Subframe Multiplexing for FPGA Manufacturing Test Configuration

Erik Chmelar
echmelar@crc.stanford.edu

The cost of FPGA manufacturing test is driven by the number of test configurations and the time required to configure an FPGA under test. The high degree of regularity of an FPGA, which is an array of interconnected logic blocks, results in manufacturing test configurations that are also highly regular. Subframe multiplexing takes advantage of the regularity in test configurations to greatly reduce the configuration time, the dominant portion of the test time. This design for test technique requires only a minor modification to the existing configuration hardware---the addition of several two-bit multiplexers---while maintaining the same configuration memory architecture. Results are shown for Xilinx Virtex and Virtex-E series FPGAs, yielding a configuration time reduced by several orders of magnitude compared to the best current technique.

An FPGA Implementation of Block Truncation Coding for Gray and Color Images

Sherif M. Saif, Hazem Abbas, Salwa M. Nassar
Mentor Graphics Egypt, 51 Beirut Str., Heliopolis, Cairo, Egypt
hazem_abbas@mentor.com

This paper presents an FPGA implementation for the Block Truncation Coding (BTC) image compression technique. Images are divided into equal blocks. The BTC calculates the mean of each block for which a two-level quantization is performed so that a zero value is stored for the pixels with values smaller than the mean. The rest of the pixels are represented by the value one. The implementation exploits the inherent parallelism of the algorithm to provide efficient algorithm-to-architecture mapping. FPGA implementation of the BTC is composed of three modules: the input module, to receive input pixels; the quantizer module, to classify pixels to one of the two levels; and divider circuits to obtain the two quantized values. The implementation is performed for gray and color images. The Xilinx-VirtexE BTC implementation has shown to provide about 23.4×10^6 of pixels/second processing rate which is about 3500 times faster than an Intel Pentium III-550-MHz processor.

An FPGA Implementation of Beneš Permutation Networks

Richard N. Pedersen, Anatole D. Ruslanov, Jeremy R. Johnson
Department of Computer Science, Drexel University,Philadelphia, PA USA
{anatole, jjohnson}@cs.drexel.edu, Rich.Pedersen@smiths-aerospace.com
Technical report DU-CS-03-02
http://www.cs.drexel.edu/page.php?name=publications.html

This work discusses an FPGA implementation study of the Beneš Permutation Network (BPN). The BPN, originally developed for connecting devices in telephone switching, is a circuit of size O(n log n) and O(log n) depth, built from 2 x 2 switches, which is capable of performing an arbitrary permutation. The BPN provides an asymptotic improvement in area over the straightforward network built with multiplexers, and the work presented here shows that an FPGA implementation uses less area for networks as small as size 4. The implementation presented in this paper uses a special-purpose tool to synthesize and place and route the circuit. The place and route tool can be used to systematically explore alternative place and route strategies and was used to obtain significantly better area utilization and timing performance compared to general-purpose tools. In addition, several general improvements and extensions were discovered that further improve performance and reduce area.

Routing Architecture for Multi-Context FPGAs

Andrea Lodi, Roberto Giansante, Carlo Chiesa, Luca Ciccarelli,
Mario Toma, Fabio Campi
{andrea.lodi,rgiansante,cchiesa,lciccarelli,mtoma,fcampi}@deis.unibo.it

Multi-context FPGAs are a convenient solution for run-time reconfiguration, but they suffer from large area occupation. This is mainly due to programmable interconnect configuration memories which need to be replicated as many times as the number of contexts. This limitation can be overcome with a DBM (Decoder Based Multicontext) routing architecture which introduces a decoding stage between configuration memories and routing structures, so that the number of SRAM cells can be highly reduced. Both the switch and connect blocks have been designed, at the architectural and schematic level. A suitable router for DBM structures has also been developed. The results show that the approach adopted requires on average only 5% more tracks whereas the device area occupation is reduced up to 40%. We also show that DBM routing architecture does not affect the critical path delay, due to the reduction of the routing wire length.

A Flexible Hardware Architecture for 2-D Discrete Wavelet Transform

Richard Carbone and Andreas Savakis
Department of Computer Engineering, Rochester Institute of Technology
Rochester, NY 14623 USA
savakis@mail.rit.edu
http://www.rit.edu/~axseec/

We present the design and implementation of a flexible hardware architecture used to perform the 2-D Discrete Wavelet Transform (DWT). A generic VHDL IP core is developed that is configurable to perform the forward and inverse 2-D DWT on an image using a user-defined biorthogonal wavelet. The DWT is performed using the hardware-efficient Lifting Scheme method versus the traditional convolution-based methods minimizing memory requirements, eliminating the need for expensive floating-point arithmetic, and increasing speed. This DWT core is synthesized to a Xilinx Virtex FPGA and configured to perform the Cohen-Dabuechies-Feaveau 5/3 and 9/7 DWTs, which are used in the JPEG2000 still image compression standard.

A VHDL MPEG-7 Shape Descriptor Extractor

Bret Woz and Andreas Savakis
Department of Computer Engineering, Rochester Institute of Technology
Rochester, NY 14623 USA
savakis@mail.rit.edu
http://www.rit.edu/~axseec/

The MPEG-7 video compression standard allows for multimedia metadata description. By providing robust descriptors and an effective framework for storing them, MPEG-7 is designed to provide a means of navigation through audio-visual content. In particular, MPEG-7 provides a region based shape descriptor, the Angular Radial Transform (ART), for use in image and video annotation and retrieval. For this work, an FPGA based ART extractor was designed and simulated for a Xilinx Virtex-E XCV300e in order to provide a speedup over software based extraction. The design created is capable of processing over 69,4400 pixels a minute. This design utilizes 99% of the FPGA's logical resources and operates at a clock rate of 25 MHz.

