

A new FPGA architecture suitable for DSP applications*

Wang Liyun(王丽云)[†], Lai Jinmei(来金梅), Tong Jiarong(童家榕), Tang Pushan(唐璞山),
Chen Xing(陈星), Duan Xueyan(段雪岩), Chen Liguang(陈利光), Wang Jian(王健),
and Wang Yuan(王元)

ASIC and System State Key Laboratory, Fudan University, Shanghai 201203, China

Abstract: A new FPGA architecture suitable for digital signal processing applications is presented. DSP modules can be inserted into FPGA conveniently with the proposed architecture, which is much faster when used in the field of digital signal processing compared with traditional FPGAs. An advanced 2-level MUX (multiplexer) is also proposed. With the added SLEEP MODE PASS to traditional 2-level MUX, static leakage is reduced. Furthermore, buffers are inserted at early returns of long lines. With this kind of buffer, the delay of the long line is improved by 9.8% while the area increases by 4.37%. The layout of this architecture has been taped out in standard 0.13 μm CMOS technology successfully. The die size is $6.3 \times 4.5 \text{ mm}^2$ with the QFP208 package. Test results show that performances of presented classical DSP cases are improved by 28.6%–302% compared with traditional FPGAs.

Key words: FPGA; uniform GRM structure; DSP modules

DOI: 10.1088/1674-4926/32/5/055012

EEACC: 2570

1. Introduction

Owing to the lower non-recurring engineering (NRE) costs and quick commercialization, FPGAs are playing a more important role in modern circuit design. In traditional FPGAs, CLBs (configurable logic blocks) are the most popular units to realize logic functions^[1–3]. However, due to their long delay time, CLBs cannot satisfy the needs of high speed digital signal processing. It is a trend to insert DSP modules into FPGAs for the purpose of improving FPGA performance. Although some commercial FPGA vendors have invented FPGAs with DSP modules, however, there is little in the literature about how to achieve this. In this paper, a new FPGA architecture that facilitates DSP insertion is proposed. Logic function can be realized by CLBs and IPs together, so circuit speeds can be accelerated.

Since multipliers are most often used in digital signal processing modules (like FFT, FIR and IIR) and contribute much to key path delay, they are selected as the DSP module used in the proposed architecture.

To implement this architecture successfully, a uniform and repeatable routing resource is designed. With this routing resource, other IP modules besides multipliers can be inserted into FPGAs successfully. The entire architecture and routing resources are described below.

2. The entire architecture design

The design, named FDP (FuDan Programmable logic chip) 2009-II, is composed of 16×8 TILES, 192 programmable IOBs, eight 18k bit block RAMs (BRAMs), eight 18×18 multipliers, configuration circuits and programmable routing resources (shown in Fig. 1). All of this will be described in detail.

2.1. TILE design

The TILE is composed of a CLB and its associated routing resources (as shown in Fig. 2).

The proposed CLB consists of four slices that are proposed by Ref. [19]. The slice is mainly composed of 4 input LUTs, carry chain cells, AND logics (for 1 bit multiplication), XOR logics (for adder), MUXs and registers. With all of the cells, combinational and sequential logic could be realized expediently.

2.2. Multiplier design

Eight 18×18 bit signed multipliers based on the Baugh–Wooley algorithm^[4] are presented. They are high speed

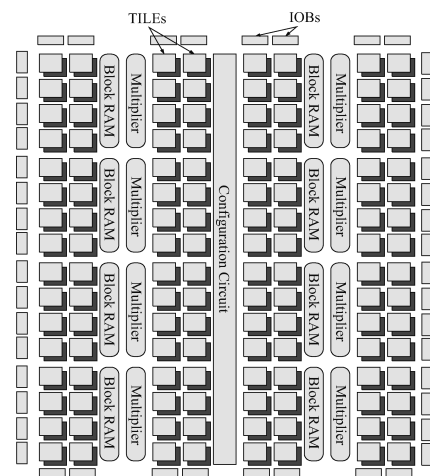


Fig. 1. Entire architecture.

* Project supported by the National Natural Science Foundation of China (No. 60776023), the National High Technology Research and Development Program of China (No. 2007AA01Z285), and the Plan of Shanghai Science and Technology, China (No. 08706200101).

