International Journal of Electrical & Electronics Research. (IJEER)
Volume 4, Issue 1, Pages 10-15, March 2016, ISSN: 2347-470X

Double Pumping Low Power Technique for Coarse - Grained Reconfigurable Architecture

S. Munaf[1]
Assistant Professor
Department of ECE, Sri Ramakrishna Institute of Technology, Coimbatore-641042, India
munafece@gmail.com

Dr. A. Bharathi[2]
Professor
Department of IT, Bannari Amman Institute of Technology, Sathyamangalam-638401, India

Dr. A. N. Jayanthi[3]
Associate Professor
Department of ECE, Sri Ramakrishna Institute of Technology, Coimbatore-641042, India

Abstract— Coarse-grained reconfigurable architectures (CGRAs) require many processing elements (PEs) and a con- figuration memory unit (configuration cache) for reconfiguration of its PE array. Though this architecture is meant for high performance and flexibility. Power reduction is very crucial for CGRA to be more competitive and reliable processing core in embedded systems. We propose a DDR SDRAM (Double Data Rate Synchronous Dynamic Random Access Memory) architecture to reduce power-overhead caused by reconfiguration. The power reduction can be achieved by using the characteristics like double pumping the data bus and an I/O buffer between the memory and the data bus of DDR SDRAM. All modules have been designed at behavioral level with VHDL coding and to Simulate in Xilinx ISE navigator.

Keywords— Coarse- grained reconfigurable architecture, configuration cache, embedded system, loop pipelining, low power.

1. INTRODUCTION

To provide high quality multimedia on mobile and embedded systems, efficient algorithms for audio/video data transfer and processing have been developed. These algorithms are complex and have computation-intensive and data-parallel characteristics. For such applications, two extreme approaches are used for their implementation: software running on a general purpose processor hardware in the form of application - specific integrated circuit (ASIC). In the case of general processor, it is flexible enough to support various applications but they may not provide sufficient performance to cope with the complexity of the applications. In the case of ASIC, one can optimize the implementation in terms of power and performance but they rely on application.

A coarse-grained reconfigurable architecture (CGRA) can provide the advantage of both the approaches. CGRA has higher performance than general purpose processor and wider applicability than ASIC. In spite of the previous advantages, the deployment of CGRA is prohibitive due to its significant power consumption. This is due to incorporation of many computational resources such as algorithmic logic unit (ALU), multiplier, divider, and configuration cache to perform frequent memory read operations for dynamic reconfiguration in every cycle.

The configuration cache is the main component in CGRA that provides distinct feature for dynamic configuration. Even though configuration cache plays an important role for high performance but suffers from large power consumption. Therefore, reducing power consumption in the configuration cache has been a serious concern for reliability of CGRA.

For Low power CGRA design, this paper provides an optimized Low Power Technique in the configuration cache and its hardware implementation. In this paper, we suggest a novel power-conscious architectural technique called to Double Pumping Technique(DPT) reduce the data transfer time and also increase the operating speed and power consumption in configuration cache. DPT is universal approach in reducing power and enhancing performance for CGRA because it can be achieved by closing the power-performance gap between clock periods. High performance achieved increasing the data transfer rate. This has been demonstrated by using real application benchmarks and gate level simulations. This paper is organized as follows. After mentioning the related work in Section II, we discussed base architecture in Section III. In Section IV, we present the motivation of our approach. In Sections V and VI, we suggest power-conscious techniques based DPT.

2. RELATED WORK

Many kinds of coarse-grained reconfigurable architecture have been proposed with the increasing interests in reconfigurable computing in recent years [1]. These CGRAs can be classified into two cases: mesh-based reconfigurable array and linear reconfigurable array. Mesh-based reconfigurable arrays arrange their processing elements (PEs) mainly as a rectangular 2-D array with both horizontal and vertical connections, which supports efficient parallelism. In the case of linear reconfigurable arrays, they support

pipelined execution for stream-based applications with static or dynamic reconfiguration. MorphoSys [8] and REMARC are representations of mesh-based architectures. MorphoSys consists of Tiny_{R}ISC processor, Reconfigurable Cell array, frame buffer, context memory and DMA controller. RC array is an8x8 array of ALUs that performs 16-bit operations based on single instruction, multiple data (SIMD) programming model. REMARC consists of a global control unit and an 8x 8 array of nano processors. The configuration for each nano processor is stored in the 32-entry instruction RAM to support multiple instruction stream, multiple data stream (MIMD) execution model as well as SIMD model. RaPiD and PipeRench have linear array structure. RaPiD architecture provides different computing resources and these resources are irregularly distributed on one dimension which are static reconfigured. Instance of CGRA considering area and performance, they do architectures have been implemented with various technologies. In our concept we are implementing DDR2SDRAM as reconfiguration of cache and PE array.

