Reducing the Cost of Floating-Point Mantissa Alignment and Normalization in FPGAs

Yehdhih Ould Mohammed Moctar¹ Nithin George² Hadi Parandeh-Afshar² Paolo lenne² Guy G.F. Lemieux³ Philip Brisk¹

¹University of California Riverside ²Ecole Polytechnique Fédérale de Lausanne (EPFL) ³University of British Columbia

> International Symposium on Field Programmable Gate Arrrays Monterey, CA, USA, February 22-24, 2012

Floating-point on FPGAs

- Best practice for HPC
 - Convert application into a deep, parallel pipeline
 - Altera's floating-point datapath compiler
 - Maxeler Technologies
 - ROCCC 2.0 (UC Riverside)
- Optimize for <u>throughput</u>, not latency
 - Reduce area
 - Fit more operators onto a fixed-size device
 - Shifters are a big bottleneck



Floating-point Addition Cluster

[Verma et al. FPL 2010]

- Similar to Altera's FP datapath compiler
- Add 2-16 single-precision FP operands at once
 - Denormalize in parallel up-front
 - Normalize the result at the end

Shifters are the area bottleneck when synthesized on an FPGA



Basic Logic Element (BLE)

FPGA Architecture (2/3)



Versatile Place and Route (VPR) CLB Architecture

FPGA Architecture (3/3)



Focus on Multiplexers

- Shifters are built from multiplexers
- FPGAs have lots of multiplexers
 - Focus on C-block and intra-cluster routing



Static Multiplexer (Standard FPGA)

Static-or-Dynamic Multiplexer (Patented by Xilinx—Alireza Kaviani) 6/32

Static vs. Dynamic Control

Under static control, one signal can route to any of the 8 multiplexer inputs.



Under dynamic control, 8 signals must route to the 8 multiplexer inputs in the correct order.



How are the dynamic control signals generated and how are they routed into the dynamic multiplexer?

Example: Conditional Swap



Example: Conditional Swap



Let's (Not) Try the C-Block



 Must route each signal on ONTO SPECIFIC SEGMENTS IN THE ROUTING CHANNEL!

Let's Try the Intra-cluster Routing



Strict Ordering Imposed on Signals Routed to CLB Inputs



Interconnect Topology Issues (1/2)



Both muxes implement the same logic function 13/32

Interconnect Topology Issues (2/2)



Changing the topology fixes the problem

Example: 4-bit Left Shift



Programmable Inversion



Routing Challenges (1/2)

- Traditional FPGAs provide a lot of flexibility to the router
 - C-block muxes
 - Intra-cluster routing muxes
 - Equivalence of LUT inputs



Routing Challenges (2/2)

- SD-Mux flexibility in the intra-cluster routing?
 - C-block muxes provide normal flexibility
 - Must route each net to a **<u>specific</u>** Intra-cluster routing mux input (CLB input)
 - LUTs offer no flexibility



Macro-Cells



Main Result

 The macro-cell routed successfully!

For a 27-bit shifter

 Routed all nets from normal CLB layer to pre-specified CLB inputs in the SD-Mux Enhanced layer



FPGA with Macro-cells (1/3)



FPGA with Macro-cells (2/3)



FPGA with Macro-cells (3/3)



Floating-point Addition Clusters [Verma et al., FPL 2010]



Experimental Setup

• VPR 5.0

- Project started several years ago
- Assumes intra-cluster routing is full-crossbar
 - We abstract away internal topology issues
- Significant modifications to P&R
 - Compute routes for the macro-cells
 - P&R large circuits with macro-cells

Parameter	Value	Parameter	Value
LUT Size	6	Fc input	0.15
Cluster Size	8	Fc output	0.1
Channel Width	96	Technology*	65nm CMOS
Cluster Inputs	36	Tile Area**	18940
* Berkeley predictive models		** Min-width transistors	

IWLS Benchmarks

- 10 largest benchmarks chosen
 - Much larger than MCNC, ISCAS, etc.
- Modified each netlist to add macro-cells
 - Macro-cells were kept off the critical paths

Benchmark	Description
ac97_ctrl	Interface to external AC 97 audio codec
aes_core	Advanced Encryption Standard (AES)
des_perf	16-cycle pipelined DES/3-DES Core
ethernet	10/100 Mbps IEEE 802.3/802.3u MAC
mem_ctrl	Embedded memory controller
pci_bridge32	Bridge interface to PCI local bus
systemcaes	Area-optimized AES implementation
usb_func	USB 2.0 compliant core
vga_lcd	Embedded VGA/LCD controller
wb_conmax	Wishbone Interconnect Matrix IP Core

Benchmark Overview



No Impact on Routing Delay!



 Locked-down resources (obstacles due to non-critical macro-cells) do not affect the critical path!

Impact on Min-channel Width



Router Runtime (not in paper)

PathFinder runtime



Limitations

- Real FPGAs use sparse crossbars for intra-cluster routing
 - Muxes may be smaller than 27:1
 - Did not model internal connections
- Did not model...
 - Area overhead of extra muxes, configuration bits, programmable inversion, etc. in the CLB
 - FP adder cluster frequency/latency
 - Energy consumption
- DSP blocks can shift too
 - ... but a precious resource for many HPC apps

Conclusion

- Use the intra-cluster routing to perform shifting
 - <u>Motivation</u>: floating-point
 - <u>Outcome</u>: ~30% reduction in area per operator
- Macro-cells address the major CAD challenges
 - We can route nets to pre-specified CLB inputs within a macro-cell
 - P&R treats macro-cells like soft IP
 - P&R cannot optimize across macro-cell boundaries
 - No negative impact on P&R results