[†] Corresponding author. Email: 071021037@fudan.edu.cn

Received 28 October 2010, revised manuscript received 14 December 2010

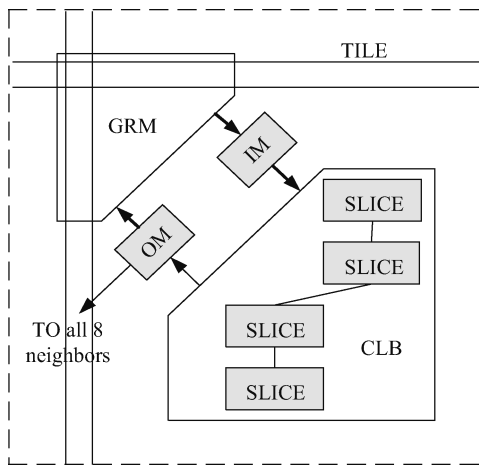


Fig. 2. A TILE.

Table 1. Simulation results.

Delay	Transistor level simulation	Post-layout simulation
FF (ns)	3.14	5.11
TT (ns)	3.80	6.17
SS (ns)	4.78	7.78
Tool	Hspice	Hspice
Library	SMIC 1013_v2p6.lib	SMIC 1013_v2p6.lib

complement multipliers that focus on bit level manipulation. A 4×4 bit Baugh–Wooley multiplier model is depicted in Fig. 3, which is used to describe the proposed structure of a 18×18 bit multiplier.

Since the multiplier is the key component in this design, its delay is also carefully simulated besides logic function verification. Both transistor-level and post-layout simulations are carried out by Hspice in three technology corners: FF, TT, and SS. Simulation results are shown in Table 1.

2.3. Block RAM design

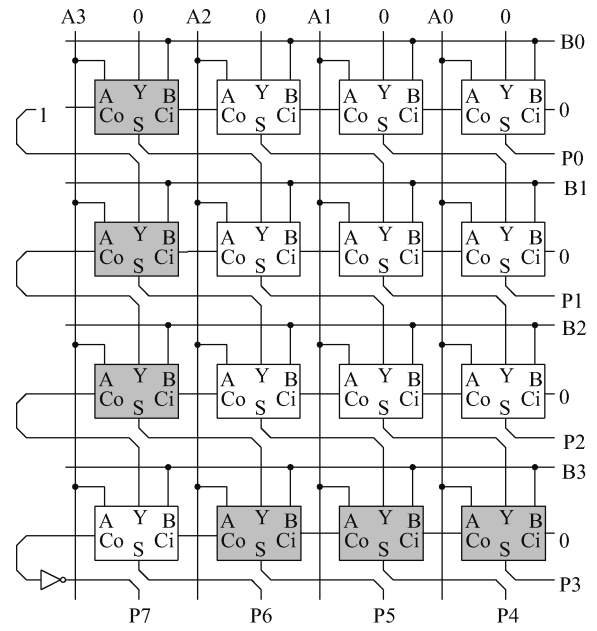
Eight 18k bit embedded dual-port Block RAMs are proposed, as shown in Fig. 4. With the bit-width adjust technique, each 18k bit BRAM can be configured as six kinds of RAM: $16k \times 1$ bit, $8k \times 2$ bits, $4k \times 4$ bits, $2k \times 9$ bits, $1k \times 18$ bits or 512×36 bits. The proposed block RAMs can work in three modes: NO CHANGE, READ FIRST and WRITE FIRST.

2.4. Programmable IOB and configuration circuit design

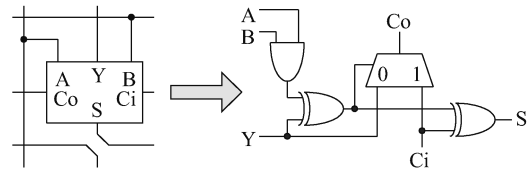
Programmable IOB is the interface between outside chip pins and internal configurable modules. In this chip, each IOB can be configured as an input block, an output block or a bidirectional block. LVCMOS and LVTTTL standards are both supported.

