GRVI Phalanx

A Massively Parallel RISC-V FPGA Accelerator 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
- 

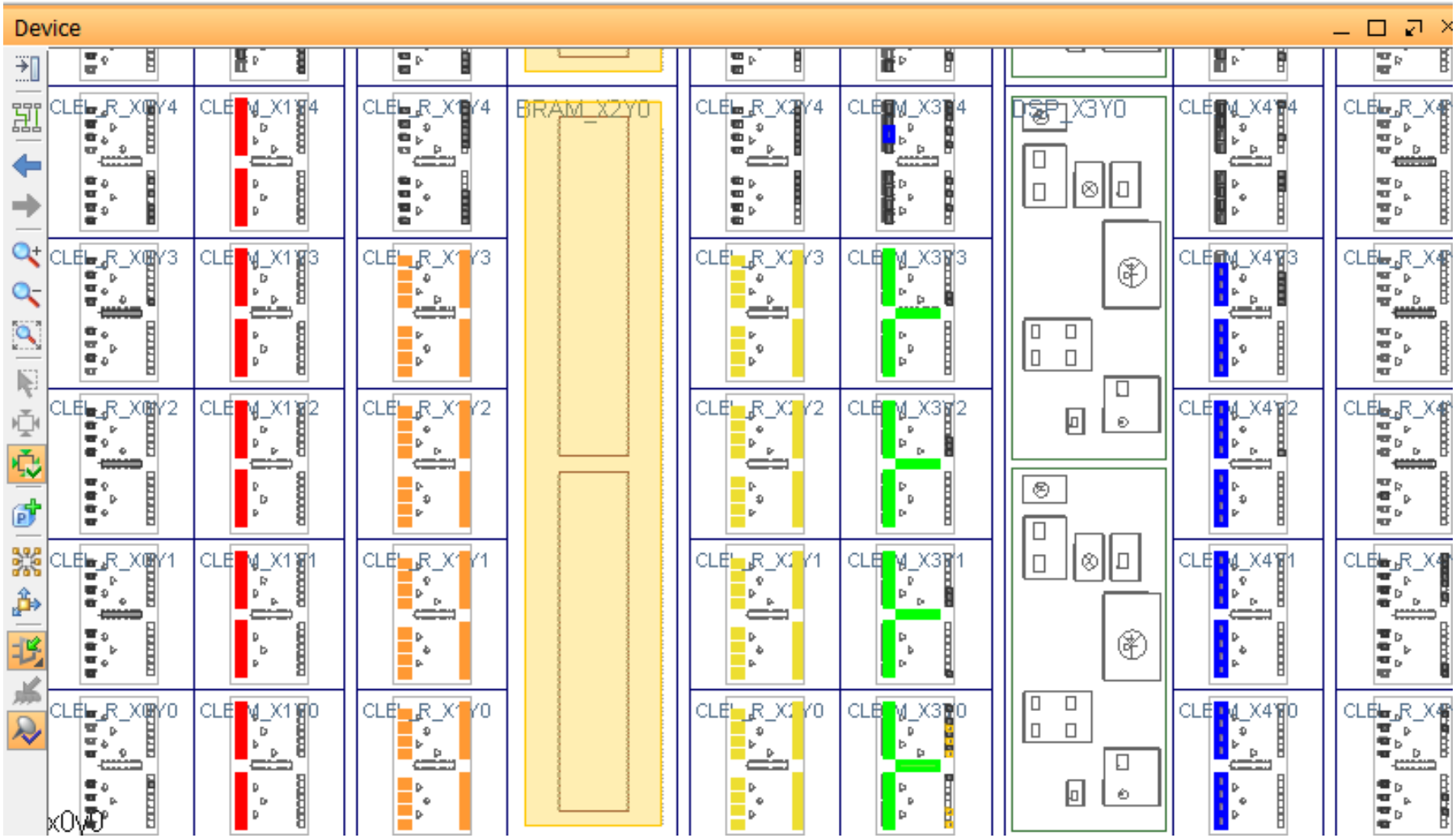
!
- … FPGA-efficient?

