Low-Power Multi-Processor System Architecture Design for Universal Biomedical Signal Processing

Li-Fang Cheng, Tung-Chien Chen, and Liang-Gee Chen

Graduate Institute of Electronics Engineering, National Taiwan University, Taiwan

{fannycube, djchen, lgchen}@video.ee.ntu.edu.tw

Abstract— As the study of the wireless body area sensor network (BASN) keeps growing, corresponding applications such as bio-signal processing for personal healthcare are gaining more attention. To realize a miniature and multi-functional system for biomedical applications, the design of on-sensor mircoprocessor is considered to be a solution. However, how to balance the power consumption and system scalability remains a challenge. In this work, we try to involve the concept of multi-core computing systems into biomedical applications. A fully programmable system architecture based on the proposed two-way pipeline processing unit (two-way PPU) is introduced and evaluated. The proposed two-way PPU is a computing unit that adopts the general purpose processor (GPP) to provide high system programmability, and is easy to be cascaded for system extension. According to the implementation results, the proposed architecture can save up to 91% of energy compared with conventional multi-core architecture in a four-core biomedical processing microsystem.

I. INTRODUCTION

The emerging development of body area sensor network (BASN) is changing the way how people can involve in their health. Through the connection of wireless front-end sensors and multi-functional biomedical signal processors (BSPs), real-time bio-signal processing and retrieval of instantaneous health status beyond conventional medical environments may become available [1]. For instance, some individual-centered healthcare applications, such as remote heart monitoring and long-term epileptic seizure tracking, are limited by the form factor of the monitoring devices and the insufficiency of advanced processing functions. These applications can thus gain benefit from the development of powerful on-sensor BSPs.

There are two major concerns for the hardware design of on-sensor BSPs. First of all, the energy efficiency is a significant factor. For most biomedical applications, continuous monitoring of bio-signals is often required, and the power consumption of corresponding BSPs should be minimized in order to support reasonable battery life. In addition, the flexibility of the BSP design is also important due to various physiological situations of each individual. The development of on-sensor BSPs with limited functions for some specific applications is a common solution so far. However, it would be more reasonable to design the BSP with proper reconfigurability to meet the requirements of different users.

Recently a variety of hardware architectures for biomedical signal processing have been proposed. Compared with the system architecture fully consisted of application–specific
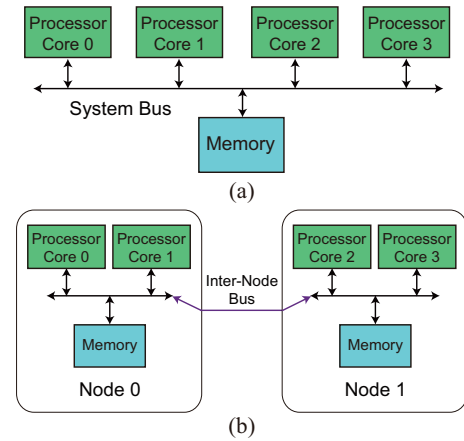

Fig. 1. Two typical architectures using multiple computing cores. (a) The SMP-based type, and (b) the NUMA-based type.

integrated circuit (ASIC), the system integrated with the GPP can provide more programmability. However, according to our knowledge, most of the current designs integrate at most one GPP into the system [2] [3], and only simple tasks are demonstrated on the GPP. When there are multiple complicated algorithms to be programmed, using only one GPP may not be enough to maintain similar performance. Therefore, a system architecture with multiple GPPs is a possible direction to solve the problem.

In this work we propose a pipelined multi-core system architecture that aims to support universal biomedical signal processing. We try to design and optimize both the flexibility and energy efficiency of a fully reconfigurable system by using multiple GPPs. The rest of the paper is organized as follows. The proposed programmable architecture is described in section II. Section III gives the implementation results and discuss the performance on real biomedical applications. Finally, section IV summarizes this work and gives the conclusion.

