal energy of the raw neural data to a threshold calculated during the initial training period. The spike detection process involves two 8- bit multiplications, a 17-bit subtraction and an 18-bit comparison, determining the critical path and thus the overall cycle time of the processor. Designed to operate at a 0.25 V supply with $50mV_{pp}$ supply noise, the detector exhibits switching activities from 0.2 to 0.01 depending on the operating condition and signal statistics, consuming mostly leakage power. Since the processor has already operated at the leakage-dominated regime, additional pipelined or parallel data processing would not reduce overall power consumption [14].

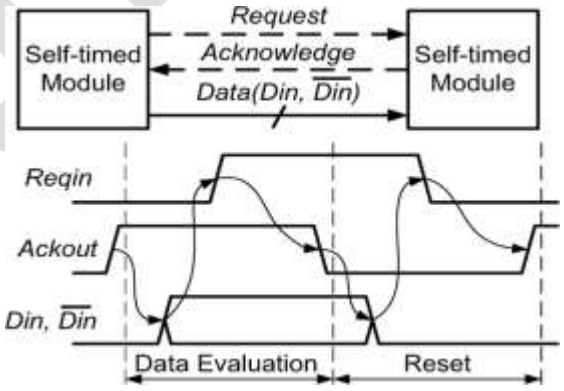

Fig. 2. Dual-rail asynchronous handshaking protocol.

The differential dynamic CMOS logic topology employed in this design significantly reduces the number of leakage paths by realizing arithmetic functions with fewer logic gates than static CMOS-based design, as demonstrated in a booth 2 decoder. The use of logic

gate with higher fan-in can further reduce overall gate counts and leakage paths, at the same time degrades circuit robustness in the presence of variations.

The upper bound of gate fan-in is determined by two worst-case process corners. At the slow NMOS/fast PMOS (SF) corner, the evaluation logic path consisting of long stacked NMOS transistors may become too weak to pull down the output node but. On the other hand, the leakage current contributed by parallel NMOS transistors could severely degrade output noise margin at the fast NMOS/slow , PMOS (FS) corner.

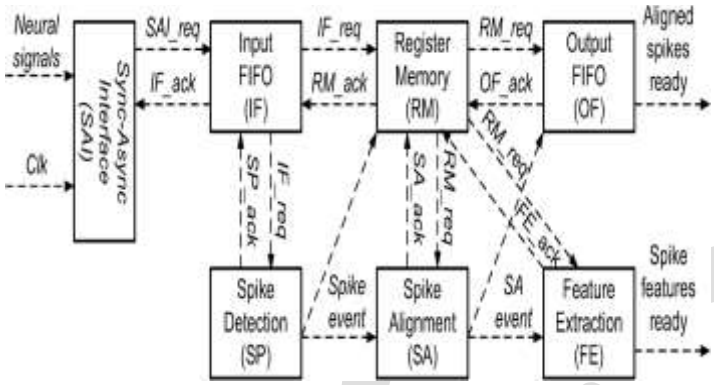

Fig. 3. Asynchronous control flow of self-timed neural signal processor.

As a result, most of logic gates in the data path were selected to have fan-ins of 4 achieving better leakage characteristics while maintaining output signal integrity. At the architectural level, algorithms such as booth encoding are employed to minimize the overall gate count. further reducing the total number of leakage paths by 1.5X The protocol circuitry regulates the dataflow and manages the transition between different operating modes. Therefore, immediately after a computation completes, a self-timed data path enters its standby mode to minimize leakage. The employment of low-leakage logic topology together with asynchronous timing effectively reduces the leakage of the processor data path by 10X compared to a static CMOS based synchronous design during normal operation.

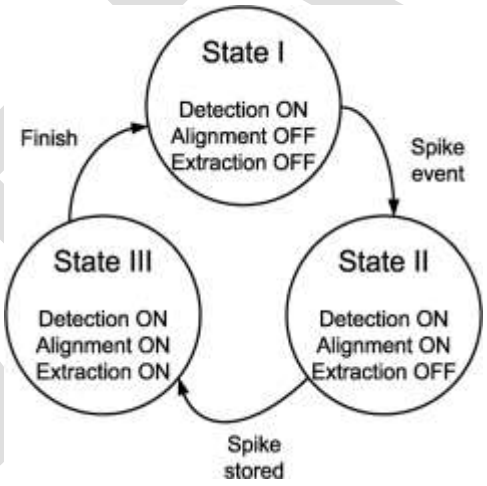

Fig. 4. State transition diagram of neural signal processor.

C. Spike alignment and feature extraction:

The identified spikes are then aligned to the sample with the maximum derivative to improve the classification accuracy. The aligned spikes are finally transformed into a feature space that better separates spikes from different neurons. The spike alignment and feature extraction (SAFE) module shown in Fig. 1 is activated only when a spike event is detected, therefore exhibiting a variable block activity ranging from 0.2 to 0.02 depending on the actual spike activity. The power consumption of the entire processor highly depends on the SAFE module activity. As a result, the active time of SAFE module has a significant impact on the overall power consumption. In a traditional synchronous digital system with a single clock domain, the processing latency of each module is fixed and determined by the operating clock frequency. In order to minimize the power contribution of the synchronous SAFE module, a

second clock or voltage domain specifically for SAFE module operation can be introduced. However, the use of multiple clock/ voltage domains does not only increase overall system complexity, but also introduces additional reliability issues at low supply voltages.