3. BASE ARCHITECTURE

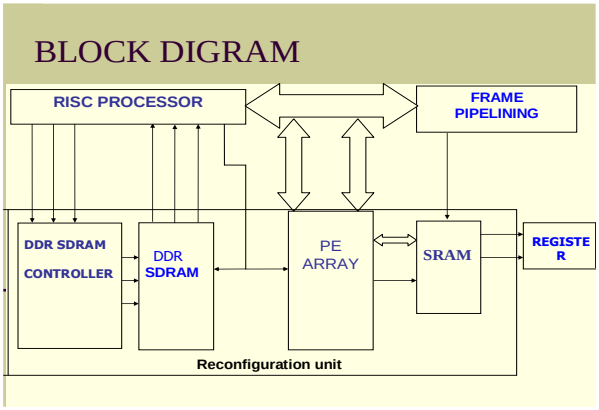


Figure 1. New CGRA Schematic

The above architecture involves the RISC processor which has the data bus and address bus. The fetching, decoding, execution functions are done here .For the functioning of these we need to have some kind of controller. These functioning can be done by DDR SDRAM controllers. These functions will reduce the overall complexity. Processing elements (PE) executes arithmetic and Logical Operations .The input data will be processed according to the RD/WR, Chip select, Data bus and address busses. The pipelining technique will be used to reduce the power.

The configuration cache is the main component in CGRA that provide distinct feature for dynamic configuration. Therefore, reducing power consumption in the configuration cache has been a serious concern for reliability of CGRA. In this unit contain cache controller and memory unit.

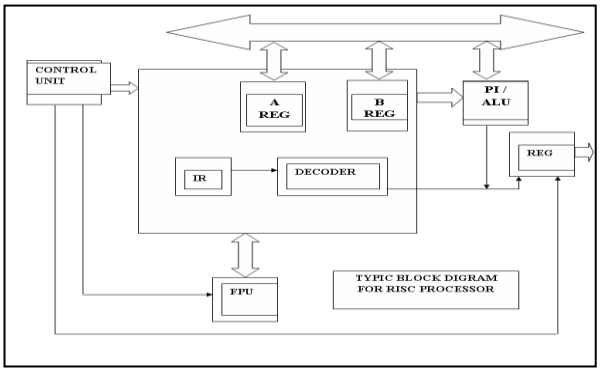


Figure 2.Risc Processor Schematic

CGRA has higher performance 16 bit processor and wider applicability than ASIC. The Processor which has the data bus and address bus. The fetching, decoding, execution functions are done here. The Control Signals from the control units send to Execution Unit and Memory Unit (ROM), based on the context word in the IR register the corresponding functions are processed

The data computations are done in **A** (accumulator) register and **B** register with help of ALU, finally results are stored in 16 bit **A** register and processor output register. The ALU functions are predefined in ROM. This processor ALU operations are done in the DDR RAM.

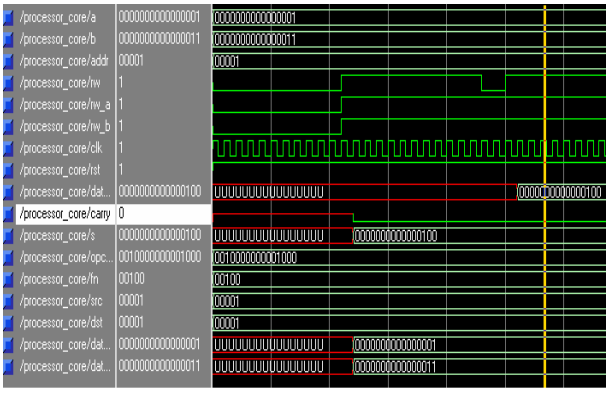


Figure 3.RISC Processor Module Result

The above figure 4 shows the simulated result for RSIC processor modules. The modules inner blocks are designed with VHDL coding and simulated by model sim. The simulated result shows the processor's arithmetic and logical computation output for given two 16 bit data stored in Reg **A** and Reg **B**, the same result on output register and it is taken as processor module final output.