Using an FPGA Coprocessor for Improving Execution Speed of TRT-LUT - One of the Feature Extraction Algorithms for ATLAS LVL2 trigger

C. Hinkelbein, A. Khomich, A. Kugel, R. Männer, M. Müller
Institute of Computer Science V, University of Mannheim, GERMANY
{hinkelbein, khomich, kugel, maenner, mmueller}@ti.uni-mannheim.de
http://www-li5.ti.uni-mannheim.de/fpga/

This work investigates the suitability of using an FPGA coprocessor for speedup track finding algorithm for ATLAS Level 2 trigger. Two realizations of the same algorithm have been compared: C++ realization tested on a computer equipped with dual Xeon 2.4 GHz CPU, 64Bit/66MHz PCI bus, 1024 Mb DDR RAM main memories with Red Hat Linux 7.1; and hybrid C++ and VHDL realization tested on the same PC equipped in addition by MPRACE board (FPGA-Coprocessor board based on Xilinx Virtex-2 FPGA and made as 64Bit/66MHz PCI card developed at the University of Mannheim). In the TRT-LUT algorithm, the most time consuming parts were implemented in VHDL and using the FPGA coprocessor. This realization can give us speed-up by factor ~2 for hybrid FPGA/CPU realization in comparison with CPU-only implementation.

Hardware Co-Simulation in System Generator of the AES-128 Encryption Algorithm

Daniel Denning, Malachy Devlin, James Irvine
Institute of System Level Integration
Alba Centre, Alba Campus, Livingston, EH54 7EG, UK
daniel.denning@sli-institute.ac.uk

We discuss the use of System Generator to hardware co-simulate in the FPGA versions of the AES-128 encryption algorithm. We show that the FPGA co-simulation of the AES can be achieved over 3 different bus types (TCP/IP, board-level TCP/IP, and PCI). One of the FPGA co-simulations is over 3 times faster running over a TCP/IP network distance off approximately 600 kilometres, than running a normal Simulink simulation on the host PC. Another hardware co-simulation time increases in the region of 4000% running over the PCI bus attached to the host PC. By having this FPGA co-simulation option, some of the IP cores in an FPGA system can be co-simulated, there by freeing up processing power on the host-PC for further developments in a system.

Buffer Schemes for Runtime Reconfiguration of Function Variants in Communication Systems

Dirk Eilers, Helmut Steckenbiller, Rudi Knorr
Fraunhofer Institute for Communication Systems
Hansastr.32, 80686 Munich, GERMANY
dirk.eilers@esk.fraunhofer.de

This contribution introduces distributed buffering schemes for runtime reconfiguration in adaptive processing systems, e.g., real-time streaming media applications. With dynamic reconfiguration, the area-cost of field-programmable logic can be reduced by reuse of area resources, and potential for adaptive signal processing techniques can be enabled. The challenge with runtime reconfiguration is the reconfiguration latency. Given the limitations regarding reconfiguration latency with traditional approaches, we propose distributed buffering schemes, to hide latency for fast adaptive systems. To verify our approach, we do analytic and simulative modeling. The results show that our approach can guarantee zero loss-rates and reduce the additional delay (reduction of delay over 65% with estimated application profiles) in cases, where the traditional approach would not reach a desired Quality of Service. Dependent on the worst case application profile, it enables potential for runtime reconfiguration of adaptive signal processing under real-time constraints. Finally, we derive an architecture template with implementation details.

FPGA-Based Implementation of Single-Precision Exponential Unit

Christopher Doss, R.L. Riley Jr.
North Carolina A & T, Dept. of EE, 551 McNair Hall
Greensboro, NC 27411 USA
cdoss@ncat.edu

This work explores the feasibility of implementing a floating-point exponentiation unit on reconfigurable computing systems. A table-driven exponentiation unit was implemented using synthesizable VHDL. The hardware implementation was performed using the same methodology as Bui *et. al*. We first implemented the individual components required to perform the exponentiation operation. These include such operations as IEEE-754 floating-point addition, multiplication, comparison, and power of two. These modules were then linked together to form the overall unit. The designs were synthesized, placed and routed. Results indicate that today's FPGAs are well suited for this operation. The exponentiation unit was simulated, synthesized, and placed and routed to a Xilinx Virtex 2 4000 part with a speed grade of 5. The unit requires 5564 of the available 23,040 slices, and is capable of operating at 87 MHz.

High-speed Systolic Array for Gene Matching

G. Caffarena, S. Bojanic, J. A. Lopez, C. Pedreira, O. Nieto-Taladriz
Dep. de Ingenieria Electronica
E.T.S.I. Telecomunicacion, Universidad Politecnica de Madrid
C. Universitaria s/n, 28040. Madrid, Spain
gabriel@die.upm.es

One of the main challenges in bioinformatics nowadays is to create a framework to compare efficiently new DNA sequence information to large existing databases. Optimal methods, such as the Smith-Waterman algorithm, provides more sensitive results than heuristic algorithms such as FASTA and BLAST, with the drawback of increased computational complexity. FPGAs implementations of Smith-Waterman exploit the intrinsic parallelism in the algorithm and achieve computation time reductions of several orders of magnitude. In this paper we present an implementation of a Smith-Waterman linear systolic array that doubles throughput of existing ones with a minimum increase of area. The processing speed improvement is achieved by means of assigning the computation of four edit distances to a single systolic cell. Implementation results on Xilinx devices XC2V6000 and XCV1000E are reported. An array that processes 8000-nucleotid DNA sequences achieves 3200 BCUPS operating at 200 Mhz on a XC2V6000-5 device.