Those may offset the potential power saving and make this approach impractical for this application. A SAFE module with asynchronous self-time operation processes data adaptively according to the different operating conditions, resulting in variable processing latencies rather than a fixed latency found in conventional synchronous designs. This ensures that the SAFE module is active only for the shortest possible time in all conditions, avoiding any idle time and thus minimizing the power. The asynchronous SAFE module demonstrates 3.6X less power consumption than a synthesized synchronous counterpart.

PERFORMANCE ANALYSIS

Asynchronous timing strategies have been advocated to enhance the speed and ease the global clocking problem in high-performance digital systems [9]. However, the built-in timing mechanisms that automatically adapt to different operating conditions make asynchronous design an attractive alternative for the realization of ultralow-energy computation. Fig.5 shows the performance characteristics of generic synchronous and asynchronous systems. The black curve represents a typical delay distribution of digital system under process variations, while the red and blue curves represent resulted performance characteristics with synchronous and asynchronous timing strategies, respectively. By employing an additional handshaking protocol circuitry, an asynchronous self-timed system can track actual circuit behaviour under both static and dynamic variations. As a result, it realizes average-case performance in the presence of variability, representing an ultimate form of adaptive digital system.

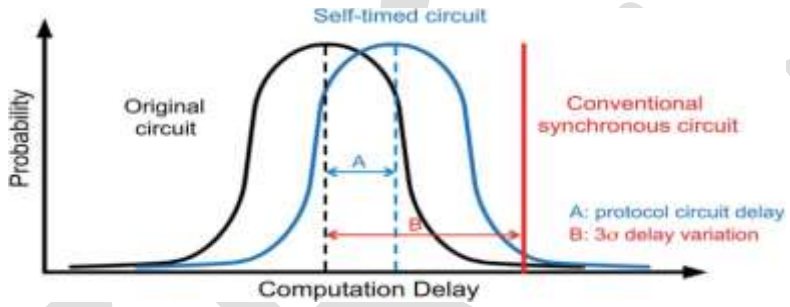

Fig. 5. Performance characteristics of original (black), synchronous (red) and asynchronous (blue) digital circuits.

This makes asynchronous design very promising for ultra-low voltage operation where variability has a strong effect on circuit performance. More detailed performance comparisons between synchronous and asynchronous designs can be found in [10], [11]. Even more importantly, asynchronous implementations can also achieve better leakage behaviour at both the logic and block levels, by operating the circuit in different power modes adapting dynamically to variations. Fig. 6 shows a comparison of power profiles between a synchronous and an asynchronous system. For biomedical applications operating at stable temperature environments, if 10% supply change and 3σ process variations are considered, a 0.25 V. Synchronous design must have a timing margin at least 9x larger than its typical delay, This extra timing slack results in extra idle time and leakage energy. An asynchronous design, on the other hand, responds immediately after a computation finishes, enabling it to either proceed to the next operation without idling, or to be switched to a low-leakage standby mode to minimize power consumption. An asynchronous design can exploit this leakage behaviour, adaptively adjusting to latency variations and switching between different power modes to minimize power consumption.

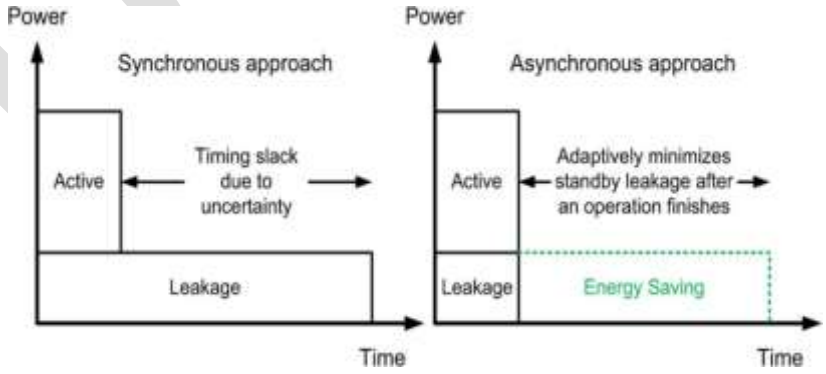

Fig. 6. Power profiles of synchronous and asynchronous systems.

CONCLUSION

We present a robust and energy-efficient computation architecture by employing an asynchronous self-timed design methodology. The proposed strategy allows for an adaptive adjustment to latency variations, and supports for an inherent leakage minimization under process variations and changing operating conditions, all of which are major issues in scaling regimes that favour major reduction in supply voltages.