II. PROPOSED SYSTEM ARCHITECTURE

A. Review of Conventional Multi-core Architectures

Figure 1 (a) illustrates the typical symmetric multiprocessing (SMP) architecture of the processor with multiple computing cores. Each core of the processor uses the single shared bus to access one shared memory. Under this architecture the
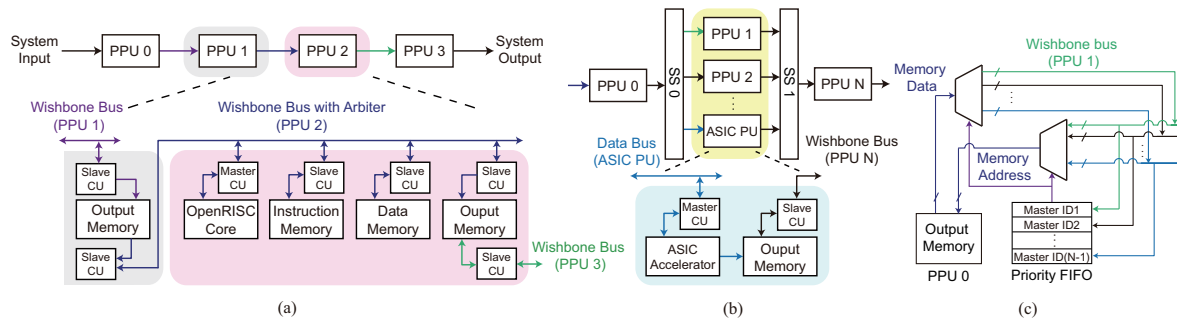
Fig. 2. (a) The proposed system architecture using the proposed two-way pipeline processing units (PPUs). (b) The parallel extension of the proposed architecture with the slave switcher (SS). The ASIC processing unit (PU) can be connected with the two-way PPUs by adopting the same I/O protocol. (c) The block diagram of the slave switcher (SS 0) in (b).

overall system efficiency depends on the frequency of simultaneous memory access of the multiple cores. An improved type of NUMA (non-uniform memory access) architecture is given in fig. 1 (b). The NUMA architecture grouped the multi-core system into different clusters (or nodes) to enhance the performance of memory access. In a NUMA system, the speed to access the memory increases if the computing core and its target memory are in the same cluster. However, the multiple cores in the same cluster still share the same bus when accessing the memory. Besides, the performance would decrease in cases with frequent inter-node data access.

B. The Two-way Pipeline Processing Unit (PPU)

Due to its improvement on intra-node memory access, the NUMA architecture seems to be a proper choice for our multi-core design. The current problem is the inefficiency for inter-node data access. For example, if we would like to realize a pipeline processing flow using multiple nodes, the bus conflict rate would increase because the sequent pipeline nodes must exchange a certain amount of stored data for processing. However, for most biomedical signal processing, the processing flow can be divided into multiple serial stages, such as preprocessing, transformation, feature extraction, and classification, and data transmission between different stages is unavoidable. In order to adopt the advantages of the NUMA system as well as solve its limitation for data exchange, we propose an architecture based on the two-way pipeline processing unit (two-way PPU) for the design of multi-core biomedical microsystems.

Figure 2 (a) illustrates the proposed system architecture and the block diagram of the two-way PPUs. Each processing unit is consisted of four major modules: the processing core, the private memory used for the storage of instruction and data, and the output memory with both internal and external connections. Each bus is with an arbiter for control, and each module connects to the arbiter through a master or slave control unit (CU). Note that in this study we use the 32-bit OpenRISC, OR1200 [4], as the computing core for testing. In the proposed architecture, each computing core connects to the instruction and data memory through the intra wishbone bus directly. The processing cores do not share the bus to access

the memory as the way described in fig. 1 (a) and (b) when there is no need to fetch the data from other computing cores.