The configuration block realizes three functions: chip resetting with some delay after power on; initializing and downloading user bit streams; FPGA starting and controlling of read back bits from the FPGA after configuration is completed.

For conventional FPGAs, the minimum bit-stream unit to be downloaded is one frame^[5,6]. The frame length is hundreds



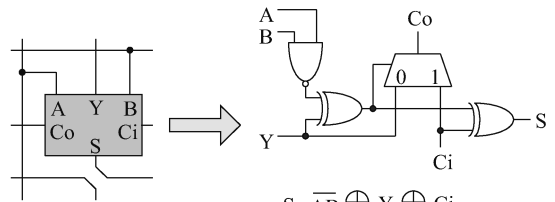
(a)



$$S = AB \oplus Y \oplus Ci$$

$$Co = AB * Y * Ci$$

(b)



$$S = \overline{AB} \oplus Y \oplus Ci$$

$$Co = \overline{AB} * Y * Ci$$

(c)

Fig. 3. A 4×4 bit Baugh–Wooley multiplier.

to thousands of bits for different sizes of FPGA chips. In this design, a new configuration circuit is proposed. This circuit features an addressable configuration register and an internal frame decoder that makes a 32-bit memory cell of the FPGA addressable. So the minimum bit-stream is decreased to 32 bits. That provides a faster configuration speed and more flexible partial configuration operations^[7].

2.5. Uniform and repeatable programmable routing resource

The proposed routing resources include global routing, terminal routing and routing resources for different modules. Terminal routing is used for dealing with end points of long lines and hanged lines. Hanged lines are kinds of multi-length lines that touch the chip border before finishing their length. The

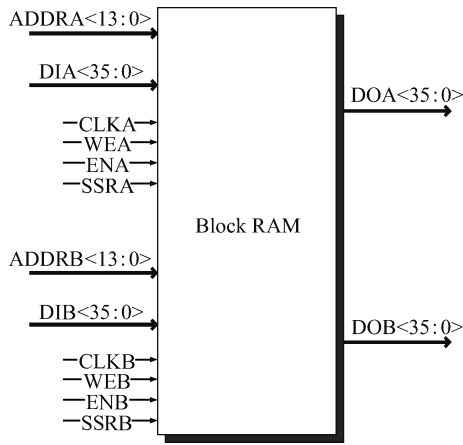


Fig. 4. An 18 kbit embedded dual-port block RAM.

global routing resource includes a unidirectional length-2 line, a unidirectional length-6 line and a bidirectional long line. The last part is the key point in routing design.

To insert a multiplier into the FPGA, the routing resource for the modules (CLB, IOB, Multiplier and BRAM) is divided into 3 parts: global routing multiplexers (GRM), input multipliers (IM) and output multipliers (OM), as shown in Fig. 5. The global routing resource includes lines in four directions: north, west, south and east. If a signal needs to be transmitted in a different direction, it could be done through the GRM. The GRM is also responsible for communication between the CLB or multiplier and the global routing resource. The IM and OM are composed of sets of multiplexers. The inputs of the IM come from the GRM and the outputs of the IM are connected to the inputs of the CLB or multiplier. The outputs of the CLB or multiplier are connected to the inputs of the OM, and the OM feeds them to the GRM or IMs of all 8 neighbors.

The GRM is the same for the CLB, IOB, BRAM and multiplier, while the IM and OM are not. There are special IM and OM for different modules, considering the number of IOs for them is different. As long as the GRM is the same, if the IM and OM are brought with a multiplier at the same time, the CLBs in the array can be replaced, so the multipliers can be inserted anywhere that is required.

2.5.1. Structure of uniform and repeatable GRM

The connection relationship in the GRM is shown in Fig. 6. In the GRM, there are length-2, length-6 and long lines from the global routing resource. The circuits are mainly composed of MUXs and buffers. A signal transmits from one line to another through the MUX. There are 4 kinds of MUX: 16 to 1 MUX for length-2 line, 12 to 1 MUX for length-6 line, 4 to 1 MUX for long line horizontally and 10 to 1 MUX for long line vertically. To get the best delay optimization, the driving relationship in the GRM is described as follows. Long lines can drive length-6 lines; length-6 lines can drive length-2 lines; length-6 and length-2 lines can be the input of the IM. The output pins of the OM can drive all three kinds of lines described above. The output of the module can drive length-2 and length-6 lines. Length-2 and length-6 lines can be driven by the same lines from the other direction.