Soft Processor Area&Energy Efficiency

- Simpler, smaller processors → 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^{+-} + MUL/H^{opt}
  - 300-375 MHz (KU-2), 1.3-1.6$^{?}$ CPI
  - ~320 6-LUTs
  - **~1 "MIPS"/LUT**

Modern FPGAs are Enormous (KU040)

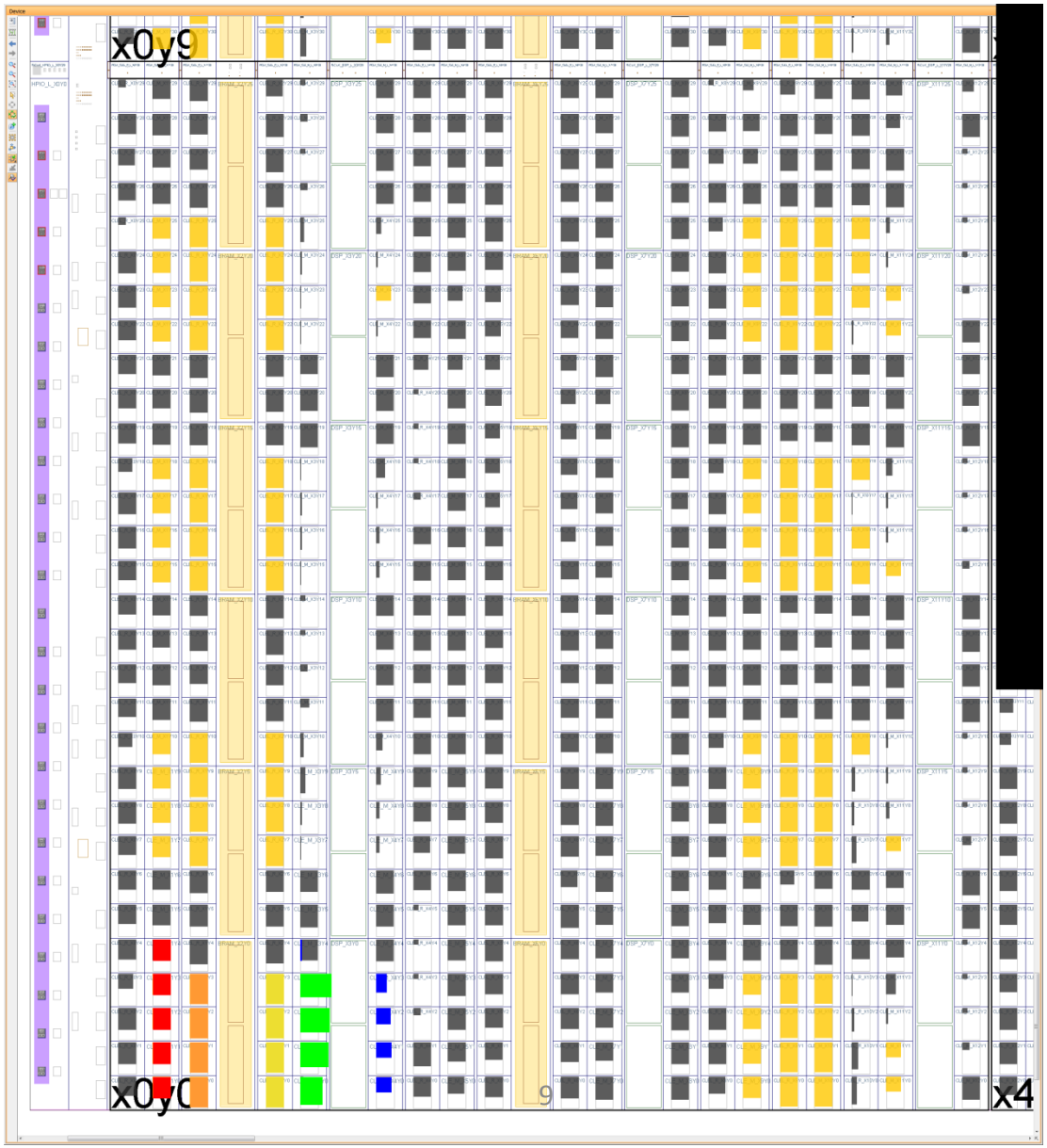
240,000 LUTs ÷ 300 = 800 PEs? But 600 4 KB BRAMs?

