A Low Power Area Efficient Full Custom 3-Read 3-Write General Purpose Register in 65nm Technology

Youzhong Li, Lijun Zhang*, Qixiao Zhang, Ziou Wang, Lingfeng Mao
School of Urban Rail Transportation
Soochow University
Suzhou, P. R. China
zhanglijun@suda.edu.cn

Abstract—The design of a full custom 32×32 bit general purpose register (GPR) file with three read ports and three write ports for a microprocessor in SMIC 65nm Logic Low-Leakage CMOS technology was presented. During the design process, a full custom method with specific circuit construction including robust cell array and optimized decoder circuit was proposed in order to control power consumption and design area. For test mode only, an internal circuitry was built to resolve multiple write-port address collisions. With respect to conventional semi-custom solution, the full custom approach achieves 28% saving of power as well as 43% of area. Its operating frequency can reach up to 900MHz, the occupied area is 0.0311mm^2 and the average power dissipation comes to 5.83mW at 1.2V supply voltage which is also superior to some previous designs.

Keywords-GPR; low power; area efficient; robust cell; 65nm

I. INTRODUCTION

General purpose register file is an integral part of the microprocessor design and its performance is critical to the whole chip since it consumes a significant fraction of processor's power and total area. According to former study, GPR accounted for about 12% power dissipation of the total processor, in particular 41% of data supply energy in Stanford ELM processor [1]. And also, one additional work shows that the area of the register file is estimated to be over four times larger than the 64KB primary data cache [2]. Therefore, apart from considering operating speed, concerned about power consumption and area is an important issue.

To satisfy the requirements of low power dissipation and small area cost, a number of approaches have been reported previously. Full custom GPR architecture can reduce power and silicon area significantly compared to GPR file based on traditional standard library [3]. Since not all operands need the full-bit width of a register entry, a bit-partitioned register file was proposed to tackle the energy and area problems by M. Kondo et al. [4]. However, complicated construction was likely to increase the design cost, while reducing the reliability of circuit. Furthermore, a delayed write-back queue combined with an operand pre-fetch technique which comprise of an operand pre-fetch buffer and request queue was used to reduce port number of register file [5]. Although just the former technique is able to save both 40% power consumption and area occupation, they caused the delay increased as well as performance sacrificed.

In this paper, we present a 32-word×32-bit 3-read 3-write GPR file under full custom design methodology. As comparison, the GPR synthesized by conventional semi-custom solution have also been studied. A separate read and write word line 8T SRAM architecture was adopted as the GPR's cell array for both area saving and robustness. The read and write decoder circuit was re-optimized, which is capable of not only reducing power consumption, but decreasing the complexity of the peripheral circuits.

The organization of this paper is as follows. In Section I, a brief background and overview of the design is introduced. In Section II, the details of the GPR implementation include circuit and layout is proposed. Section III presents the comparison of full-custom and semi-custom design as well as previous studies. And Section IV obtains the conclusion of the whole work.

II. DESIGN IMPLEMENTATION

A. Implementation Overview

The designed 3-read and 3-write GPR with 32×32-bit memory cell array is based on SMIC 65nm LL process. Fig. 1 shows the block diagram of the GPR file. Expect for the 32×32 memory cell, it consists of read/write pre-decoder, read/write decoder, finite state machine, sense amplify module and write-through-read control circuit. All of the input data, address, enable, valid, and output data signals are latched for avoiding unpredictable variations when operations are performed. In addition, to resolve multiple write-port address collisions, an internal circuitry that deliberately set the write priority of each port was built for test mode.

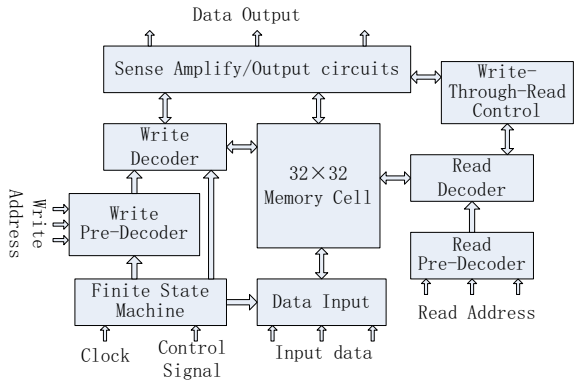

Figure 1. The GPR file architecture block diagram with 3-read and 3-write ports.