The Gigahertz FPGA: Design consideration and Applications

Jong-Ru Guo, C. You, M. Chu, R. Heikaus, K. Zhou, O. Erdogan, J. D. Diao, B.S. Goda*, R.P. Kraft, J.F. McDonald
Rensselaer Polytechnic Institute, Troy, NY, 12180 USA
* United State Military Academy, West Point, N.Y. 10096 USA

This paper describes the implementation of a large scale SiGe FPGA that serves as a high speed FPGA test platform. In the FPGA core, 20 x 20 building blocks (Basic Cells) are used to implement logic applications. This chip contains 10^6 devices including SiGe NPNs and MOSFETs. This chip is fabricating with the IBM SiGe 7HP process with cut off frequency of 120GHz. The target running frequency of this FPGA is 10GHz. Clock repeaters are added for improved clock distribution. A test circuit whose building block cell runs up to 10GHz is fabricated and measured by the same process. Future work and some potential applications of the SiGe FPGA are also described.

Transistor Grouping and Metal Layer Trade-offs in Automatic Tile Layout of FPGAs

Ian Kuon, Aaron Egier, and Jonathan Rose
Edward S. Rogers Sr. Department of Electrical and Computer Engineering
University of Toronto, Toronto, Ontario, Canada M5S 3G4
{ikuon|aegier|jayar}@eecg.utoronto.ca
http://www.eecg.toronto.edu/~jayar/pubs/ATL/ATL.html

The physical layout of modern commercial FPGAs is one of the last bastions of manual VLSI layout. Our recent work has automated the FPGA layout process from architectural description to mask-level layout of the repeated FPGA tile. Here we improve on that work using two approaches: 1) by making better choices for the grouping of circuitry into the cells used for the layout and 2) through better allocation of metal. The new groupings improve the FPGA tile area by between 10% and 14%. That, together with the superior metal layer allocation allows us to automatically lay out a very accurate capture of a Xilinx Virtex-E tile that is only 54% to 92% larger than the real thing. With two additional metal layers, our tile area is only 13% larger. In addition, we show that a standard cell implementation is 102% larger than the real Xilinx Virtex-E.

Bit-level Super-Systolic Array for FIR Filter with a FPGA-based Bit-serial Semi-Systolic Multiplier

Jae-Jin Lee, Gi-Yong Song
School of Electrical and Computer Engineering
Chungbuk National University, Cheongju Chungbuk, 361-763, Korea
ceicarus@just.chungbuk.ac.kr, gysong@chungbuk.ac.kr

To achieve higher degree of concurrency in a systolic array, it is desirable to make cells of a systolic array themselves a systolic array as well, leading to a structure called super-systolic array. This paper proposes a bit-level super-systolic FIR filter with a FPGA-based bit-serial semi-systolic multiplier. Compared to the word-level systolic FIR filter and corresponding super-systolic filter, the proposed design is very compact in that it needs only two 1-bit I/O ports in addition to significant improvement on hardware complexity. The input to the implementation proposed in this paper is a sequence of bit-string in contrast to the distributed arithmetic which assumes input of parallel bit-string. Also, the arrangement of the cells of a semi-systolic multiplier is tuned to fit to the structure of the FPGA. (This work was done as a part of Information & Communication fundamental Technology Research Program supported by Ministry of Information & Communication in republic of Korea.)

High Level Area, Delay and Power Estimation for FPGAs

Tianyi Jiang, Xiaoyong Tang, Prith Banerjee
Electrical and Computer Engineering Northwestern University
2145 Sheridan Road, Evanston, IL 60208 USA
{jiang, tang, banerjee}@ece.northwestern.edu

This paper describes an approach for high-level estimation of area, delay and power for FPGA synthesis. This approach has been integrated within the PACT compiler framework which has an automated design space exploration pass that determines the effects of various compiler optimizations on the synthesized hardware. Such a pass needs early estimation of area, delay and power. Towards this end, we have developed area and delay models for various RTL level operators such as adders, multipliers, and logical operators, which are parameterized with the bit widths of the devices. We have also derived high-level equation based power macro-models which take into account input switching activities, input spatial correlation and input bit width. These models are derived by actual synthesis of the RTL operators using back-end logic synthesis and place-and-route tools. Experimental results show that these area, delay and power models are accurate and efficient.

Fast Adders in Modern FPGAs

Jianhua Liu[1], Michael Chang[1], Chung-Kuan Cheng[1],
John MacDonald[2], Nan-Chi Chou[2], Peter Suaris[2]
[1]University of California, San Diego
[2]Mentor Graphics
{jhliu, mlchang, kuan}@cs.ucsd.edu,
{john_macdonald, nanchi_chou, peter_suaris}@mentor.com

Binary addition is one of the most frequent operations in computation systems. Dedicated carry logic in modern FPGA devices allows ripple-carry adders to outperform other kinds of adders. However, a long carry chain is still time- consuming in a wide bit-width adder. We propose a new methodology to partition the long carry chain into short segments and organize these segments by applying carry-select or carry-skip schemes. Therefore, the resulting adder can take advantage of fast carry ripple locally, reduce the long signal delay globally, and produce high performance calculations.