Level 1: Clusters

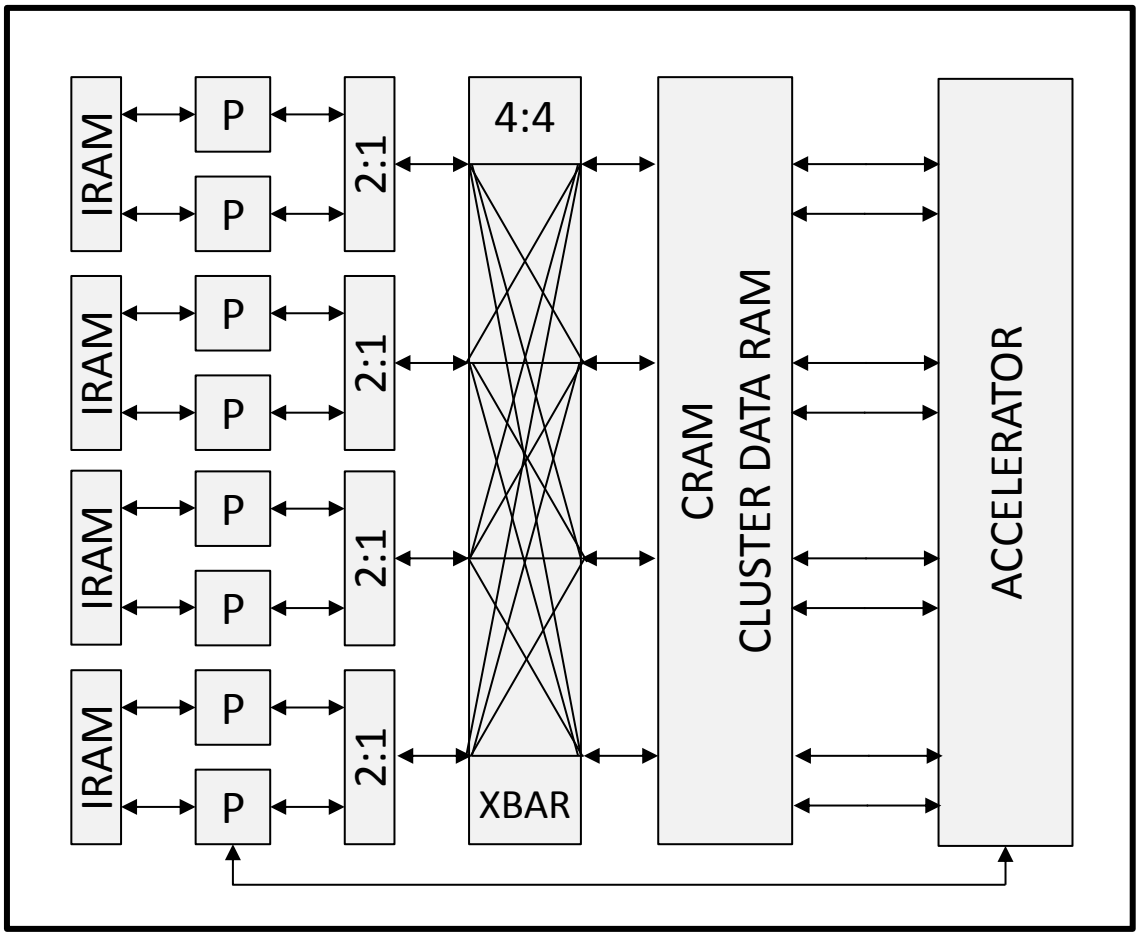
- UltraScale: ~400 LUTs/BRAM
  - Two PEs can share an instruction BRAM (IRAM)
  - All PEs share one cluster data RAM (CRAM)