The function of the GPR contains three modes, which are read, write and write-through-read, depending on the values of the address, enable and valid signals. The address comparison for a write-through-read operation is performed

978-0-7695-5125-8/13 $26.00 © 2013 IEEE
DOI 10.1109/CSA.2013.178

inside the GPR. If the write-through-read is efficient, the write-port data inputs are sent directly to the read-port data outputs and written into the cell arrays in one clock cycle instead of performing a write followed by a read. Such an action would improve the working speed, save power dissipation obviously. When the read and write enable signals were valid, read address and write address was compared. As a result, if the addresses were equal, the write-to-read operation was performed.

B. Memory Cell Array

When carry out the GPR memory cell design, we pay close attention to area and robustness. The conventional 6T cell structure with two nMOSs pass-gate has fundamental stability problem that may cause data destruction in storage node when read operation is performed. Thus, it is not suitable for using as multi-port GPR file cell, especially when multi-port read performed in one cell. And although the cell introduced in [6] which added isolating inverters between coupled inverters and read bit line has superiority in stability, it consumes much precious area [6]. To overcome these disadvantages, a robust 8T cell [7] with separate read write word lines and different read write ports was used, as shown in Fig. 2. The single array cell contains totally 16 MOS transistors, MN4~MN10 are transmission transistors for writing operation of three write ports, and MN11~MN16 are separate read transistor which combine with three read bit lines.

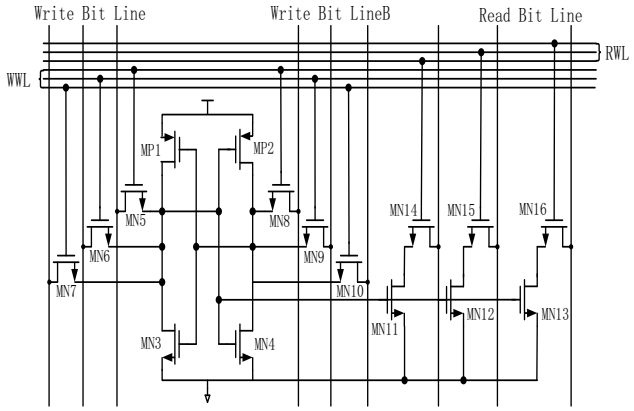

Figure 2. The Separate read and write 8T GRP cell array

In order to evaluate the stability of our 8T-SRAM based cell, the static noise margin (SNM) is simulated by Seevinck's method [8]. The results are listed in Table I, and the simulated results of traditional 6T cell are also given. It shows that the two cross-coupled inverters provided a significantly larger SNM. Without read disturbs, the robustness of 8T- SRAM based GPR cell is improved effective.

As we can see that the additional two read transistors in 8T GPR cell may lead to an increase in cell area over traditional 6T cell, yet greatly improvement of cell stability is obtained. The single cell area of out 8T cell, however, comes to 16.30um which is much smaller than the cell area (20.52um) in [6]. On the other hand, due to only those columns that need to be accessed will switch, the active power in out 8T cell array can be reduced. Furthermore, different from 6T cell array, there is no unnecessary read of unselected columns because those columns reside on a separate write word line.

TABLE I. COMPARED THE SNM OF TRADITIONAL 6T CELL AND OUR 8T CELL

Corner	Mode	6T SNM(mV)	8T SNM(mV)
TT	1 port-read	316	434
	2 port-read	242	434
	3 port-read	198	434
FF	1 port-read	358	485
	2 port-read	294	485
	3 port-read	231	485
SS	1 port-read	247	370
	2 port-read	178	370
	3 port-read	146	370

C. Read and Write decoder

For input read and write addresses with the addition of control signals, a pre-decode operation is needed. By this way, the faster working speed can be achieved. The read and write pre-decoder includes a latch, a 2-to-4 decoder and a 3-to-8 decoder of each port so as to generate the 12 bit temporary address. As mentioned before, an internal circuitry which used to resolve multiple write-port address collisions was built into the write address decoder part. Fig. 3 depicts the structure of write decoder integrated with address collisions resolution circuit where ACL, BCL and CCL are address input control signals, and WENWL is write enable signal. When there is more than one port writes data to the same address of GPR array, the port priority is defined. In this work, we set the A port with the highest priority, B followed and C the lowest as shown in Fig. 3. With this approach, we combined the write decoder and address collisions resolution circuit together ingeniously, so that the write address conflicts would be resolved without any additional module. It is helpful to reduce power consumption of the whole GPR and save silicon area.