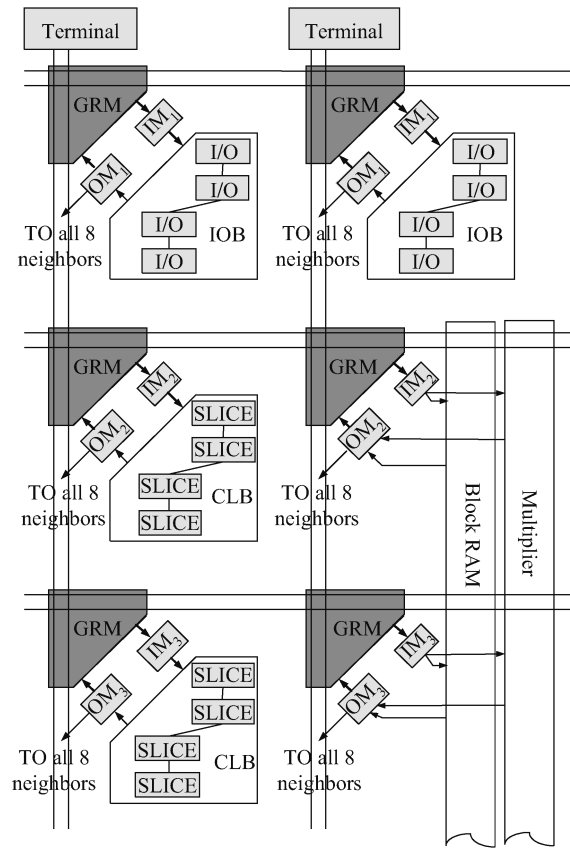


Fig. 5. Partial uniform and repeatable routing resource.

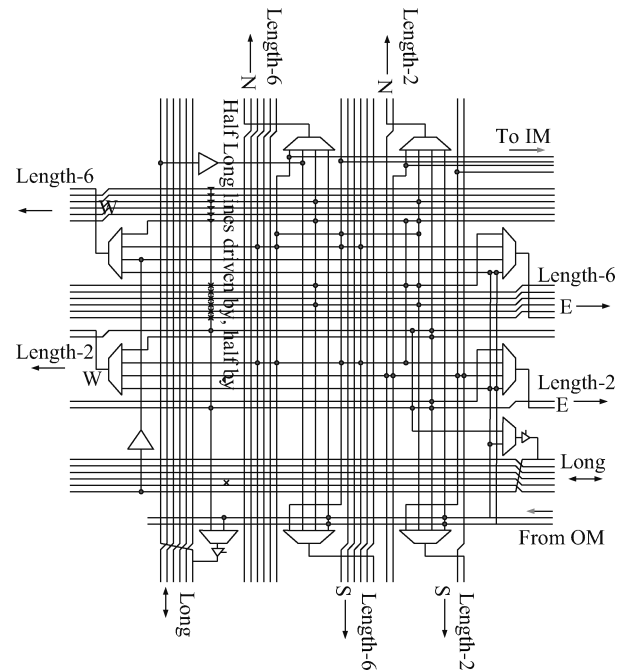


Fig. 6. Connection relationship in the GRM structure.

2.5.2. Terminal routing design

When a multi-length line touches the border of the chip, the length is not finished and the end point appears. To solve this problem, we adopt the return line, which guarantees the re-

Table 2. Comparison of MUX design styles.

Property	Style		
	Tree	Flat	2-level
Delay level	$\lceil \log_2 N \rceil$	1	2
capacitance	$\lceil \log_2 N \rceil$ (distributed)	$N - 1$ (lumped)	$2\sqrt{N}$
Configuration bits	$\lceil \log_2 N \rceil$	N	$2\sqrt{N}$
Pass gates	$2N - 2$	N	$N + \sqrt{N}$

N is the fan-in of the multiplexer.