BRAMs	PEs	IRAM	CRAM
1I + 2D = 3	2	4 KB	8 KB
2I + 4D = 6	4	4 KB	16 KB
4I + 8D = 12	2	16 KB	32 KB
4I + 8D = 12	**8**	**4 KB**	**32 KB**

8 GRVI / 12 BRAM Cluster



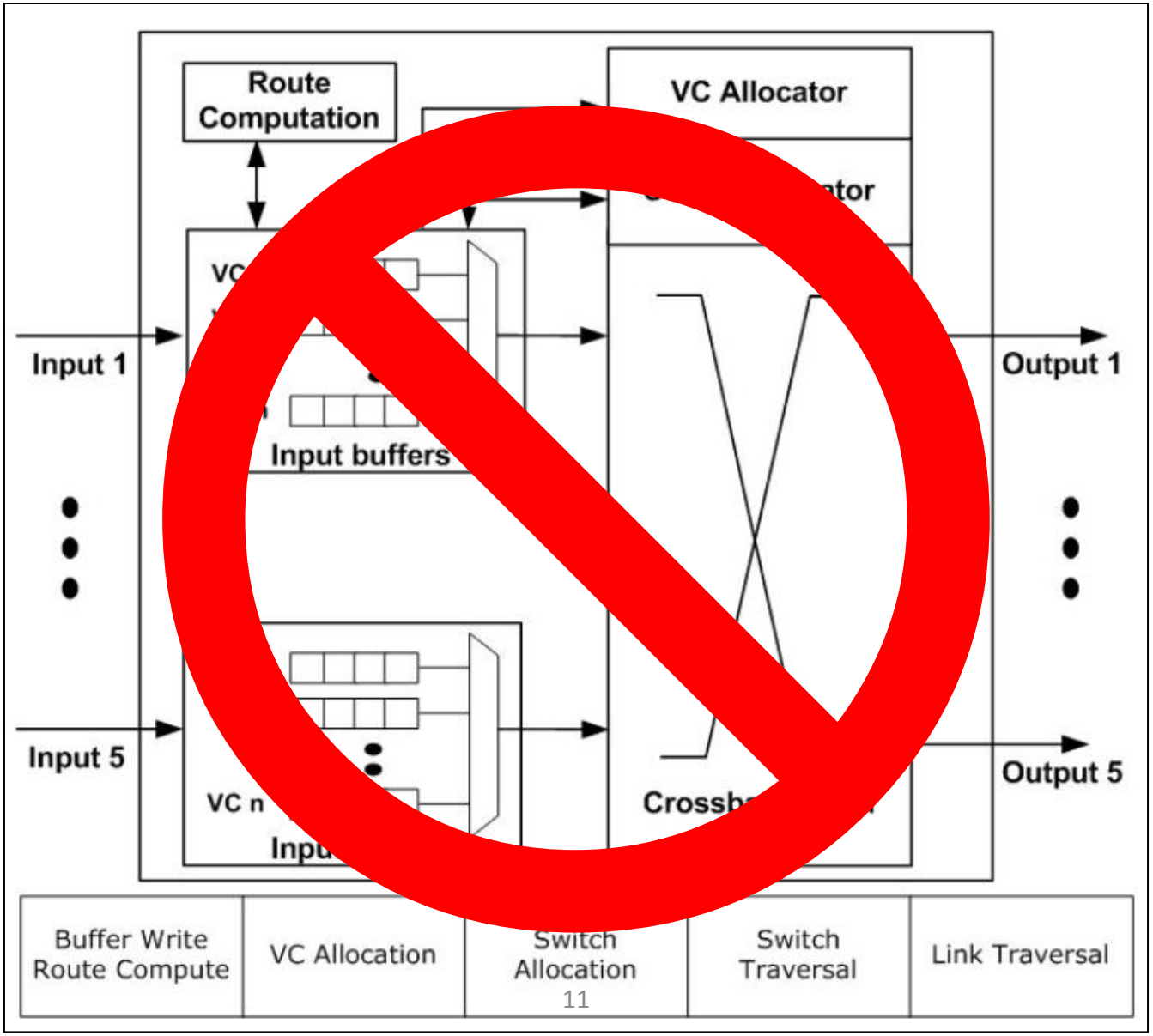
8 GRVI Cluster



600 BRAMs ÷ 12 BRAMs = 50 clusters

How to interconnect them?