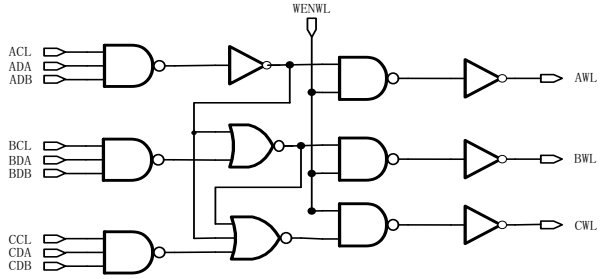

Figure 3. The structured of optimized write address decoder

During read operation, a comparison between read address and write address is carried out when read enable signal is valid. This determines whether the output data is taken from the cell array or input port directly or keep unchanged. Considering these, the read address decoder is more complicated with respect to write one. Finally, we adopted one three-input NOR gate and two two-input NAND gate to achieve lower power consumption and smaller occupied area.

D. Layout

The layout of the GPR was accomplished by full-custom method as well, as shown in Fig. 4. An advantage of the full-custom layout is the opportunity to optimize the cell array and placement of peripheral circuit for smaller silicon area. From Fig. 4, we can see that the 32×32 cell array is in the centre and occupies about 53% of the whole chip. The data-in and data-out module is symmetrical in bottom and top of side in order to avoid crowd. The read decoder and write decoder is symmetrically placed on both side of cell array for the purpose of shortening the signal path. With regards to the cell, a compact layout was achieved, which run the read and write word lines horizontally across the cell width like the word line in 6T cells.

III. Analysis and Simulation Results

The GPR file in this work would be implemented in SMIC 65nm 1P6M Low-Leakage (LL) process. Fig. 5 (a) and (b) give the frequency and power consumption of the GPR at different voltage supply and temperature respectively. From Fig. 5 (a), we could find that as the voltage supply increases, the operating frequency shows an upward trend. Moreover, when the voltage is the same, the better performance is achieved at low temperature. Checking Fig. 5 (b), it is clearly seen that the lower supply voltage, the smaller energy dissipation. At the optimum operating voltage of 1.2V, the GPR can work well at 917MHz with 5.52mW power consumption.

As mentioned before, the primary full-custom design results along with traditional synthesized one are given in Table II. Compared with the synthesized design, the full-custom one obtains a dynamic power decrease of 28% as well as considerable leakage power decline due to well-designed circuit. Furthermore, significant silicon area reduction up to 43% was realized by optimizing the layout and placement relative to synthesized design.

To make the comparison more comprehensive and explicit, the normalized power was defined as following:

$$P_{norm} = \frac{Power[nW]}{Freq[MHz] \bullet \left(n_{read-ports} + n_{write-ports}\right) \bullet n_{bits}} \quad (1)$$

Moreover, the power per port per frequency is obtained according to equation (2)

$$P_{pppf} = \frac{Power[\mu W]}{Freq[MHz] \bullet \left(n_{read-ports} + n_{write-ports}\right)} \quad (2)$$

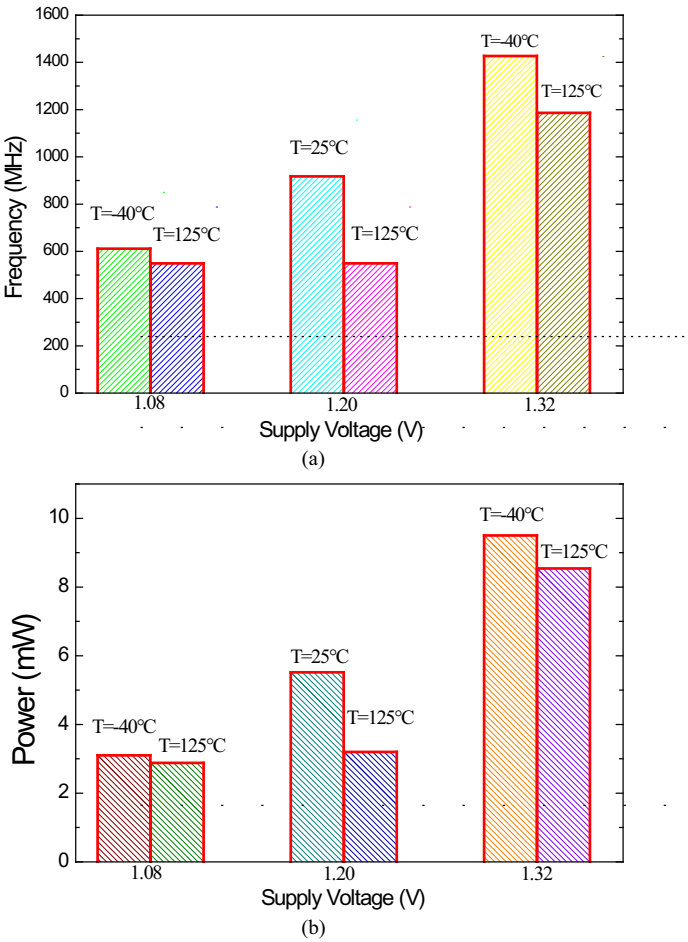

Figure 5. Simulation results at each voltage supply and temperature. (a) Frequency. (b) Power consumption

