GRVI and 2GRVI HBM2 Phalanxes, U280
Add custom function units/cores/memories to suit 320 6 cycles
RV32I +
H
H
16 nm
C
H
cluster bandwidth
C
C
A
New 64b cluster interconnect, 64b 32B writes/32B reads
2019 Q2
(300 MHz)
U280
Nontrivial to access and transport 5 cycles
- 1 / cycle
H
H
C
C
H
IDs
4 GB HBM2 DRAM STACK
A
Never queue in Y compute offload at AXI bridges
Deconstruct PEs into minimal core plus cluster
C
C
C
C
400 / 300 MHz
Compose cores & accelerator(s), & send/receive 32 byte to a NoC
AXI
C
C
C
C
C
C
C
C
C
2015 Q4
Add ~60 design
C
C
C
C
C
C
C
C
C
C
H
optional† (+100 LUTs)
C
H
C
MS Catapult, Amazon AWS F1, Alibaba, Baidu,
typical but optional 3 stg: IF|DC|EX
HW: VU3xP in
C
C
C
C
C
C
C
C
C
C
C
V1 shortcomings
• 32b pointers poor for AWS F1 / big data / OpenCL kernels
• 32b accesses waste half 64b UltraRAM cluster bandwidth
• In-order parch stalls on loads = 5 cycles in an 8-Pe cluster
• Packed, congested, insuff. pipelining = low freq (300 MHz)
• DRAM (10s GB/s) uncompetitive vs. GPUs (100s GB/s)

2019: game changer: FPGAs ↔ HBM2 DRAM
• Xilinx UltraScale+ VU33P and Intel Stratix 10 MX families
• Xilinx: two HBM2 stacks; 32 AXI-HBM bridge/controllers
• AXI—MCS switch: any AXI port can access any controller
• Reads/writes up 32 x 256b x 450 MHz = 460 GBs

Let’s make it easy to use this bandwidth
• TB/s to BRAMs/ UltraRAMs, 10s GB/s to HBM2 DRAM

V2: redesign GRVI Phalanx for HBM FPGAs
• New latency tolerant RV64I PEs
• New 64b cluster interconnect, 64b UltraRAM banks
• New 32B/cycle deep pipeline NoC-AXI RDMA bridges
• Add PCIe XDMA mastering (1 AXI-HBM channel)
• Add many more NoC ring columns
• Add Fmax: floorplanning, pipelining, FIFOs: target 400+ MHz
• Add SDAccel-for-RTL shell if, OpenCL kernel runtime
• 15 cols x 256b x 400 MHz – peak 192+192 GB/s/R/W†
• T: work-in-progress / pending / current plan

Cluster: 0-8 PEs, 128 KB RAM, accel’s, router
• Compose cores & accelerator(s), & send/receive 32 byte messages via multiported banked cluster shared RAM

Soc: 15x15-3 array of clusters + HBMs + PCIe
• 15 columns x NoC-AXI RDMA bridge + 2xAXI-HBM bridge

2GRVI – a simple, latency tolerant RV64I PE
• 400 LUTs (sans shared barrel shifter), up to 550 MHz
• Register scoreboard: only stall on use of a busy register
• Out of order retirement; concurrent execution
• Call me save relogs, block copies: now 1 load/cycle
• 2 stg: DC [EX – 3 stg: IF/DC] [EX – 4 stg: IF/DC] [EX] [WB]
• 4 stage (superpipelining) has L=2 ALU – CPI ↑ 25%
• Plan: further tolerate latency with two hardware threads

FPGA soft processor area and energy efficiency
• Simpler CPUs -- more CPUs -- more memory parallelism
• Deconstruct PEs into minimal core plus cluster-shared concurrent FUs: shifts, mul, custom FUs, memory ports (!)

But two hard problems
• Software: Porting & maintaining workload as accelerator
• Hardware: Composite 100s of cores, 100G NICs, many DRAM/HBM channels, with easy timing closure

Mission: GRVI Phalanx FPGA accelerator kit
• GRVI: FPGA-efficient RISC-V processing element cores
• Phalanx: array of clusters of PE, SRAMs, accelerators.
• Hoplite NoC: FPGA-optimal directional 2D torus soft NoC
• Local shared memory, global message passing, PGAS

Software-first, software-mostly accelerators
• Run your C+/OpenCL kernels on 100s of soft processors
• Add custom function units/cores/memories to suit
• More 10 sec reccompiles, fewer 5 hour synthesize/place/route
• Complements high level synthesis & OpenCL—FPGA flows

2017: V1: 1680 core GRVI Phalanx in a U29P

GRVI: A Massively Parallel RISC-V® FPGA Processor Framework
A 1680-core, 26 MB SRAM Parallel Processor Overlay on Xilinx UltraScale+ U29P
Jan Gray | Gray Research LLC | Bellevue, WA | jan@fpga.org | http://fpga.org