Improving Execution Speed of FPGA using Dynamically Reconfigurable Technique

Roel Pantonial, Md. Ashfaquzzaman Khan, Naoto Miyamoto[1], Koji Kotani, Shigetoshi Sugawa, Tadahiro Ohmi[1]

Graduate School of Engineering, Tohoku University
[1]New Industry Creation Hatchery Center, Tohoku University
Aza Aoba 6-6-10, Aramaki, Aoba-ku, Sendai, 980-8579, Japan
Phone/Fax: +81-22-795-3977 / +81-22-795-3986, E-mail: miyamoto@fff.niche.tohoku.ac.jp

***Abstract*- This paper studies issues concerning dynamically reconfigurable FPGA (DRFPGA). It reports the architecture and performance of Flexible Processor III (FP3), a newly proposed DRFPGA. The FP3 employs a new shift register-type temporal interconnect to reduce operation delay. Designed and fabricated in 0.35um 2P3M CMOS technology, FP3 works correctly as a multi-context FPGA. Our experimental results show that there exist cases where the best user circuit speed was achieved when 2 contexts were in use for a benchmark circuit. This is because of the reduction of buffers in the critical path by temporal partitioning.**

I. Introduction

Dynamically customizable and reconfigurable hardware architecture for a specific task "on demand" is one of the most important issues to bring out a novel-computing paradigm in the era of system LSIs. Dynamically Re-configurable FPGAs (DRFPGAs) allow dynamic reuse of logic and interconnect resources by having more than one on-chip configuration memory planes. Each memory plane is called a context and switching between contexts, that is reconfiguration, can be done by changing a dedicated signal that determines from which context configuration data should be read. In the DRFPGAs, the user circuit is temporally partitioned into several sub-circuits such that sequential execution of these sub-circuits yields the same results as the original user circuit. The cycle of a sub-circuit is called 'micro-cycle' and that of the user circuit is called 'user cycle'. One user cycle is made up of one pass through all the micro-cycles as shown in Fig. 1. To support temporal data transfer among sub-circuits, DRFPGAs must have proper interconnect architecture called temporal communication module (TCM)[1,2] or micro-register [3].

While dynamic reconfiguration obviously increases logic density, it also accompanies additional delay due to context switching, which includes TCM data store and load delay. All previously reported works i.e. Trimberger's Time-Multiplexed FPGA [3] and Flexible Processor [1,2], suffered from drastic fall in execution speed when more than one context were in use.

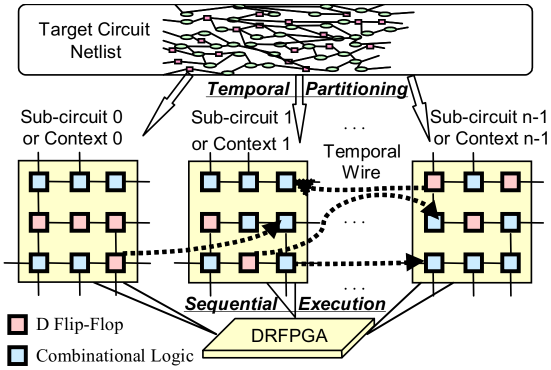

Fig. 1 Implementing a user circuit on a DRFPGA

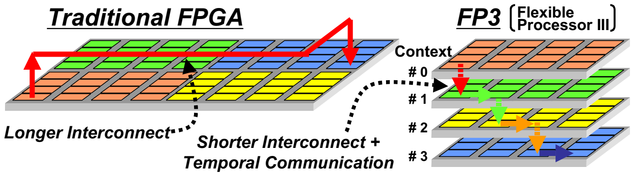

Fig. 2 Reduction of interconnect length in Flexible Processor III (FP3)

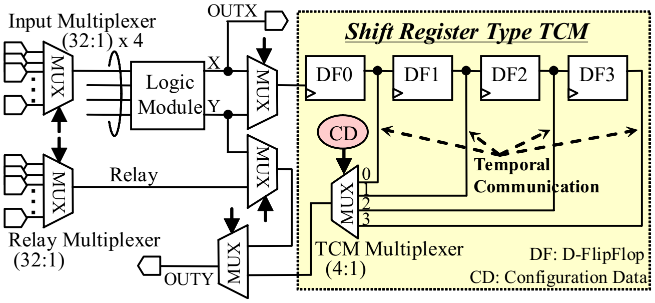

Fig. 3 Logic Element (LE) of FP3 including the proposed TCM

In the case where interconnect delay is far dominating than logic delay, dividing user circuit temporally can improve execution speed, since temporal partitioning divides a long spatial interconnect into several shorter ones via temporal communication as shown in Fig. 2. Flexible Processor III (FP3), our newly developed multi-context DRFPGA, reduced the TCM delay by employing a novel shift register-type TCM that makes the temporal data store and load time short enough to be negligible.