Addressing Application Integrity Attacks Using a Reconfigurable Architecture

Joseph Zambreno[1], Rahul Simha[2], Alok Choudhary[1]
[1]Dept. of ECE, Northwestern University, Evanston, IL 60208 USA
[2]Dept. of CS, The George Washington University, Washington, DC 20052 USA
{zambro1, choudhar}@ece.northwestern.edu
simha@gwu.edu

Growing concerns regarding application security and software piracy have motivated research in systems that ensure an increased level of tamper resistance while limiting the effectiveness of malicious attacks. These approaches have ranged from simple code restructuring techniques containing run-time checks to complex cryptographic systems. In this work we propose a reconfigurable software protection architecture that places an FPGA between the highest level of on-chip cache and main memory. Instructions requested by the processor are passed through the FPGA component after being fetched from memory. The task of the FPGA is to validate and also possibly transform the instructions in some fashion before sending them back to the processor. As the FPGA configuration can be customized to individual applications, the resultant system can be flexible in meeting both security and performance requirements. Our initial results show that a strong level of security can be obtained with only a limited performance overhead.

SPFD-Based One-to-Many Rewiring

Katsunori Tanaka[1], Shigeru Yamashita[2] and Yahiko Kambayashi[1]
[1]Kyoto University, Kyoto, 606-8501, Japan
[2]Nara Institute of Science and Technology, Ikoma, Nara, 630-0101, Japan
{ktanaka, yahiko}@db.soc.i.kyoto-u.ac.jp, ger@is.nara-aist.ac.jp

This presentation gives a new SPFD-based method to make a target wire redundant with higher possibility than the previous rewiring methods. After placement and routing, we obtain or estimate how much the existing and candidate wires interferes such physical design, and thus, we find out the most critical one of the existing wires. In order to remove it, the previous methods add new input wires to only one LUT. Actually, in the context of gate/cell-based logic optimization, it is well-known that input wire addition to many gates/cells is likely to make the critical one redundant. However, in an FPGA circuit, since an LUT can realize an arbitrary function for a specified number of inputs, such wire addition has never been considered to be useful. In this presentation, we provides with an SPFD-based condition for such wire addition to improve the FPGA circuit performance. We also present several experimental results to show the effectiveness of the proposed condition.

SHAPER: Synthesis for Hybrid FPGA Architectures containing PLA Elements using Reconvergence Analysis

Manoj Kumar.A, Jayaram.B and Kamakoti.V
Dept of Computer Science,IIT-Madras,
Chennai-600036 INDIA
amk_141@yahoo.co.in, amk@peacock.iitm.ernet.in

This paper discusses the technology mapping problem on Hybrid Field Programmable Architectures (HFPA). HFPAs are realized using a combination of the following programmable logic devices, namely, Lookup Tables (LUTs) and Programmable Logic Arrays (PLAs). The HPFAs provide the designers with the advantages of both LUT-based Field Programmable Gate Arrays and PLAs. Specifically, use of PLAs can result in reduction of area required for mapping a circuit. Designing a methodology that maps a given circuit on to the HFPA that exploits the above-mentioned advantages to the maximum is a problem of very great research and commercial interest. This paper presents the SHAPER, which maps the circuits onto HFPAs using reconvergence analysis. Empirically, it is shown that SHAPER yields 18% better area-reduction than the previous known algorithms.

Least-Significant Bit Optimization Techniques for FPGAs

Mark L. Chang, Scott Hauck
Dept of EE, University of Washington, Seattle, WA USA
{mchang, hauck}@ee.washington.edu
http://acme.ee.washington.edu

We present a methodology for FPGA datapath precision optimization subject to user-defined area and error constraints. This work builds upon our previous research which presented a methodology for optimizing for dynamic range--the most significant bit position. In this work we define the problem of least significant bit optimization and propose optimization techniques that provide finer control of area-to-error tradeoffs than more traditional methods. We present some preliminary results describing the effectiveness of our techniques on typical signal and image processing kernels.

FPGA Modelling for High-Performance Algorithms

Martin Danek[1], Josef Kolar[2]
[1]Institute of Information Theory and Automation, Czech Academy of Sciences
[2]Dept. of Computer Science & Engineering, Czech Technical University
danek@utia.cas.cz, kolar@fel.cvut.cz

The poster deals with topological modelling of FPGA circuits for timing-driven algorithms. It presents a method for analysing and deriving topological placement/routing models from architectural description of existing FPGAs. A metric is introduced that reflects information loss in more abstract global routing models. The metric captures both the loss of topological information and the decrease in precision of possible signal delay estimation based on the model. The practical use of the metric is demonstrated for a wire-type model and a global routing model derived from the Xilinx XC4000 family and for two linear and one Elmore-based signal delay estimation models. This research has been partially supported by the Grant Agency of the Czech Republic under Project No. 102/04/2137.

A Constraints Programming Approach to Communication Scheduling on SoPC Architectures

Christophe Wolinski[1], Krzysztof Kuchcinski[2], Maya Gokhale[3]
[1]IRISA, IFSIC France
[2]Department of Computer Science, Lund University, Sweden
[3]Los Alamos National Laboratory,Los Alamos, NM, USA, maya@lanl.gov

This paper presents a novel approach to scheduling communications among concurrent hardware processes mapped onto a "System on a Programmable Chip." Point-to-point, broadcast and multi-cast ommunication types are supported. The algorithm has been prototyped on the Processor-Coupled Polymorphous Fabric for the Altera Excalibur Arm architecture. The communication schedule problem has been specified using Constraints Programming. The advantages of our method are the following: Application of a general constraint solver makes it possible to express many different sorts of constraints in a uniform manner. All imposed constraints are handled by the solver concurrently which increases the chances of obtaining optimal results. The scheduler guarantees that loops periods are the same for each iteration so the smaller controllers can be generated. The method has been illustrated with a Fabric-based implementation of the K-means clustering algorithm for which an optimal communication schedule has been achieved.