TABLE II. The Comparison of Full-Custom (FC) and Synthesized (SYN) Design Results

Parameter	FC	SYN	Unit	Ratio[%]
Total Area	0.0311	0.0548	mm^2	43
Dynamic Power	5.52	7.65	mW	28
Worest Power	9.50	12.17	mW	22
Leakage Power	18.59	25.33	μW	27

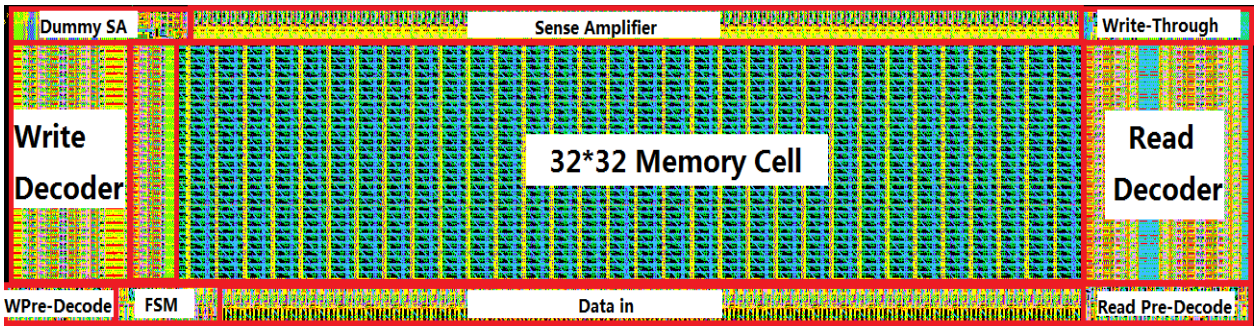

Figure 4. The layout of the GPR file

TABLE III. COMPARISON OF THE PRESENT GPR WITH SOME PREVIOUSLY PUBLISHED WORKS

Reference	$N_{words} \times N_{bits}$	Read Ports	Write Ports	Process [nm]	Supply [V]	Frequency [MHz]	Power [mW]	P_{pppf} [μW/MHz]	P_{norm} [nW/MHz]	Area [mm²]	A_{norm} [μm²]
ISSCC' 06[9]	16×64	1	1	65	1.2	8800	198	11.25	175.8	0.017	8.30
ESSCC' 07[10]	48×42	1	1	65	1.2	6300	47	3.73	88.81	0.031	4.79
ASYNC' 09[11]	32×32	2	2	65	1.0	1200	7.5	1.56	48.83	0.0088	2.15
VLSI' 09[12]	30×32	3	2	65	1.2	200	1.2	1.20	37.50	0.023	4.79
ISSCC' 10[13]	64×32	1	1	32	1.0	7500	72	4.80	150.0	0.076	18.55
ISSCC' 11[14]	144×78	4	2	45	0.9	2300	59	4.28	54.81	0.088	1.31
JOS' 12[6]	32×32	4	2	65	1.2	800	7.2	1.50	46.88	0.046	7.49
This Work	32×32	3	3	65	1.2	900	5.52	1.02	31.94	0.031	5.05

Similarly, the normalized area was given by (3)

$$A_{norm} = \frac{Area[\mu m^2]}{n_{words} \bullet n_{bits} \bullet \left(n_{read-ports} + n_{write-ports} \right)} \quad (3)$$

All the parameters of our design including P_{norm}, P_{pppf} and A_{norm} are listed in table III. Meanwhile some previously published works are also presented as comparison. The results show that the power consumption is reduced significantly except for [12]. Especially the normalized power reaches 31.94nW/ bit·MHz which is the lowest one. Even through the silicon area is not the smallest, it still has remarkable advantage over other recently published register files.