4. RECOFIGURATION UNIT

We have implemented the base architecture shown in Fig. 1 at the RT-level with VHDL [11].We have synthesized a gate-level circuit from the VHDL description and analyzed power cost. The synthesis has been done using Design Compiler with 0.18- m technology. We have used Design-Ware library for the multipliers (carry-save array synthesis model). DDRSDRAM Macro Cell library is used for the frame buffer and configuration cache. Model Sim and Prime Power or Xilinx tools are used have been used for gate-level simulation and power estimation, respectively. To obtain power breakdown data, we have used 2-D-FDCT as the kernel for simulation-based power measurement. The simulation has been done for the typical case under the operating condition of 100 MHz frequency, 1.8 V and 27 C temperature. As be observed from Fig.4a the CGRA spends about 89% of the total power consumed in RAA. Fig. 4b shows more detailed power breakdown the RAA.

The RAA spends about 48% of the total power consumed in the PE array, which consists of many components such as ALUs, multipliers, shifters, and register files. The PE array consumes maximum power, which is quiet natural because coarse-grained architecture aims to achieve high performance and flexibility with plenty of resources. The configuration cache spends about 43% of the total power, which is the second maximum power constraint. Even though the frame buffer uses the same kind of SRAM in the [13]. But in this paper we are combine both PEs and cache in a single unit named as reconfiguration unit , this reconfiguration is known as architecture reconfigure technique .

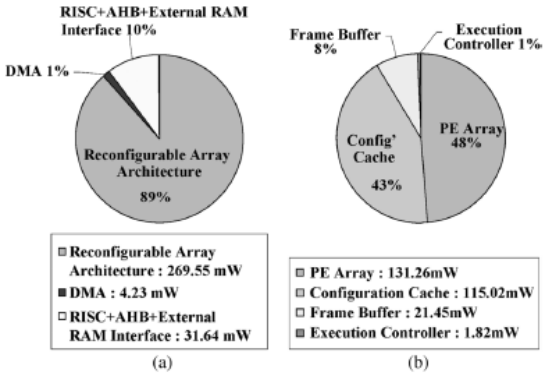


Fig. 4. Power cost breakdown for CGRA running 2-D-FDCT: (a) entire CGRA and (b) RAA

5. MOTIVATION

We propose a DDR SDRAM (Double Data Rate Synchronous Dynamic Random Access Memory) architecture to reduce power-overhead caused by reconfiguration. The power reduction can be achieved by using the characteristics like double pumping the data bus and an I/O buffer between the memory and the data bus of DDR SDRAM. Figure5 describe the data transfer.

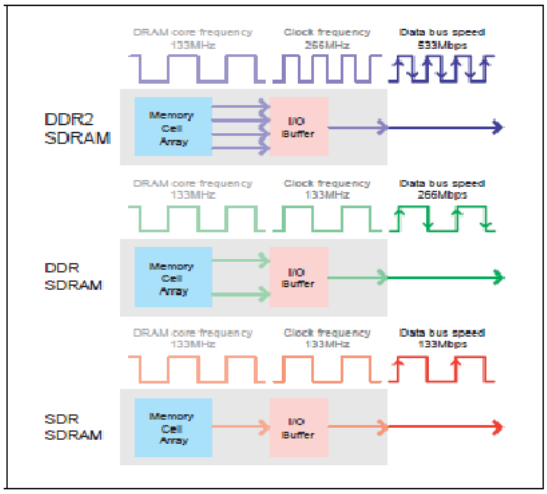


Figure 5. Double Data Transfer Functional Diagram

6. POWER-CONSCIOUS TECHNIQUES BASED DPT

We re-configure the entire section of the processes is changing the cache section by DDR SDRAM. The architecture involves the RISC processor which has the data bus and address bus. The fetching, decoding, execution functions are done here .For the functioning of these we need to have some kind of controller. These cache and controllers are going to be replaced by DDR SDRAM controllers. These functions will reduce the overall complexity. The PE (processing element) will be changed to have its ALUs and Shift Functions. The input data will be processed according to the RD/WR, Chip select, Data bus and address busses. The pipelining technique will be used to reduce the power .All these processes will be done to simulate in Xilinx ISE navigator.