In order to exchange the data between different PPUs efficiently, an extra output memory module is adopted to store the data that may be accessed by other cores. To achieve this function, the output memory module for each PPU is connected to both intra and inter wishbone buses. That is, this memory can be accessed not only by its PPU, but also other PPUs whose wishbone bus is connected to it. Through a pre-defined master-slave control protocol that is recognized by all PPUs, the PPUs can be cascaded and pipelined.

A simplest system architecture using the proposed PPU is given in fig. 2, and an example of the pipelined system working schedule is also shown in fig. 3. In the case demonstrated here, a specific memory address is set to be an enable flag for the next stage to realize the pipelined processing flow. The protocol for data exchange can be further designed and modified flexibly at the software level by the designer.

C. Expansion of the Architecture

In addition to the characteristic of pipelining, the proposed system architecture using two-way PPUs can be further extended to a parallel architecture through the control unit of the inter-buses. An intermediate switching module, slave switcher (SS) consisted of a MUX and a DMUX module is designed to control the connection. Figure 2 (b) shows an example of a parallel expansion of the PPU-based architecture with the slave switcher. The block diagram of the slave switcher is also given in fig. 2 (c). In this system, all the PPUs in the second stage work in parallel. The PPU at the third stage (PPU N) can then read the input data from either of the PPU through the slave switcher. The slave switcher connects the wishbone bus of PPU N to the output memory of PPUs in the second stage. The dimension of the system can thus be adjusted flexibly to meet the requirements of target applications. As for biomedical signal processing, the parallel architecture is especially suitable for channel expansion or high-dimensional feature extraction.

Finally, it should be noted that the architecture can be easily integrated with other design as long as the protocol of data exchange is the same as that of the PPUs. For example, the
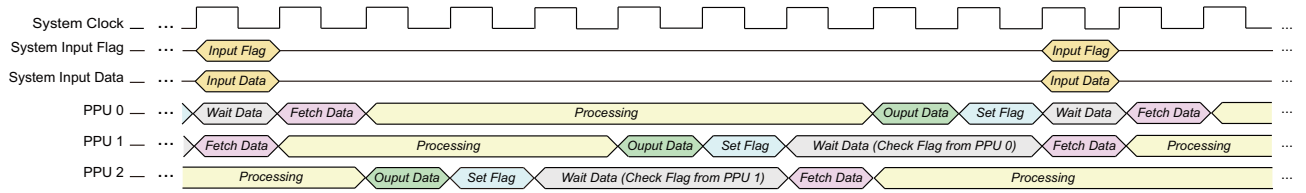
Fig. 3. The general working schedule of the PPUs in fig.2. The duration of each working phase depends on the programmed applications.

TABLE I
SYNTHESIZED RESULTS OF THE PROPOSED SYSTEM USING THE
TWO-WAY PIPELINE PROCESSING UNITS

Process	UMC 65nm Low Leakage CMOS
Supply Voltage	1.08 Volt
Operating Frequency*	15MHz
Area	1.09mm^2
Memory	128KB
Logic Gate (w/o memory)	101k

*The supported maximum clock speed of this implementation.

TABLE II
A COMPARISON OF THE POWER CONSUMPTION USING THE SMP-BASED
SYSTEM AND THE PROPOSED TWO-WAY PPU

System	Operating Frequency	Power
Four-core SMP	110MHz	10.7mW
Proposed	12MHz	903μW

computing cores in the proposed architecture can be replaced with specific computing units, such as ASIC accelerators demonstrated in fig. 2 (b).

III. VLSI IMPLEMENTATION RESULTS AND PERFORMANCE EVALUATION

A. Hardware Implementation Results

The system shown in fig. 2 (a) is implemented in this work to evaluate the proposed system architecture using the two-way PPUs. Four PPUs are used to construct a four-stage pipeline system. A serial-to-parallel programming interface is used for configuring the instruction memory of the two-way PPUs. Table I summarizes the overall implementation results. In this work, the UMC 65nm low–leakage process is adopted for circuit synthesis. The system operating frequency is 15MHz, while it can be adjusted case by case. Totally 101k logic gates and 128KB SRAM are used, where the size of of the instruction and data memory for each two-way PPU is 16KB respectively. In order to improve the power efficiency, the clock gating technique is adopted in this implementation. More details of the power analysis is given as follows.

