GRVI Phalanx

A Massively Parallel RISC-V FPGA Accelerator

Jan Gray
jan@fpga.org
Introduction

• FPGA accelerators are hot
  – MSR Catapult. Intel += Altera. OpenPOWER + Xilinx

• FPGAs as computers
  – Massively parallel. Specialized. Connected
  – High throughput. Low latency. Low energy

• Great! But for two challenges
  – Software: app → ??? → HW. OpenCL! OpenCL?
  – Hardware: compose 100s of accelerators
    with 100 Gb/s networks and DRAM/HBM channels?
Phalanx: An Accelerator Accelerator

• Infrastructure to make it easier to
  – Run your application on an FPGA
  – Connect everything together
• Processor+accelerator clusters + NOC

• Acceleration requires an efficient processor core with OSS infrastructure
• !
• ... FPGA-efficient?
Soft Processor Area&Energy Efficiency

• Simpler, smaller processors $\rightarrow$ more processors
• **Jan’s Razor**: “In a CMP, **cut** inessential resources from each CPU, to maximize CPUs per die.”
• Share other FUs with your cluster

• Sweat every LUT
Austere RV32I Datapath – ~250 LUTs
GRVI ("Groovy")
Gray Research RISC-V RV*I

• Purpose: efficient parallel processing element.
• Scalar, 2-3 stage pipeline RV32I<sup>+-</sup> + MUL/H<sub>opt</sub>
  – 300-375 MHz (KU-2), 1.3-1.6<sup>?</sup> CPI
  – ~320 6-LUTs
  – ~1 “MIPS”/LUT
Modern FPGAs are Enormous (KU040)

240,000 LUTs ÷ 300 = 800 PEs? But 600 4 KB BRAMs?
Level 1: Clusters

- UltraScale: ~400 LUTs/BRAM
  - Two PEs can share an instruction BRAM (IRAM)
  - All PEs share one cluster data RAM (CRAM)

<table>
<thead>
<tr>
<th>BRAMs</th>
<th>PEs</th>
<th>IRAM</th>
<th>CRAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1I + 2D = 3</td>
<td>2</td>
<td>4 KB</td>
<td>8 KB</td>
</tr>
<tr>
<td>2I + 4D = 6</td>
<td>4</td>
<td>4 KB</td>
<td>16 KB</td>
</tr>
<tr>
<td>4I + 8D = 12</td>
<td>2</td>
<td>16 KB</td>
<td>32 KB</td>
</tr>
<tr>
<td>4I + 8D = 12</td>
<td>8</td>
<td>4 KB</td>
<td>32 KB</td>
</tr>
</tbody>
</table>
8 GRVI / 12 BRAM Cluster
600 BRAMs ÷ 12 BRAMs = 50 clusters

How to interconnect them?
5-port Virtual Channel Router?
Hoplite 2D Router

- Rethink FPGA NOC router architecture
  - No S/R/flits, no VCs, no buffering, no credits
  - Just move the bits!
  - Simple; frugal; wide; fast!
  - Default: deflecting dimension order routing

- **1%** of area×delay of FPGA-optimized VC routers
2D Directional Torus ‘Hoplite NOC’
Example Hoplite 2D Torus NOC

256b links @ 400 MHz = 100 Gb/s links; <5% of FPGA
Hoplite 2D Router

• Altera, Xilinx optimal area and delay
  – 1 LUT/bit of link width
  – D-FF \( \rightarrow \) (...) \( \rightarrow \) 1 LUT \( \rightarrow \) D-FF
  – 1b-1024b wide \( \rightarrow \) 1-400 Gb/s links, easy

• Will change FPGA design, IP, tools, devices
  – Everything is connected / site doesn’t matter much
10×5×8 = 400 GRVI Phalanx (KU040)

100% of BRAMS; 73% of LUTs
GRVI Phalanx: Hoplite NOC

<40 LUTs/processor; ~6% of FPGA LUTs
GRVI Phalanx
Very Preliminary Power Data

• ~250 mW/cluster = ~30 mW/core all in
• Will go up, will come down
400 RISC-Vs!

- $50 \times 32 \text{ KB} = 1.6 \text{ MB total CRAM}$
- $\leq 100,000 \text{ MIPS}, 600 \text{ GB/s CRAM BW}$
- Send/receive 32B/cluster/cycle, til NOC saturates

- Multicast Hoplite NOC (soon)
  - Load all clusters’ IRAM in 1K cycles = 3μs load kernel
  - Broadcast 100 Gb/s packet messages to all sites
2x2x8 = 32 GRVI Phalanx
Artix-7-35T. 40 Gb/s links. $35 Q1. ~$1/RISC-V core
Accelerated Parallel Programming Models (Aspirational)

- SPMD: OpenCL kernels, work group/cluster; ‘Gatling gun’ packet processing
- MIMD: task parallel models
- MP: streaming data through process networks
- Accelerated via
  - Custom GRVI and cluster function units
  - Custom memories, interconnects
  - Custom accelerators on CRAM, on NOC
Status

• Bringing up Thoth message passing system
• Next steps: crawl, walk, run
  – Debug/trace over NOC
  – Hoplite/AXI4 bridges: Zynq, DRAM, Ethernet, PCIe
  – Arria 10 FPUs!
  – OpenCL stack
  – Bridge to Chisel RISC-V infrastructure?
Accelerating FPGA Acceleration

• GRVI Phalanx is “All Programmable”
  – 100s of GRVI RISC-V soft processors and accelerators
  – Local shared memory, global message passing
  – Extreme bandwidth I/O
  – Accelerated parallel programming models

• All connected with a Hoplite NOC

Thanks!

Q&A?
BACKUP MATERIAL
The Software Problem

• Evolving application $\Rightarrow$ ??? $\Rightarrow$ FPGA?

• OpenCL!
  – Software defined data, compute, synch

• OpenCL?
  – Much parallel software is not OpenCL
  – FPGA is specialized to a kernel
  – Hours per design spin?
See Also

• Jan’s Razor (‘02)
  http://www.fpgacpu.org/log/mar02.html#020305

• Mapping CMPs to FPGAs (‘05)
  http://ramp.eecs.berkeley.edu/Publications/Mapping%20CMPs%20to%20Xilinx%20FPGAs.ppt

• The Past and Future of FPGA Soft Processors

• Hoplite
  http://fpga.org/hoplite
Austere RV32I Subsets and Hacks

- Config sans byte/halfword ld/st
- Config sans shifts – always use mul/mulh
- Sans perf counters
- Mul/mulh sans div
- Fmul/fadd sans fdiv

- Branch/jump delay slots. *No, really!*

1/5/2016
$8 \times 12 = 96$ Hoplite Phalanx 😊

http://www.militaer-wissen.de/griechische-phalanx/
GRVI Phalanx: 10×5 Cluster Folded Torus
GRVI Phalanx: GRVI PEs and BRAMs
About Gray Research LLC

Gray Research LLC [http://fpga.org] is a hardware and software development consulting firm located in Bellevue, WA, USA.

We specialize in design and implementation of solutions and IP for energy efficient, FPGA-optimized parallel compute accelerators, extreme bandwidth Hoplite NOC routers, soft processor cores, and related tools.

We are pleased to support the RISC-V Foundation as a means to boost collaboration and innovation in computer architecture research and practice.
Connecting the World