II. Shift Register-Type TCM

When a user circuit is divided temporally, data communication exists among the sub-circuits. Such communication is called temporal wire and the function of TCM is to support such temporal wire communication.

Figure 4 shows the measured interconnect delay of FP3. In order to reduce temporal data store and load time, the TCM is arranged in the structure of a shift register (Fig. 3). This is a newly proposed architecture and is better than the previously proposed ones [1,2,3] in following ways:

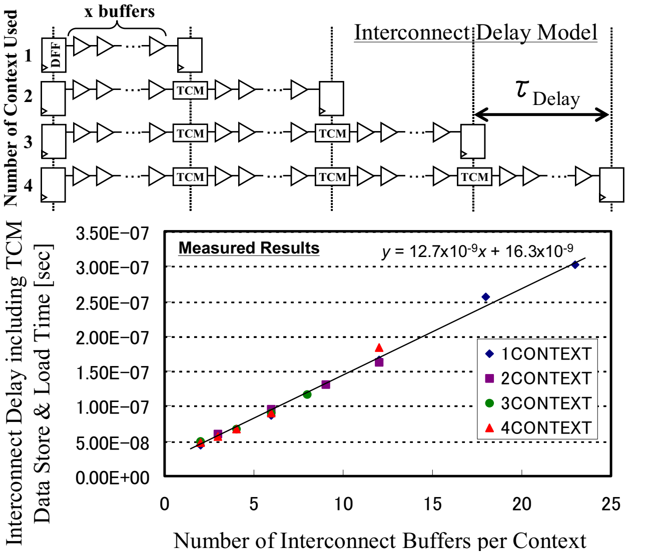

Fig. 4 Interconnect delay using the TCM. Proposed TCM enables to communicate temporal data with negligible overhead

1-4244-0630-7/07/$20.00 ©2007 IEEE.

1. *It reduces hardware cost, and reduces temporal data save and retrieve time to minimum.*
2. *It makes the execution sequence of a micro-cycle simple by eliminating load and store sequences, thus allowing high execution speed.*

Since the TCM overhead is negligible as shown in Fig.4, interconnect delay is constant against any number of contexts used. Using shift register makes temporal data move from DF#n to DF#n+1 in the TCM (Fig. 3) at the end of each micro-cycle. Further, this move is sequential and can be easily handled by CAD tool. Thus, the proposed TCM improves performance and reduces hardware cost without sacrificing any FPGA's functionality and flexibility.

III. Chip Fabrication and CAD Tool Developments of FP3

Fig. 5 shows the chip die photo of FP3. It has a 24x24 array of logic elements (LEs) and 4x24 IOBs (I/O Blocks) on all sides of the chip. Each LE, including its interconnect, can be configured in just one clock cycle using a 32-bit dedicated configuration data bus. Thus, at most only 600 clock cycles are necessary for a context configuration.

A full automatic Partitioning-Placement-Routing tool called PELOC has been developed to map any user circuits on FP3. Its snapshot is shown in Fig. 6. It takes a verilog netlist as its input and generates configuration data.

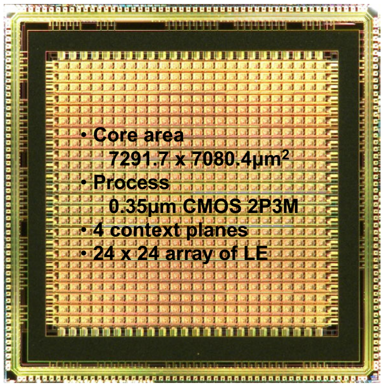

Fig. 5 FP3 chip die photo

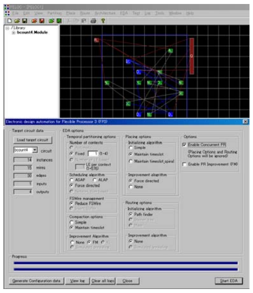

Fig. 6 PELOC snapshot. PELOC is a CAD tool for FP3

IV. Measurement Results and Discussions

The FP3 chip has been fabricated and its correct behavior has been confirmed using an LSI tester. Fig. 7 shows the FP3 performance when 4-bit and 8-bit binary counters were implemented with the number of contexts 1 to 4. For 4-bit binary counter, *micro-cycle delay* is almost the same for any number of contexts used, while, for 8-bit binary counter, utilizing only 1 context generates 2.43 times longer delay than the others. For 4-bit binary counter, *user-cycle delay*, which is defined as the micro-cycle delay times the number of contexts used, is linearly increasing, while, for 8-bit binary counter, there is a lowest-bound when using 2 contexts. From these results, using 1 context does not always produce minimum operation delay, and the optimal number of context differs from each user circuit. **This obviously indicates that there exists case where multi-context execution of DRFPGAs is faster than single-context execution of the conventional FPGAs.**