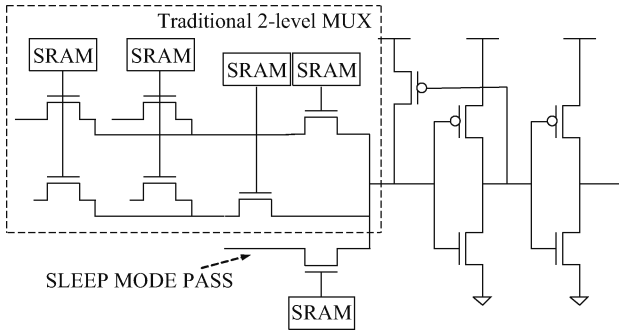


Fig. 7. Structure of the advanced 2-level MUX.

peatability of the single module routing resource and the uniform loading on the same kind of line. In terminal routing, long lines still drive length-6 lines, just as the connection relationship of the GRM^[8].

2.5.3. Advanced 2-level MUX

The basic unit in the GRM, IM and OM is the MUX + buffer structure. The MUX is the most popular unit used in FPGA routing design^[9-16], which ensures the programmability of the routing resource.

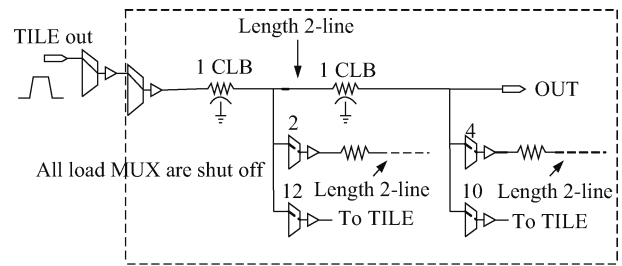
A 2-level MUX^[12, 15] is used popularly in a modern routing switch design owing to better balance between the area and delay than the other 2 MUXs (shown in Table 2). However, a defect exists. When it is idle, all of the SRAM bits are configured as logic “0”, and there is uncertain voltage at the buffer input (as shown in the square of Fig. 7), which causes static leakage. To solve this problem, an advanced 2-level MUX is proposed. With the added SLEEP MODE PASS (shown in Fig. 7), logic “1” goes to the buffer input and static leakage is reduced.

2.5.4. Long lines with buffers at early returns

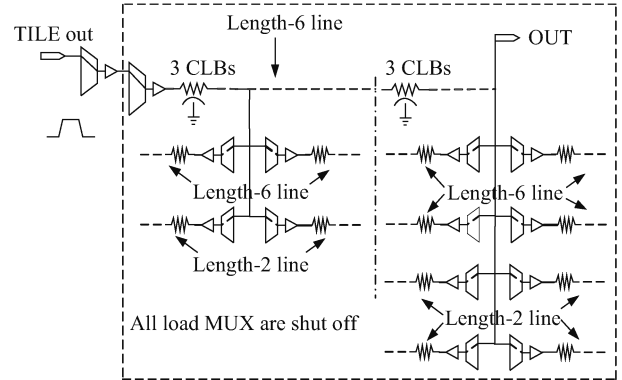
Long lines span the entire chip horizontally or vertically with early returns connecting to a CLB every 6 CLBs. Each early return is fed to some MUXs for a length-6 line. There is more loading along the line as signals travel further in traditional long line design. So the delay is too much to bear. To solve this problem, we inserted an extra buffer at the early turn, as shown in Fig. 8(c). The simulation result shows that with this buffer, the delay of the long line is improved by 9.8% while the area is increased by 4.37%.

2.5.5. Routing resource simulation

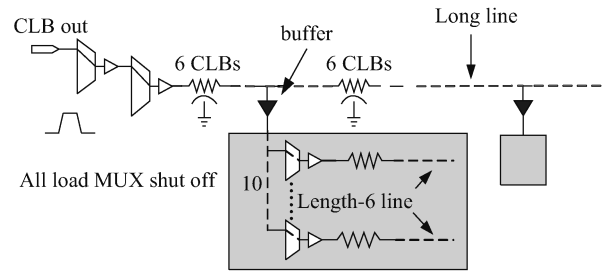
Since routing performance is very important in entire chip performance, performances of length-2, length-6 and long line



(a)



(b)



(c)