A Novel Coarse-Grain Reconfigurable Data-Path for Accelerating DSP Kernels

M.D. Galanis[1], G. Theodoridis2, S. Tragoudas[3], D. Soudris[4], and C.E. Goutis[1]
[1] VLSI Design Lab., ECE Dept., Univ. of Patras, Rio 26110, Greece
[2] Physics Dept., Aristotle Univ. Thessaloniki 54124, Greece
[3] Elect. & Comp. Eng. Dept., Southern Illinois Univ., Carbondale 62901, USA
[4] VLSI Design & Testing Center, ECE Dept., Democritus Univ., Xanthi 67100, Greece
mgalanis@vlsi.ee.upatras.gr

In this paper, an efficient implementation of a high performance coarse-grain reconfigurable data-path on a mixed-granularity reconfigurable platform is presented. It consists of several coarse grain components of the same type, a reconfigurable inter-component network, and a centralized register bank. The universal type of coarse grain component is shown to increase the system's performance due to significant reductions in the latency. A flexible interconnection network facilitates the data transfers between the coarse grain components and also from or to the register bank. An automated methodology for mapping DSP and multimedia kernels on the data-path is also presented. Chaining of operations is optimally exploited, and the architecture allows for simple and efficient algorithms for scheduling, live signal reduction, and component binding. Experimental results verify the impact of our architectural decisions and design automation methods.

Improving the Reliability of FPGA Circuits Using Triple-Modular Redundancy (TMR) & Efficient Voter Placement

Michael J. Wirthlin
Department of ECE, 459 CB , Provo, UT 84602 USA
wirthlin@ee.byu.edu

Triple-modular redundancy has been proposed as a technique for improving the reliability of FPGA circuits in the presence of radiation-induced SEUs. This technique masks circuit faults by voting on the output of three identical circuit modules. A critical design decision in any TMR system is the placement of voters between the TMR modules. This paper presents a technique for selecting the appropriate location of voters within feedback paths of an FPGA circuit. Voting within the feedback path will repair corrupted state variables and tolerate online repair of the FPGA configuration.

Multi-Resource Aware Partitioning Algorithms for FPGAs with Heterogeneous Resources

Navaratnasothie Selvakkumaran[1], Abhishek Ranjan[2], Salil Raje[2], George Karypis[1]
[1]Dept of C.S. University of Minnesota
[2]HierDesign, Inc.
{selva, karypis}@cs.umn.edu, {ranjan, salil}@hierdesign.com

As the chip densities increase, the modern FPGAs contain large capacity and increasingly provide heterogeneous units such as multipliers, processor/DSP cores, RAM-blocks etc, for efficient execution of crucial functions of the design. The hypergraph partitioning algorithms are generally used as a divide-and-conquer strategy, during synthesis and placement. The partitioning algorithms for designs with heterogeneous resources, need to not only minimize the cut, but also balance the individual types of resources. Unfortunately, the state-of-the-art multilevel hypergraph partitioning algorithms (hMetis,MLPart), are not capable of distinguishing the types of cells. To overcome this problem, we developed a new set of multilevel hypergraph partitioning algorithms, that are aware of multiple resources, and are guaranteed to balance the utilization of different resources. By evaluating these algorithms on large benchmarks, we found that it is possible to achieve such feasible partitions, while incurring only a slightly higher cut (3.3%-5.7%) compared to infeasible partitions generated by hMetis.

FPGA-Based Supercomputing: An Implementation for Molecular Dynamics

Ian Kuon, Navid Azizi, Ahmad Darabiha,Aaron Egier, and Paul Chow
Deparment of ECE, University of Toronto, Toronto, Ontario Canada
{ikuon,nazizi,ahmadd,aegier,pc}@eecg.utoronto.ca

Current high-performance supercomputing applications are typically implemented on large-scale general-purpose distributed or multiprocessing systems often based on commodity microprocessors. FPGAs have now reached a level of sophistication that they too could be used for such applications. We explore the feasibility of using FPGAs to implement large-scale application-specific computations by way of a case study that implements a novel Molecular Dynamics system. The system has been designed such that it is scalable and parallelizable. On the Transmogrifier 3, the system performs calculations on an 8,192 particle system in 37 seconds at 26MHz. This implementation shows that by scaling to more modern parts running at 100MHz and using a better architecture, a speedup of over 20x can be achieved compared to a state-of-the-art microprocessor. This can also be achieved at less cost, using less power and taking less space than a standard microprocessor-based system, while maintaining the computational precision required.

An FPGA Prototype for the Experimental Evaluation of a Multizone Network Cache

Paul Berube, Jose Nelson Amaral, Mike MacGregor
Dept. of Computer Science, University of Alberta, Edmonton, Alberta, T6G 2E8 Canada
{berube,amaral,macg}@cs.ualberta.ca

Network routers rely on Content Addressable Memories (CAMs) to accelerate the process of looking up the next hop of a packet. We describe our implementation of a versatile prototype for a CAM. This prototype allows the empirical evaluation of the idea of caching lookup results in a multizone cache organized according to the length of the network prefix portion of the addresses. Implementing a cache in an FPGA efficiently required the design of a new cache replacement policy, the Bank Nth Chance policy. In this paper we present results from a functional simulator that allows the comparison of this new policy with existing ones such as LRU, FIFO and Second Chance. We have a complete and functional prototype with pipelined lookups implemented in a Xilinx Virtex 2000E device; we also report frequency of operation and occupation of the device.