Circuit techniques specifically for leakage minimization are aggressively employed at both the logic and system levels. Moreover, the self-timed operation alleviates the impact of variations on processor performance. Self-timed design aggressively suppresses standby power in the presence of unpredictable variation and robustness concern. Therefore, the asynchronous design exhibits a better statistical characteristic of power performance than the synchronous counterpart.

REFERENCES:

[1] A. Wang and A. Chandrakasan, "A 180-mV sub threshold FFT processor using a minimum energy design methodology," *IEEE J. Solid- State Circuits,* vol. 40, no. 1, pp. 310–319, Jan. 2005.
[2] L. Pang and B. Nikolic, "Measurements and analysis of process variability in 90 nm CMOS," *IEEE J. Solid-State Circuits,* vol. 44, no. 5, pp. 1655–1663, May 2009.
[3] N. Ickes, G. Gammie, M. E. Sinangil, R. Rithe, J. Gu, A. Wang, H. Mair, S. Datla, B. Rong, S. Honnavara-Prasad, L. Ho, G. Baldwin, D. Buss, A. P. Chandrakasan, and U. Ko, "A 28 nm 0.6 V low power DSP for mobile applications," *IEEE J. Solid-State Circuits,* vol. 47, no. 1, pp. 35–46, Jan. 2012.
[4] S. Hanson, B. Zhai, M. Seok, B. Cline, K. Zhou, M. Singhal, M. Minuth, J. Olson, L. Nazhandali, T. Austin, D. Sylvester, and D. Blaauw, "Exploring variability and performance in a sub-200 mV processor," *IEEE J. Solid-State Circuits,* vol. 43, no. 4, pp. 881–891, Apr. 2008.
[5] N. Kurd, P. Mosalikanti, M. Neidengard, J. Douglas, and R. Kumar, "Next generation Intel core micro-architecture (Nehalem) clocking," *IEEE J. Solid-State Circuits,* vol. 44, no. 4, pp. 1121–1129, Apr. 2009.
[6] D. Bull, S. Das, K. Shivashankar, G. S. Dasika, K. Flautner, and D. Blaauw, "A power-efficient 32 bit ARM processor using timing-error detection and correction for transient-error tolerance and adaptation to PVT variation," *IEEE J. Solid-State Circuits,* vol. 46, no. 1, pp. 18–31, Jan. 2011.
[7] K. A. Bowman, J. W. Tschanz, S.-L. L. Lu, P. A. Aseron, M. M. Khellah, A. Raychowdhury, B. M. Geuskens, C. Tokunaga, C. B. Wilkerson, T. Karnik, and V. K. De, "A 45 nm resilient microprocessor core for dynamic variation tolerance," *IEEE J. Solid-State Circuits,* vol. 46, no. 1, pp. 194–208, Jan. 2011.
[8] M. Fojtik, D. Fick, K. Yejoong, N. Pinckney, D. Harris, D. Blaauw, and D. Sylvester,"Bubble Razor: An architecture-independent approach to timing-error detection and correction," *IEEE Solid-State Circuits Conf. Dig. Tech. Papers,* 2012, pp. 488–490.
[9] J. Sparsø and S. Furber, *Principles of Asynchronous Circuit Design.* Norwell, MA, USA: Kluwer, 2001.
[10] K. Stevens, P. Golani, and P. Beerel,"Energy and performance models for synchronous and asynchronous communication," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.,* vol. 19, no. 3, pp. 369–382, March 2011.
[11] T.-T. Liu and J. Rabaey, "Statistical analysis and optimization of asynchronous circuits," in *Proc. 18th IEEE Int. Symp. Asynchronous Circuits and Systems,* May 2012, pp. 1–8.
[12] V. Karkare, S. Gibson, and D. Marković, "A 130µw- , 64-channel neural spike-sorting DSP chip," *IEEE J. Solid-State Circuits,* vol. 46, pp. 1214–1222, May 2011.
[13] R. Muller, S. Gambini, and J. M. Rabaey, "A 0.013 , 5 mW, DC-coupled neural signal acquisition IC With 0.5 V Supply," *IEEE J. Solid-State Circuits,* vol. 47, no. 1, pp. 232–243, Jan. 2012.
[14] D. Marković, C. C.Wang, L. Alarcón, T.-T. Liu, and J. Rabaey, "Ultralow- power design in near-threshold region," *Proc. IEEE,* vol. 98, no. 2, pp. 237–252, Feb. 2010.
[15] M. Chae, W. Liu, Z. Yang, T. Chen, J. Kim, M. Sivaprakasam1, and M. Yuce, "A 128-channel 6 mW wireless neural recording IC with on-the-fly spike sorting and UWB transmitter," in *IEEE Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, 2008*, pp. 146–147.
[16] Tsung-Te Liu and Jan M. Rabaey, *" A 0.25 V 460 nW Asynchronous Neural Signal Processor With Inherent Leakage Suppression," IEEE J. solid-state circuits,* vol. 48, no. 4, pp.897-906, April 2013.
[17] Ian Kuon, Russell Tessier and Jonathan Rose, " FPGA Architecture: Survey and Challenges," *Foundations and Trends in Electronic Design Automation,* vol. 2, no. 2, pp. 135–253, 2007