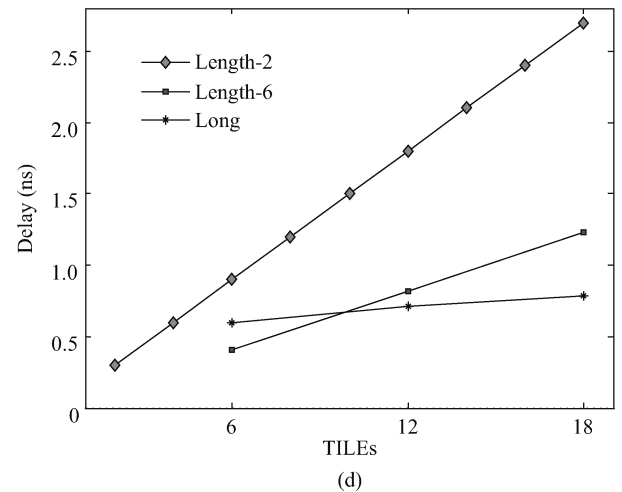


Fig. 8. Simulation model for (a) length-2 line, (b) length-6 line, and (c) long line. (d) Performance of each kind of line.

are carefully simulated before the chip is taped out.

- simulation software: Spectre;
- simulation library: SMIC 1013_v2p6_spe.lib;
- technology node: TT, FF, SS, FS, SF.

The transistor size of the multiplexer and buffer are scanned to get the best performance. The simulation model for

each kind of line and the final delay value used in the chip at TT are given in Fig. 8. From this figure, we can see that the performance of length-2 and length-6 lines is linear, while that of the long line is not. That's because for the long line there are no re-drive buffers along the line. The simulation model for each line of every programmable model (CLB, IOB, Multiplier and Block RAM) is the same.

3. Implementation and test result

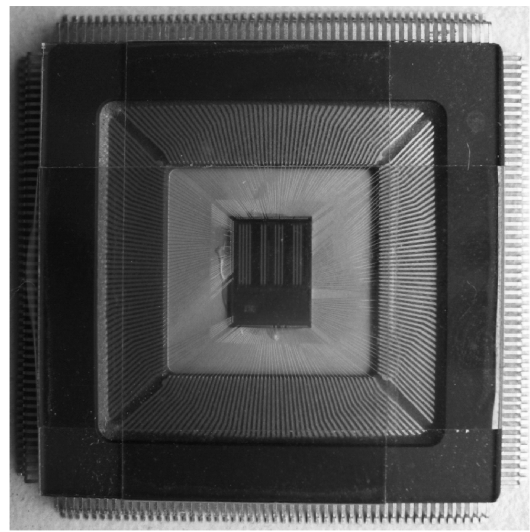
The array of logic TILES in the FDP2009-II chip is 16×8 , with BRAMs and multipliers inserted between them. We have simulated each module (CLB, IOB, BRAM and multiplier) before the layout is taped out. As for the layout, modules like CLBs, routing resources, IOBs, BRAMs and multipliers are designed by a full custom method. As for the configuration circuit, firstly the circuit is described in Verilog HDL language, and then it is synthesized by the Synopsys design compiler tool. Finally, the layout is generated by the Synopsys Astro tool. FDP2009-II chip is designed and manufactured with $0.13 \mu\text{m}$ standard CMOS technology, and the die size is $4.5 \times 6.3 \text{ mm}^2$, in which about 60% is taken up by routing resources.

To evaluate the chip FDP2009-II, a test platform (Fig. 9(b)) is built up, which includes an FDP2009-II FPGA test board, a dual channel arbitrary function generator (Tektronix AFG3235), mixed signal oscilloscopes (Tektronix MSO 4054), a multimeter and a PC. Corresponding software, FDE2009-II (advanced version of Ref. [18]), is developed to test the FPGA chip.