An Algorithmic Approach by Heuristics to Dynamical Reconfiguration of Logic Resources on Reconfigurable FPGAs

Phan C. Vinh and Jonathan P. Bowen
London South Bank University, Centre for Applied Formal Methods
103 Borough Road, London SE1 0AA, UK
{phanvc,bowenjp}@lsbu.ac.uk

Efficient management of the logic resource available is one of the biggest problems faced by the embedded systems based on FPGA, in which their functionality can be partially modified at run-time without stopping the operation of the whole system. When the sequence of reconfigurations to be performed is not predictable, resource allocation decisions have to be made on-line. Dynamical reconfiguration can be necessary to relocate a running physical configuration, and to rearrange the logic resources into the variety of physical portions. Our proposed algorithm is formally developed to enable implementing an on-line management of FPGA logic resources, supporting the rearrangement of running functions, releasing enough contiguous space for configuration of new incoming functions, and performing the defragmentation in a way completely transparent to the applications currently running. Therefore, on-line scheduling of tasks in the spatial and temporal domains becomes possible, enabling the implementation of virtual hardware concept.

Dynamically Reconfigurable Architecture for High-throughput Processing of Data Centric Applications

Magesh Sadasivam and Sangjin Hong
Department of Electrical and Computer Engineering
Stony Brook University – SUNY
Stony Brook, NY 11794-2350 USA

This paper presents a reconfigurable platform for executing data centric applications. The proposed platform realizes applications where buffers are dominated in the design. Such characteristics of application domain is mostly exhibited by loop based processing systems. To increase the throughput, the platform is centered around buffers interacting through reconfigurable interconnect. Each buffer is associated with an autonomous reconfigurable controller where reconfiguration parameters are obtained statically from the dataflow graph. The execution concurrency is maintained by two level pipelining mechanism. We illustrate that the proposed platform is highly regular and decreases dynamic reconfiguration time significantly.

FPGA Implementation of a High Speed Network Interface Card for Optical Burst Switched Networks

Pronita Mehrotra, Mrugendra Singhai, Mike Pratt, Mark Cassada, Patrick Hamilton
Advanced Networking Research, MCNC-RDI
3021 Cornwallis Road, RTP NC 27709 USA
{pmehrot,msinghai,mpratt,mcc,phamilt}@anr.mcnc.org

Current generation of FPGAs with integrated high-speed transceivers provide a useful tool for prototyping various networking applications. We discuss a case study in developing a high-speed network interface card (NIC) for communication networks. The NIC implements a low-latency signaling protocol called Just-in-Time (JIT) for Optical Burst Switched networks. The main processing engine inside the FPGA runs at 62.5MHz and can handle data streams up to 1Gbps. The NIC utilizes the FPGA's gigabit transceiver cores, standard PCS/PMA and MAC layers and uses an optical front-end card to transmit data optically on specific wavelengths. The JIT engine in the NIC processes three kinds of messages - messages from the host, the network and internal timing messages associated with various timer events. The engine implements layer 3 functions typically implemented in software - like generating signaling messages and maintaining all active connections. This allows very fast setup and teardown of connections than otherwise possible.

Low Energy FPGA Interconnect Design

Rohini Krishnan, Jose Pineda de Gyvez, Martijn. T. Bennebroek
Philips Research Laboratories, Prof. Holstlaan 4, WAY 4.71,
Eindhoven, 5656AA, The Netherlands
rohini.krishnan@philips.com
http://www.research.philips.com

FPGAs are not energy efficient largely due to their programmable, capacitively loaded interconnect. We propose a new low energy FPGA interconnect architecture that is based on low energy switch blocks using Dynamic Threshold CMOS (DTMOS) based switches and an encoded-low swing (EL) technique. The presented case study, based on circuit simulations using SPICE in CMOS 0.13 micron process technology, illustrates that a 41% energy reduction can be achieved compared to the conventional techniques. A one to one comparison between NMOS based switches and the proposed DTMOS based switches reveal that the latter have a 36% lower power-delay product. We also show through a model analysis and circuit simulations that using low swing on interconnect, a timing budget can be met at 30% less energy consumption..

In-System FPGA Prototyping of an Itanium Microarchitecture

Roland E. Wunderlich and James C. Hoe
Computer Architecture Laboratory (CALCM)
Carnegie Mellon University, Pittsburgh, PA 15213-3890 USA
{rolandw, jhoe}@ece.cmu.edu

This work is part of our on-going effort to prototype an Itanium microarchitecture on an FPGA. To conserve time and effort in model development, we described our microarchitecture in Bluespec, a synthesizable high-level hardware description language. The microarchitecture model currently supports a subset of the Itanium instruction set architecture (ISA). The model includes details such as multi-bundle instruction fetch, decode and issue, parallel pipelined execution units with scoreboarding and bypassing, and multiple levels of cache hierarchies. The microarchitecture model is synthesized and prototyped on an FPGA that interfaces directly to the memory bus of a host PC. The prototyped microprocessor core executes the supported ISA subset at 100MHz and directly references the host-PC's DRAM and I/O resources through the memory bus at up to 800MB/sec of bandwidth. This effort is a first step toward developing a convenient in-system microprocessor prototyping platform capable of executing realistic full-scale applications and operating systems.