Addressing memory in a DDR SDRAM memory requires four separate addresses: Chip Select, Bank Select, Row Address and Column Address.
In DDR SDRAM there are two memory chips connected in parallel, with unique chip enable signal. This configuration allows the two chips to share address and data lines. By selectively asserting only one chip enable single at a time, this configuration allows twice the memory depth compared with signal chip. Chip enable signals are controlled by highest level of the memory addressing modes. This addressing level is called the chip selects.

The remaining three addressing levels all take place within a single memory chip. Figure-3 shows a simplified block diagram of the internals of a DDR SDRAM memory chip. At the core of the memory chip are four 2D memory array banks. Each memory banks are addressed by both a row and column address. To determine the memory selection it is necessary to first understand the process of reading from one of the 2D memory arrays. To read from the 2D memory array involves various steps and initially involves in selecting which row in the memory array to be address. This is done by issuing an ACTIVE command to the memory. This results in the memory array outputting an entire row of data through the sense amplifier, shown in figure 1. At this point the memory chip is ready to accept read commands. Read commands include a column address, which decodes and selects required piece of data, currently outputted by the sense amplifiers, to read. Once this process is completed the 2D memory array can be returned to an idle state and this is accomplished by issuing a PRECHARGE command to the memory. The need to both activate and pre-charged the 2D memory array is that both data's cannot be transmitted on every clock cycle, since the memory bank is busy in handling other tasks. In order to mask the time required to activate and pre-charge the memory array, DDR SDRAM memory chips is utilized with four independent banks of memory. The idea behind this is that bank is being activated or pre-charged, transactions can still occur on the remaining banks.

The DDR SDRAM is designed with RTL with VHDL coding its simulated outputs are as shown in the figure8.

7. POWER EVALUATION

To demonstrate the effectiveness of our power-conscious approach, we have analyzed the power consumption of base architecture configuration cache. Table I shows comparison of power usage between the two architectures. Selected kernels were executed with 100 iterations. Compared to the base architecture we have saved up to 93.85% of the total power consumed in the configuration cache and 39.19% of that in the entire architecture.

These results show that DPT a good solution for power saving in CGRA. In the case of 32 Taps FIR showing the maximum reductions ratio, the total power consumption of proposed architecture is much less than the result of PipeRench PipeRench has been fabricated in a 0.18-micron process and shows power measurement with varying FIR filter tap sizes. The power measurement shows that the power consumption of 32 Taps FIR ranges from 600 to 700mW.

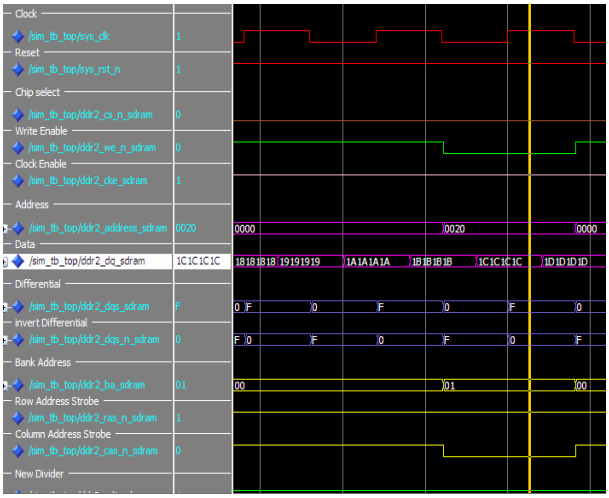


Figure 8: 128 Meg X 4 Functional simulation result

Existing result analyses are

Table I

SIZE OF CONFIGURATION CACHE AND CONTEXT REGISTERS

Size of memory elements	Architecture		Reduced (%)
	Base	Proposed	
Context registers	160-Byte	480-Byte	-
Spatial cache	5120-Byte	2560-Byte	40
Temporal cache		512-Byte	
Total amount	5280-Byte	3040-Byte	**32.73**