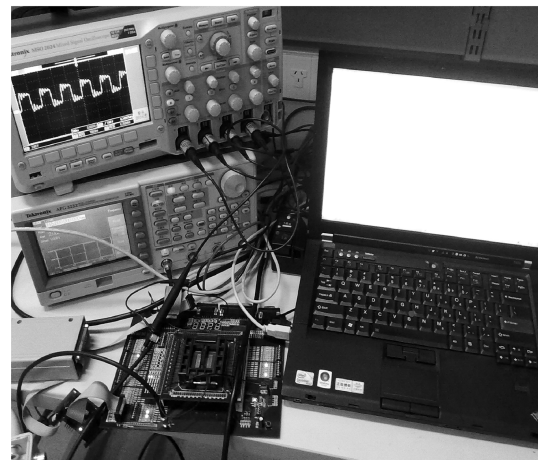
Both function and performance tests are considered. For the function test, firstly we test every single unit function in the CLBs, like LUT, Distributed RAM, shift register, fast carry chain, arithmetic unit and DFF. Then the routing resource, multiplier and BRAMs are tested. All of the test cases demonstrate that the FPGA core works correctly.

As for the performance test, we considered an 18 bit multiplier and five classical DSP benchmarks: Rotator64 (a part of FFT), IIR_4_16 (an IIR filter with a filter order of 4 and input Precision of 16), IIR_4_8 (an IIR filter with a filter order of 4 and input Precision of 8), IIR_2_16 (an IIR filter with a filter order of 2 and input Precision of 16) and FIR_7_16 (an FIR filter with a filter order of 7 and input Precision of 16). In order to evaluate the chip performance, these test cases were implemented in an FDP2009-II. These cases were also implemented in an FDP 2008^[3], a previous generation FPGA designed by Fudan University. As the Altera Stratix and Agate Angelo are the same as the FDP2009-II in terms of the technology process, their area and performance are listed in Table 3 for comparison. The Angelo is designed by Agate Logic Corporation, the first corporation to design FPGAs in China.

From these results it can be seen that the performances of the presented classical DSP cases are improved by 28.6%–302% compared with the traditional FPGA-FDP2008^[3] and 2.1X–29.3X better than Angelo. The performance of the benchmarks is better in the FDP2009-II than in the traditional FPGAs because all of these DSP benchmarks use the multiplier as the fundamental building block. We can see that two cases failed when they were implemented in the Angelo. That is because the capacity of Angelo is 4096 LUTs + FFs (Angelo's capacity is the biggest among the Agate prod-



(a)



(b)

Fig. 9. (a) The FDP2009-II FPGA chip. (b) Test environment of the FDP2009-II chip.

ucts) and its software shows that the resource is not enough. As we can see, the performance of the first two benchmarks is faster in the FDP2009-II than in the Altera Stratix, while the performance of the last four benchmarks is slower. Also, the first two benchmarks need fewer resources than the last four. The performance of the last four benchmarks is slower in the FDP2009-II than in the Stratix because the resources in the FDP2009-II (including 1024 LUTs and 8 multipliers) are much less than in the Stratix (Stratix EP1S60B956C6 includes 57120LEs and 144 DSPs) and when resources are insufficient, there are fewer choices for placement and a routing tool to choose to get the best result.

4. Conclusion

A new FPGA architecture (FDP2009-II) suitable for digital signal processing applications is proposed. It is composed of 16×8 TILES, 192 programmable IOBs, eight 18k bit BRAMs, eight 18×18 multipliers, configuration circuits, and uniform and programmable routing resources. With the proposed routing resources, DSP modules can be inserted into the FPGA con-

Table 3. Test results.

Benchmarks	Altera-Stratix EP1S60B956C6 0.13 μm		Agate-Angelo 0.13 μm		FDP 2008 ^[3] 0.18 μm		FDP 2009-II 0.13 μm	
	Area	Speed (MHz)	Area	Speed (MHz)	Area	Speed (MHz)	Area	Speed (MHz)
Mult_18	72 LEs	225	648 LUTs 72 FFs	11	217 LUTs 72 FFs	83	72 FFs	334 1 Multiplier
Rotator64	148 LEs 8 DSPs	120	2183 LUTs 147 FFs	12	1203 LUTs 301 FFs	55	40 LUTs 105 FFs	147 4 Multipliers
IIR_4_16	611 LEs 2 DSPs	67	Failed	—	2219 LUTs 152 FFs	33	282 LUTs 128 FFs	50 8 Multipliers
IIR_4_8	323 LEs 2 DSPs	78	950 LUTs 48 FFs	21	602 LUTs 66 FFs	50	138 LUTs 64 FFs	68 8 Multipliers
IIR_2_16	271 LEs 2 DSPs	87	1678 LUTs 32 FFs	13	1092 LUTs 79 FFs	58	100 LUTs 64 FFs	75 4 Multipliers
FIR_7_16	362 LEs 14 DSPs	91	Failed	—	1976 LUTs 604 FFs	63	236 LUTs 619 FFs	81 8 Multipliers