Online Placement Infrastructure to Support Run-Time Reconfiguration

Brian Leonard, Jeff Young, and Ron Sass
http://www.parl.clemson.edu/

The requirements for placing modules in an automatic run-time reconfigurable (RTR) system differ from those of ASIC and other static environments. The most notable difference is the continual addition and removal of modules from the FPGAs. We examine the effectiveness of a collection of two-dimensional placement decision algorithms in a RTR environment. New algorithms are proposed in addition to several which have been adapted from their one-dimensional counterparts. All of the algorithms have been tested on a set of benchmark applications. Six programs used for testing include examples from encryption, image processing and matrix manipulations, as well as arithemetic, assignment, and looping benchmarks. These applications are multiplexed and simulated to run on our RTR system. A simple last-accessed removal scheme with no compaction is currently implemented. The merit of each algorithm is determined by a set of factors that include fragmentation, chip utilization, decision time, and program run-time benefit.

Automatic Discovery, Selection, and Specialization of Modules in RCADE

Ranjesh G. Jaganathan, Matthew Simpson, and Ron Sass
http://www.parl.clemson.edu/

With increasingly larger FPGA devices, it is tempting to continue using conventional code generation techniques to construct hardware designs from high-level languages. However, in this paper we make a case against the application of conventional techniques to hardware design in favor of a new code generation approach that enables the use of heavily customized, parameterizable modules. This approach moves a portion of the generation functionality into tiles, which handle the negotiation of a module's parameters and subsequently the instantiation of the parameterized module. Furthermore, the tiles are automatically discovered at run-time and dynamically loaded into the design generator, which allows the system to be readily extended. This approach has been implemented in a collaborative, problem-solving design environment for reconfigurable computing. It is shown --- with respect to latency, throughput, and size --- that the proposed approach to hardware design generation is significantly more efficient than the conventional code generation techniques.

An Algorithm for Trading off Quantization Error with Hardware Resources for MATLAB based FPGA Design

Sanghamitra Roy, Debjit Sinha, Prith Banerjee
Electrical and Computer Engineering Northwestern University
2145 Sheridan Road, Evanston, IL 60208 USA
{sroy, debjit, banerjee} @ece.northwestern.edu

Most practical FPGA designs of digital signal processing applications are limited to fixed-point arithmetic owing to the cost and complexity of floating-point hardware. While mapping DSP applications onto FPGAs, a DSP algorithm designer, who often develops his applications in MATLAB, must determine the dynamic range and desired precision of input, intermediate and output signals in a design implementation to ensure that the algorithm fidelity criteria are met. The first step in a flow to map MATLAB applications into hardware is the conversion of the floating-point MATLAB algorithm into a fixed-point version using “quantizers” from the Filter Design and Analysis (FDA) Toolbox for MATLAB. We describe an approach to automate the conversion of floating-point MATLAB programs into fixed-point, for mapping to FPGAs by profiling the expected inputs to estimate errors. Our algorithm attempts to minimize the hardware resources while constraining the quantization error within a specified limit.

Roving Testing Using New Built-in-Self-Tester Designs for FPGAs

Vinay Verma[1] and Shantanu Dutt[2]
[1]Xilinx Inc. vinay@xilinx.com
[2]Univ. of Illinois at Chicago ECE Department, dutt@ece.uic.edu
http://www.ece.uic.edu/~dutt

We propose a roving tester (ROTE) that tests the PLBs of the FPGA by periodically moving across it. At any time, the ROTE occupies a certain area of the FPGA, say two columns, and tests all PLBs in that area using parallel built-in self-tester (BISTers). A significant contribution of this work are designs for 1- and 2-diagnosable BISTers. To the best of our knowledge, this is the first time that BISTer designs with provable diagnosabilities have been developed for FPGAs. We also develop functionality-specific testing methods that test PLBs in only two circuit functions that will be mapped to them (as opposed to testing PLBs in all their operational modes), for any reconfigurable fault pattern as the ROTE moves across the FPGA. The combination of our 1- or 2-diagnosable BISTer design and our functionality-specific testing technique leads to more accurate and faster test-and-diagnosis of FPGAs than achieved by previous work.

Preliminary Performance Analysis of Flex Power FPGA, a Power Reconfigurable Device with Fine Granularity

Takashi Kawanami[1,2], Masakazu Hioki[1], Hiroshi Nagase[2], Toshiyuki Tsutsumi[1,3], Tadashi Nakagawa[1], Toshihiro Sekigawa[1] and Hanpei Koike[1]
[1]Electroinformatics Group, Nanoelectronics Research Institute, National Institute of Advanced Industrial Science and Technology, JAPAN
[2]Department of Information and Computer Engineering, Kanazawa Institute of Technology, JAPAN
[3]Department of Computer Science, Meiji University, JAPAN
{t-kawanami, h.koike}@aist.go.jp

The Flex Power FPGA design is presented as a novel FPGA design offering the ability to configure the trade-off between power consumption and speed for each logic element by adjusting the threshold voltage. This design targets the reduction of static power consumption, which has become one of the most important issues in the development of future-generation devices. A method to effectively assign threshold voltages to transistors at a prescribed granularity based on a timing analysis of the mapped circuit is implemented using the VPR simulator, and the static power reduction for 70 nm technologies is estimated using MCNC benchmark circuits. Simulation results show that the average static power can be reduced to as little as 1/30 of that in the corresponding conventional FPGA. This design is also demonstrated to be effective with future technologies, where the proportion of static power will be greater.