Table II

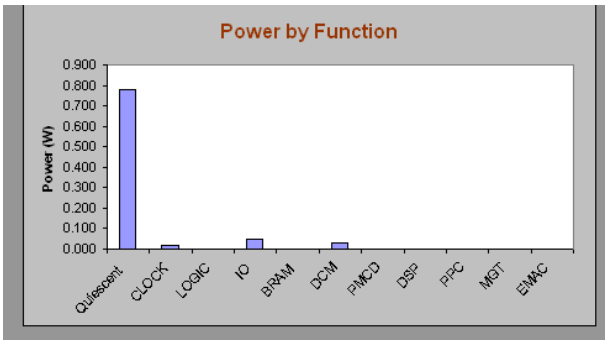
POWER COMPARISON BETWEEN BASE ARCHITECTURE AND PROPOSED ARCHITECTURE

Kernels	Power(mW)				Reduced(%)	
	Cache		Entire		Cache	Entire
	base	proposed	base	proposed		
First_Diff	104.77	17.13	360.63	241.96	83.65	32.90
Tri- Diagonal	106.18	19.25	340.35	240.28	81.87	29.40
Dot_Product	72.90	18.48	321.05	245.01	74.65	23.68
Complex_Mult	107.65	19.56	371.13	269.21	81.83	27.46
Hydro	89.34	19.53	295.20	211.19	78.14	28.46
ICCG	127.37	20.20	326.73	219.56	84.14	32.80
24-Taps FIR	142.23	19.44	330.15	207.36	**86.33**	**37.19**
MVM	120.09	18.17	309.54	207.62	84.86	32.92
ITRANS	127.41	43.50	246.88	162.97	65.85	33.98
2D-FDCT	115.02	22.60	269.55	177.84	80.22	34.02
SAD	113.94	46.16	237.86	170.08	59.48	28.49

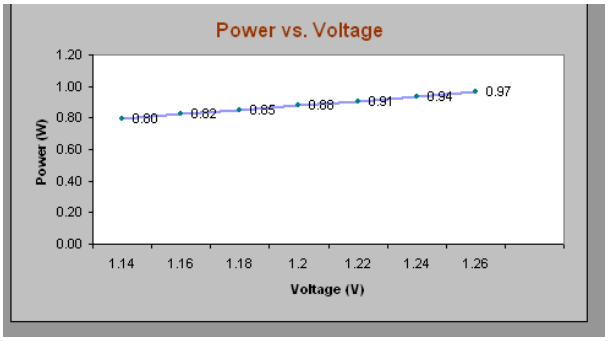
New 32 DDR taps FIR power reduced to 93.85%

Our proposed reconfiguration technique is used mean we can improve memory size as well as reduced power consumption are achieved. The testing results and graphs shows the experimental analysis.