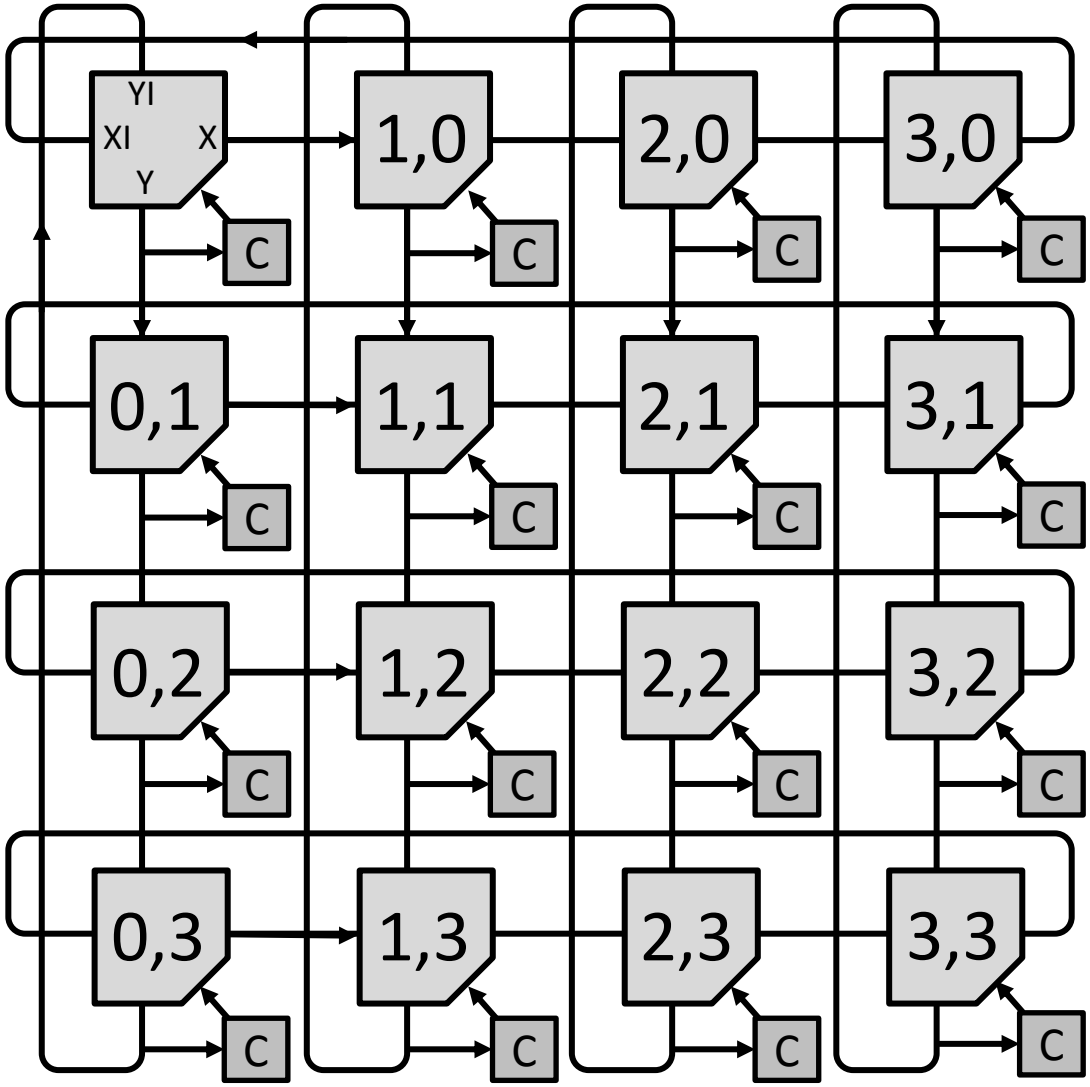
5-port Virtual Channel Router?