B. Performance Evaluation

In order to evaluate the PPU-based system architecture, we use a four-stage cascaded finite impulse response (FIR) filter for simulation and comparison. The FIR filter is commonly used for the preprocessing of bio-signals. It is a convolution-based procedure which can be expressed as

$$y[n] = \sum_{i=0}^{N} a_i x[n-i] \qquad (1)$$

, where $x[n]$ is the input signal, a is the filter coefficients and N is the taps of the filter. In this evaluation case, a four-stage FIR is used for testing. Each computing core performs a 200-tap FIR filter respectively. We assume that the sampling rate of the

input signal is 250Hz with 12-bit resolution, which is a general sampling configuration for some bio-signals such as clinical electrocardiogram (ECG) signals. The system is required to perform a real-time operation under this situation.

Table II gives the estimated power consumption of the proposed system using the testing case. The power estimation for a four-core SMP system like fig. 1 (a) is also given for comparison. All the power data given here is simulated on gate–level design and analyzed by primetime. In the SMP system, the global memory is estimated to be the summation of the instruction and data memory of the four PPUs in the proposed system. Since the memory is shared by four cores through the global bus, the required operation frequency and power consumption increase in this case.

Figure 4 illustrates the effect of the power reduction of the proposed two-way PPU-based system with different number of cores. In this simulation, each core in the system performs the FIR case the same as above. It can be observed that the proposed system can work more efficiently than the SMP system. The percentage of power saving increase as the number of cores increases. In the four-core system, about 85% of the power is saved by using the proposed system. An average of 44% of power consumption of the proposed system can be further saved by adopting the clock gating technique according to the simulation. Therefore, about 91% of the total power can be reduced through this implementation.

C. Example of Biomedical Signal Processing: Running ECG Analysis Using the Proposed System

In the last part, we also choose the ECG analysis as an example to demonstrate how the proposed architecture can support real biomedical applications. There are several research fields in ECG analysis, such as PQRST feature extraction, ischemia detection or arrhythmia detection. Since the health status and thus system requirements for each individual can be very different, a flexible platform like the proposed system is suitable for this kind of applications.

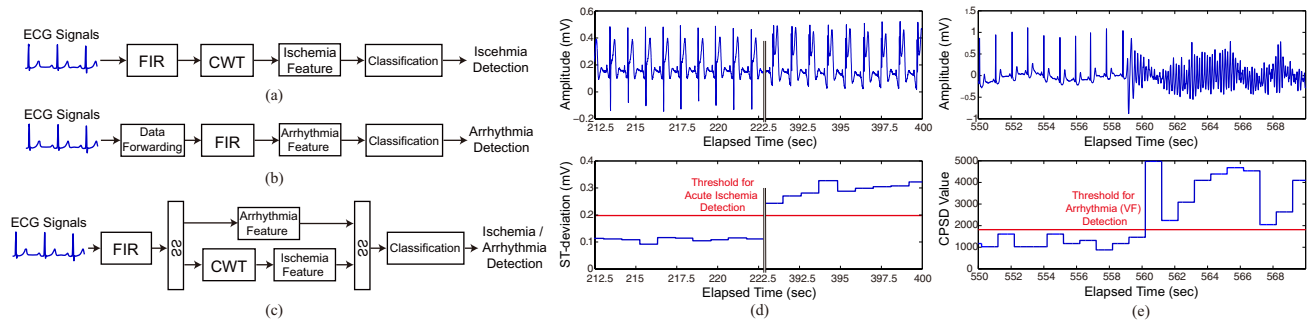Here, two different algorithms are programmed on the