On the Design of a Function-Specific Reconfigurable Hardware Accelerator for the MAC-Layer in WLANs

Thilo Pionteck, Thorsten Staake, Thomas Stiefmeier,
Lukusa D. Kabulepa, Manfred Glesner
Darmstadt University of Technology
Institute of Microelectronic Systems, Karlstr. 15, 64283 Darmstadt, Germany pionteck@mes.tu-darmstadt.de

This work presents the hardware design of a dynamically reconfigurable function unit (RFU) to accelerate computation-intensive tasks in Medium Access Control (MAC) layers of WLANs. The function unit is integrated in a pipelined 32 bit RISC processor and provides full hardware support for the Advanced Encryption Standard (AES) as specified in upcoming WLAN standards such as IEEE 802.11i. Dynamic reconfiguration allows the processor to use arithmetic components and memory elements of the RFU not only for AES, but also for additional tasks common in the MAC-layer. With our approach it is possible to accelerate Reed-Solomon-Code generation, Cyclic Redundancy Checks as well as other encryption standards like SQUARE, Magenta and Twofish by supporting Galois Field multiplication and table look-ups. The integration of the reconfigurable unit in the processor core results in an architecture that can simultaneously support control-flow and data-flow oriented tasks. This architecture was prototyped onto a Virtex2 FPGA.

Divide and Concatenate: A Scalable Hardware Architecture for Universal MAC

Bo Yang[1], Ramesh Karri[1], David A. McGrew[2]
[1]ECE Department, Polytechnic University, Brooklyn, NY, 11201 USA
[2]Cisco Systems, Inc. San Jose, CA 95134 USA
yangbo@photon.poly.edu, ramesh@india.poly.edu

We present a cryptographic architecture optimization technique called divide-and-concatenate based on two observations: (i) the area of a multiplier and associated data path decreases quadratically and their speeds increase gradually as their operand size is reduced. (ii) in hash functions, message authentication codes and related cryptographic algorithms, two functions are equivalent if they have the same collision probability property. In the proposed approach we divide a 2w-bit data path into two w-bit data paths and concatenate their results to construct an equivalent 2w-bit data path. We applied this technique on NH hash. When compared to the 100% overhead associated with duplicating a straightforward 32-bit pipelined NH hash data path, the divide-and-concatenate approach yields a 94% increase in throughput with only 40% hardware overhead. The NH hash associated message authentication code UMAC architecture with collision probability 2-32 that uses four equivalent 8-bit divide-and-concatenate NH hash data paths yields a throughput of 79.2 Gbps with only 3840 FPGA slices when implemented on a Xilinx FPGA.

Implementation of Elliptic Curve Cryptosystems over GF(2^n) in Optimal Normal Basis on a Reconfigurable Computer

Sashisu Bajracharya[1], Chang Shu[1], Kris Gaj[1], Tarek El-Ghazawi[2]
[1] George Mason University, [2] The George Washington University
kgaj@gmu.edu http://cpe02.gmu.edu/rcm

During the last few years, a considerable effort has been devoted to the development of reconfigurable computers, machines that are based on the close interoperation of traditional microprocessors and Field Programmable Gate Arrays. Several prototype machines of this type have been designed, and demonstrated significant speed-ups compared to conventional workstations for computationally intensive problems, such as codebreaking. In this paper, we demonstrate an efficient implementation of Elliptic Curve scalar multiplication over GF(2^n) in Optimal Normal Basis, using one of the leading reconfigurable computers available on the market, SRC-6E. We show how the hardware architecture and programming model of this reconfigurable computer has influenced the choice of the optimum program partitioning scheme. The detailed analysis of the control, data transfer, and reconfiguration overheads is given in the paper. The end-to-end speed-ups in the range from 895 to 1300 compared to the microprocessor implementation are demonstrated depending on the chosen partitioning scheme.

A Left-Edge Algorithm Approach for Scheduling and Allocation of Hardware Contexts in Dynamically Reconfigurable Architectures

Remy Eskinazi Sant'Anna[1], Manoel Eusebio de Lima[2], Paulo Romero Martins Maciel[2]
[1]Polytechnic School, Pernambuco University, Recife - PE – Brazil
[2]Center of Informatics, Federal University of Pernambuco, Recife - PE – Brazil
{res, mel, prmm}@cin.ufpe.br

This work aims to describe a methodology for scheduling and allocation of hardware contexts, in applications with high degree of parallelism, in a Run-Time-Reconfiguration (RTR) proceeding for a reconfigurable FPGA. The Scheduling approach is based on the hardware resource distribution in the FPGA architecture. The Scheduler is modeled as a Petri Net and the best performance yields the best scheduling. The hardware contexts allocation is based on a Left-Edge algorithm principle for rationalization of resources in scheduling approach. The adaptation of the algorithm considers that pre-located areas for loading of the contexts in the architecture are used.

Power Analysis and Estimation Tool integrated with XPOWER

E. Todorovich, E. Boemo, F. Cardells, J. Valls
School of Engineering, Universidad Autónoma de Madrid, España
etodorov@ii.uam.es

A power estimation tool for FPGA technology was developed. The current version is able to estimate total and individual node average power consumption in combinational blocks. The tool is statistical-based, allowing the user to specify the tolerated error for a given confidence level. Due to the fact that the power for individual nodes can be obtained, the software is able to connect this information with the physical positions of the nodes. In order to demonstrate the tool features and usability, a set of arithmetic blocks for digital radio have been utilized as benchmark circuit. Main results have been compared with actual power measurements for Virtex and Virtex-II devices.