veniently. Since multipliers are one of the most used modules in signal processing and contributes a lot to key path delay, they are selected as the DSP modules used in the proposed architecture. Compared with a pure CLB structure, the performance of this architecture is improved. The design of this architecture is taped out successfully. Test results show that the performances of the presented classical DSP cases are improved by 28.6%–302% compared with a traditional FPGA^[3].

References

- [1] Betz V, Rose J, Marquardt A. Architecture and CAD for deep-submicron FPGAs. Kluwer Academic, 1999
- [2] Meyer J, Kocan F. Sharing of SRAM tables among NPN-equivalent LUTs in SRAM-based FPGAs. IEEE Trans Very Large Scale Integration Systems, 2007, 15(2): 182
- [3] Wu Fang, Wang Yabin, Lai Jinmei. Circuit design of a novel FPGA chip FDP2008. Journal of Semiconductors, 2009, 30(11): 115009
- [4] Baugh C R, Wooley B A. A Tow's complement parallel array multiplication algorithm. IEEE Trans Computers, 1973, C-22(12): 1045
- [5] Schultz D P, Hung L C. Method and structure of configuring FPGAs. USA Patent, No. 6204687B1, Mar 20, 2001
- [6] Xilinx Corporation. Application Note Xapp151. Virtex series configuration architecture user guide (v1.7), Oct 20, 2004
- [7] Xie J, Wang Y, Chen L, et al. Fast configuration architecture of FPGA suitable for bitstream compression. Proc IEEE, the 8th Int Conf on ASIC, 2009: 126
- [8] Wang L, Wang Y, Chen L, et al. Uniform routing architecture for FPGA with embedded IP cores. Proc IEEE, the 8th Int Conf on ASIC, 2009: 109
- [9] Meijer M, Krishnan R, Bennebroek M. Energy-efficient FPGA interconnect design. Proc Design, Automation and Test in Europe, 2006
- [10] Anderson J H, Najm F N. Low-power programmable routing circuitry for FPGAs. IEEE/ACM International Conference on Computer Aided Design, 2004
- [11] Rahman A, Das S, Tuan T, et al. Heterogeneous routing architecture for low-power FPGA fabric. Proc IEEE Custom Integrated Circuits Conference, 2005
- [12] Lewis D, Ahmed E, Baeckler G, et al. The Stratix II logic and routing architecture. Proceedings of the ACM/SIGDA 13th International Symposium on Field-Programmable Gate Arrays, 2005
- [13] Liu H, Chen X, Ha Y. An architecture and timing-driven routing algorithm for area-efficient FPGAs with time-multiplexed interconnects. International Conference on Field Programmable Logic and Applications, 2008
- [14] Anderson J H, Najm F N. Low-power programmable routing circuitry for FPGAs. IEEE/ACM International Conference on Computer Aided Design, 2004
- [15] Lee E, Lemieux G, Mirabbasi S. Interconnect driver design for long wires in field-programmable gate arrays. IEEE International Conference on Field Programmable Technology, 2006
- [16] Lemieux G, Lee E, Tom M, et al. Directional and single-driver wires in FPGA interconnect. IEEE International Conference on Field-Programmable Technology, 2004
- [17] Wu F, Zhang H, Lai J. A delay-optimized uniform programmable routing circuit. Proc 14th Asia and South Pacific Design Automation Conference, 2009: 135
- [18] Xie Ding, Shao Yun, Lai Jinmei, et al. A FDE2009 software system for programmable logic device with hierarchical architecture. Chinese Journal of Electronics, 2010, 38(5): 1136
- [19] Pan Guanghua, Lai Jinmei, Chen Liguang, et al. The design and implementation of sequential circuits in FPGA configurable logic block. Chinese Journal of Electronics, 2008, 36(8): 1480