Fig. 5. The functional block diagrams for different processing cases using the implemented hardware: (a) the ischemia detection, (b) the arrhythmia detection. (c) An extension of the architecture to support both ischemia and arrhythmia detection. (d) The simulation results of (a) tested on an ischemia case with ST-segment elevation in European ST-T Database. (e) The simulation results of (b) tested on an arrhythmia case (VF) in MIT Sudden Death Database.
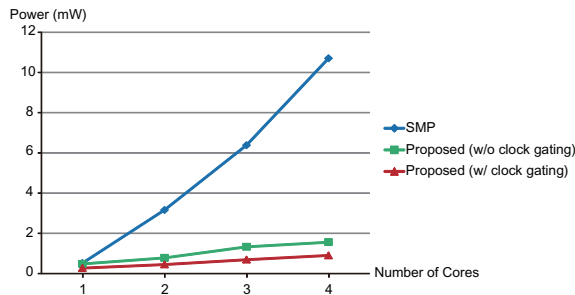

Fig. 4. The power comparison of the SMP and the proposed system.

implemented system. We use ECG data from the standard MIT-BIH databases [5] for simulation. In the first case, we programmed the algorithm based on continuous wavelet transform (CWT) for PQRST extraction in [6], and then use the ST-deviation feature to perform ischemia detection [7]. In the second case, the Chaotic Phase Space Differential (CPSD) algorithm [8] for arrhythmia detection is programmed. The functional block diagrams about how the two tasks can be arranged in the system are given in fig. 5 (a) and (b) respectively. Two examples of the simulation results are also shown in fig. 5 (d) and (e).

In the third case, in order to build a system consisted of both functions described above, the designer can simply add only two slave switchers and one more PPU, as shown in fig. 5 (c). The common preprocessing part of the two algorithms like FIR filtering can be used as a shared stage. This demonstrates the concept for system flexibility and scalability using the proposed architecture. Table III summarizes the corresponding power consumption of the three evaluation cases.

IV. CONCLUSION

In this work, we design and implement a pipeline multi-core system based on the proposed two-way PPU for general biomedical processing. The proposed system adopts the advantage of NUMA architecture and improve the efficiency of inter-core data exchange by using a two-way protocol through the partition of output memory module. According to the simulation results, the proposed system can provide up to 91% of energy saving compared with the conventional SMP architecture in a four-core processing system. Therefore, the proposed architecture is considered to be suitable for the general biomedical applications.

ACKNOWLEDGMENT

REFERENCES

[1] M.A. Hanson and et al., "Body area sensor networks: Challenges and opportunities," *Computer*, vol. 42, pp. 58–65, 2009.
[2] Fan Zhang and et al., "A batteryless 19 μw mics/ism-band energy harvesting body area sensor node soc," in *IEEE International Solid-State Circuits Conference (ISSCC)*, feb. 2012, pp. 298 –300.
[3] Shu-Yu Hsu and et al., "A sub-100-μw multi-functional cardiac signal processor for mobile healthcare applications," in *2012 Symposium on VLSI Circuits (VLSIC)*, june 2012, pp. 156 –157.
[4] OpenRISC Project, ," *http://opencores.org/openrisc,or1200*.
[5] Ary L. Goldberger and et al., "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101:e215-e220, 2000.
[6] I. Romero and et al., "R-wave detection using continuous wavelet modulus maxima," in *Comput. Cardiol.*, Sept. 2003, vol. 30, pp. 565–568.
[7] T. A. Fischell and et al., "Real-time detection and alerting for acute st-segment elevation myocardial ischemia using an implantable, high-fidelity, intracardiac electrogram monitoring system with long-range telemetry in an ambulatory porcine model," *Journal of the American College of Cardiology*, vol. 48, no. 11, pp. 2306–2314, 2006.
[8] Chien-Sheng Liu and et al., "Chaotic phase space differential (cpsd) algorithm for real-time detection of vf, vt, and pvc ecg signals," in *4th European Conference of the International Federation for Medical and Biological Engineering*, vol. 22. 2009.