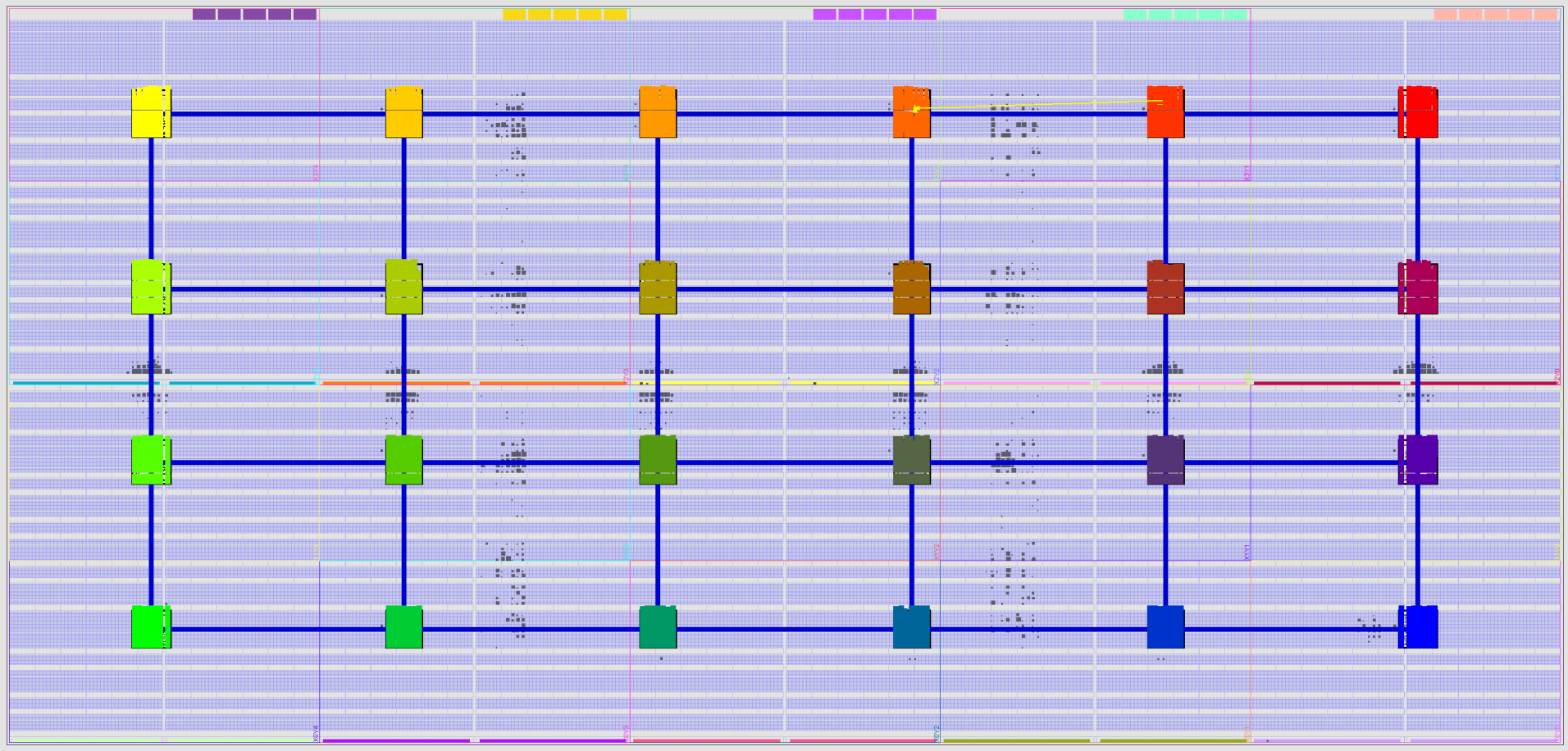
Hoplite 2D Router

- 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