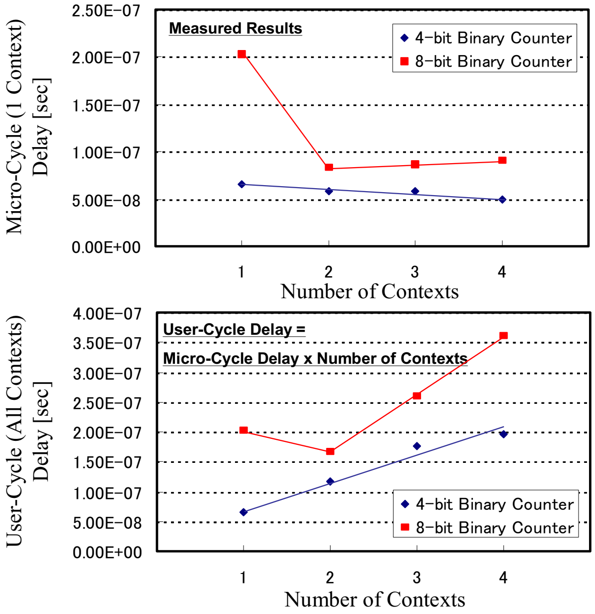

Fig. 7 Measurement results of 4-bit and 8-bitbinary counters

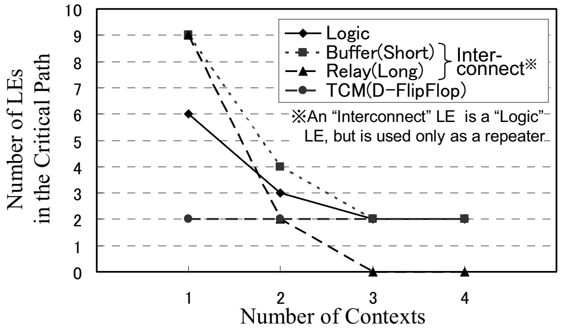

Fig.8 Number of LEs in the critical path of 8-bit binary counter after placement and routing by CAD tool (PELOC)

Fig. 8 shows the number of LEs in the critical path categorized by logic, buffer, relay and TCM. Here, a logic LE is assigned to a particular Boolean function in the netlist. A buffer is a LE that serves only as an interconnect between a source LE and a sink LE. A relay is used as an intermediate distance interconnect that reaches twice as long as the buffer. Either TCM or D-Flipflop is utilized by a TCM LE. From the figure, as the number of contexts increases by the temporal partitioning, the number of LEs, especially interconnect LEs, in the critical path decreases. We consider that this is because by temporally sub-dividing the user circuit continuously, its complexity (for example Rent's parameter) decreases until such point where it meets the FPGA's routing flexibility, and then the circuit works with minimum delay. Thus, it is very effective to use temporal partitioning and dynamic reconfiguration, in order to decrease the number of interconnect LEs as well as the critical path delay in FPGAs.

V. Conclusion

This paper reports the architecture and performance of FP3, a newly proposed DRFPGA. A shift register-type TCM was developed which enables to communicate temporal data with negligible overhead. From the measurements, we found that, for 8-bit binary counter, the best user circuit speed was achieved when 2 contexts were in use (not 1 context as conventional FPGAs). This result is interesting since it establishes possibility that the temporal partitioning reduces total critical path delay, by converting spatial interconnect delay into temporal communication delay.

ACKNOWLEDGMENTS

The FP3 in this study was designed with Cadence, Synopsys and Avant! CAD tools and fabricated through the chip fabrication program of VLSI Design and Education Center (VDEC), the University of Tokyo, with the collaboration by Rohm Corporation and Toppan Printing Corporation.

REFERENCES

[1] K.Kotani et al. ,"A Personal-use Single Chip Emulator Using Dynamically Reconfigurable Logic Array", in proc. of A-SSCC 2005

[2] T. Ohkawa, et al., "The Flexible Processor- Dynamically Reconfigurable Logic for Personal-Use Emulation System", Symposium on VLSI Circuits, Digest of Technical Papers, pp.279-282, 2003

[3] S. Trimberger, et al., "A Time-Multiplexed FPGA", in proc. of the IEEE FCCM, pp.22-28, 1997