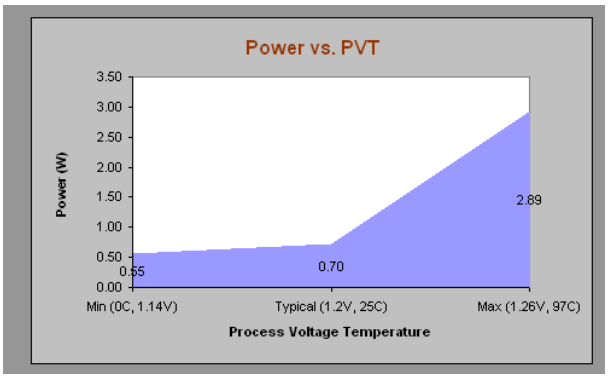
Graph I



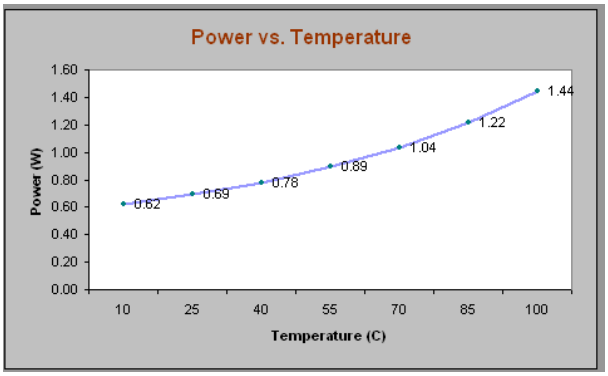
Graph II



Graph III



Graph IV



8. CONCLUSION

Coarse-grained reconfigurable architectures are considered appropriate for embedded systems -because they can satisfy both flexibility and high performance requirements. Though power consumption is crucial for the reconfigurable architecture to be used as a competitive processing core in embedded systems. Most reconfigurable architectures have a configuration cache for dynamic reconfiguration, which consumes high amount of power. In this paper, we proposed the design of the Double Pumping Technique (DPT) for low power reconfiguration cache structure supporting this technique. Our architecture can be used to achieve power-savings in a reconfigurable architecture while maintaining performance same as general CGRA. In addition, new configuration cache structure is more efficient than previous one in terms of memory size. In the experiments, we show that the proposed approach saves power even with reduced configuration cache size. Power reduction ratios in the configuration cache and the entire architecture are up to 93.85% show that definitely the overall architecture power consumption get reduce.

REFERENCES

[1] R. Hartenstein, “A decade of reconfigurable computing: A visionary retrospective,” in Proc. Des. Autom. Test Eur. Conf., Mar. 2001, pp. 642–649.

[2] B. Mei, S. Vernalde, D. Verkest, and R. Lauwereins, “Design methodology for a tightly coupled VLIW/reconfigurable matrix architecture:
A case study,” in Proc. Des. Autom. Test Eur. Conf., Mar. 2004, pp. 1224–1229.

[3] N. Bansal, S. Gupta, N. D. Dutt, and A. Nicolau, “Analysis of the performance of coarse-grain re-configurable architectures with different processing element configurations,” in Proc. Workshop Appl. Specific Process., Dec. 2003.

[4] A. Lambrechts, P. Raghavan, and M. Jayapala, “Energy-aware interconnect- exploration of coarse-grained re-configurable processors,” presented at the Workshop Appl. Specific Process., New York, Sep.
2005.

[5] H. Zhang, M. Wan, V. George, and J. Rabaey, “Interconnect architecture exploration for low-energy reconfigurable single-chip DSPs,” presented at the VLSI, Washington, DC, Apr. 1999.

[6] J. Lee, K. Choi, and N. D. Dutt, “Mapping loops on coarse-grained reconfigurable architectures using memory operation sharing,” Center for Embedded Computer Systems (CECS), Univ. California Irvine, Tech. Rep. 02-34, 2002.

[7] M. Ahn, J. W. Yoon, Y. Paek, Y. Kim, M. Kiemb, and K. Choi, “A spatial mapping algorithm for heterogeneous coarse-grained reconfigurable architectures,” in Proc. Des. Autom. Test Eur. Conf., Mar. 2006, pp. 363–368.

[8] H. Singh, M.-H. Lee, G. Lu, F. J. Kurdahi, N. Bagherzadeh, and E.M. C. Filho, "MorphoSys: An integrated reconfigurable system for data-parallel and computation-intensive applications," IEEE Trans. Comput., vol. 49, no. 5, pp. 465–481, May 2000.
[9] J. Becker and M. Vorbach, "Architecture, memory and interface technology integration of an industrial/academic configurable system-onchip (CSoC)," in Proc. IEEE Computer. Soc. Ann. Symp. VLSI, 2003, pp.107–112.

[10] Y. Kim, C. Park, S. Kang, H. Song, J. Jung, and K. Choi, "Design and evaluation of coarse-grained reconfigurable architecture," in Proc. Int.
SoC Des. Conf., Oct. 2004, pp. 227–230.

[11] Y. Kim, M. Kiemb, C. Park, J. Jung, and Choi, "Resource sharing and pipelining in coarse-grained reconfigurable architecture for domain-specific optimization," in Proc. Des. Autom. Test Eur. C nf., Mar.2005, pp. 19–24.

[12] M. Lanuzza, M. Margala, and P. Corsonello, "Cost-effective low-power processor-in-memory-based reconfigurable data path for multimedia applications," in Proc. Int. Symp. Low Power Electron. Des., Aug. 2005, pp. 161–166.

[13].Yoonjin Kim, "Low power Reconfiguration Technique for CGRC", IEEE Trans Vol.17 May 2009.