IV. CONCLUSION

This paper details the full-custom design of a 32×32 bit 3-read 3-write general purpose register file. The implementation adopts 8T SRAM cell with separate read and write word and bit line as the memory cell. This approach not only ensures the stability, also reduces the occupied area. The write decoder with address collisions resolution circuit and the read decoder integrate with address comparison circuit are elaborately design to decline both power consumption and silicon area respectively. Besides, the full-custom layout realized the optimization of cell array as well as external circuit placement for area efficient. The GPR is implemented in SMIC 65nm technology, and the simulation results demonstrate that both power dissipation and area decreased significantly with respect to synthesized consequence. Further, compared with some published works, superiority in power and area is obvious as well.

ACKNOWLEDGMENT

The authors wish to acknowledge the assistance and support from the National Natural Science Foundation of China, under Grant No.61272105, and 61076102.

REFERENCE

[1] J. D. William, B. James, B.-S. David, C. James, R. H. Curtis , P. Vishal et al., "Efficient Embedded Computing," IEEE Computer, vol. 41, pp. 27-32, July 2008.

[2] P. R. Preston, R. W. Badeau, D. W. Bailey, S. L. Bell, L. L. Biro, W. J. Bowhil et al. "Design of an 8 wide superscalar RISC microprocessor with simultaneous multithreading," IEEE Int. Solid-State Circuits Conference, pp. 334-472, February 2002.

[3] T.-S. Jau, W.-B. Yang, and C.-Y. Chang, "Analysis and design of high performance, low power multiple ports register files," IEEE Proc. Asia Pacific Conference on Circuits and Systems APCCAS 2006, pp. 1453–1456, December 2006.

[4] M. Kondo, and H. Nakamura, "A Small, Fast and Low-Power Register File by Bit-Partitioning," High-Performance Computer Architecture, 2005. HPCA-11. 11th Int. Sym. on, pp. 40-49, February 2005.

[5] N. S. Kim, and T. Mudge, "The microarchitecture of a low power register file," Low Power Electronics and Design, 2003. ISLPED '03. Proceedings of the 2003 Int. Sym. on, pp. 384-389, August, 2003.

[6] X. X. Zhang, Y. Li, B. Y. Xiong, Y. Y. Zhang, F. Y. Dong, and Z. Zhang, et al. "Robust and low power register file in 65 nm technology," Journal of Semiconductor, vol. 33, pp. 035010-1-5, March 2012.

[7] L. Chang, D. M. Fried, J. Hergenrother, J. W. Sleight, H. D. Rober, K. M. Robert, and et al. "Stable SRAM Cell Design for the 32 nm Node and Beyond," VLSI Technology, 2005. Digest of Technical Papers. 2005 Symposium on, pp. 128-129, June 2005.

[8] E. Seevinck, F. J. List, and J. Lohstroh, "Static-noise margin analysis of MOS SRAM cells," IEEE J. Solid-State Circuits, vol. 22, pp. 748-754, October 1987.

[9] S. Hsu, A. Agarwal, M. Anders, S. Mathew, R. Krishnamurthy, and S. Borkar, "An 8.8 GHz 198mW 16x64b 1R/1W variationtolerant register file in 65nm CMOS," IEEE Int. Solid-State Circuit Conference, pp. 1785-1797, February 2006.

[10] A. Agarwal, N. Banerjee, K. S. Hsu, R. K. Krishnamurthy, and R. Kaushik, "A 200mV to 1.2V, 4.4MHz to 6.3GHz, 48x42b 1R/1W Programmable Register File in 65nm CMOS," European Solid State Circuits Conference, pp. 316-319, September 2007.

[11] J. Dama and A. Lines, "GHz asynchronous SRAM in 65nm," in Proc. 15th IEEE Symposium on Asynchronous Circuits and Systems ASYNC'09, pp. 85–94, May 2009.

[12] H. O. Kim, B. H. Lee, J. T. Kim, J. Y. Choi, K. M. Choi, and Y. Shin, "Supply switching with ground collapse for low-leakage register files in 65-nm CMOS," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 18, pp. 505-509, July 2009.

[13] A. Agarwal, S. K. Mathew, S. K. Hsu, M. A. Anders, H. Kaul, and F. Sheikh, et al. "A 320mV-to-1.2V on-die fine-grained reconfigurable fabric for DSP/media accelerators in 32 nm CMOS," IEEE Int. Solid-State Circuit Conference, pp. 328-330, February 2010.

[14] G. S. Ditlow, R. K. Montoye, S. N. Storino, S. M. Dance, S. Ehrenreich, and B. M. Fleischer et al. "A 4R2W register file for a 2.3 GHz wire-speed POWER processor with double-pumped write operation," IEEE Int. Solid State Circuits Conference, pp. 256-258, February 2011.