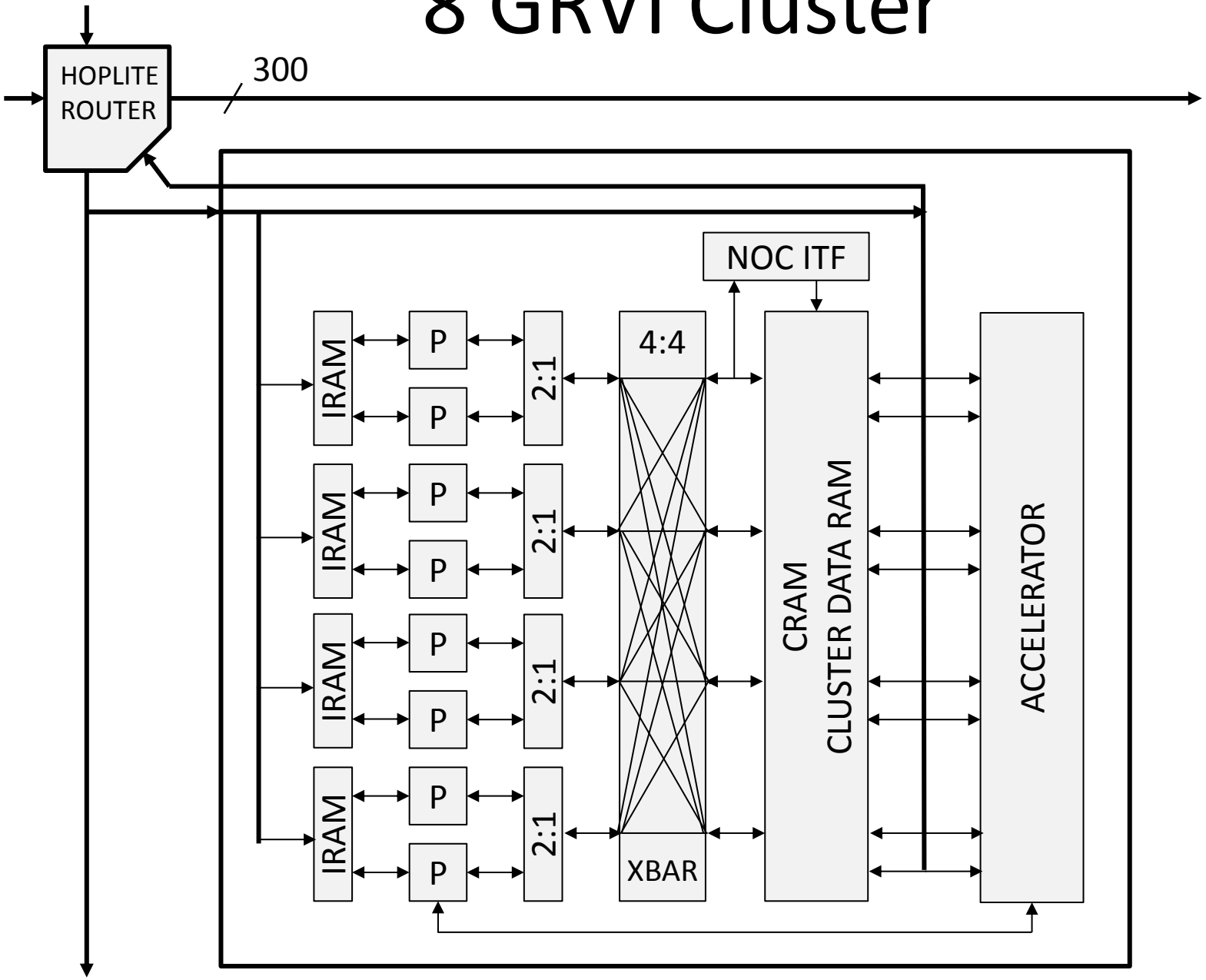
256b links @ 400 MHz = 100 Gb/s links; <5% of FPGA

Hoplite 2D Router

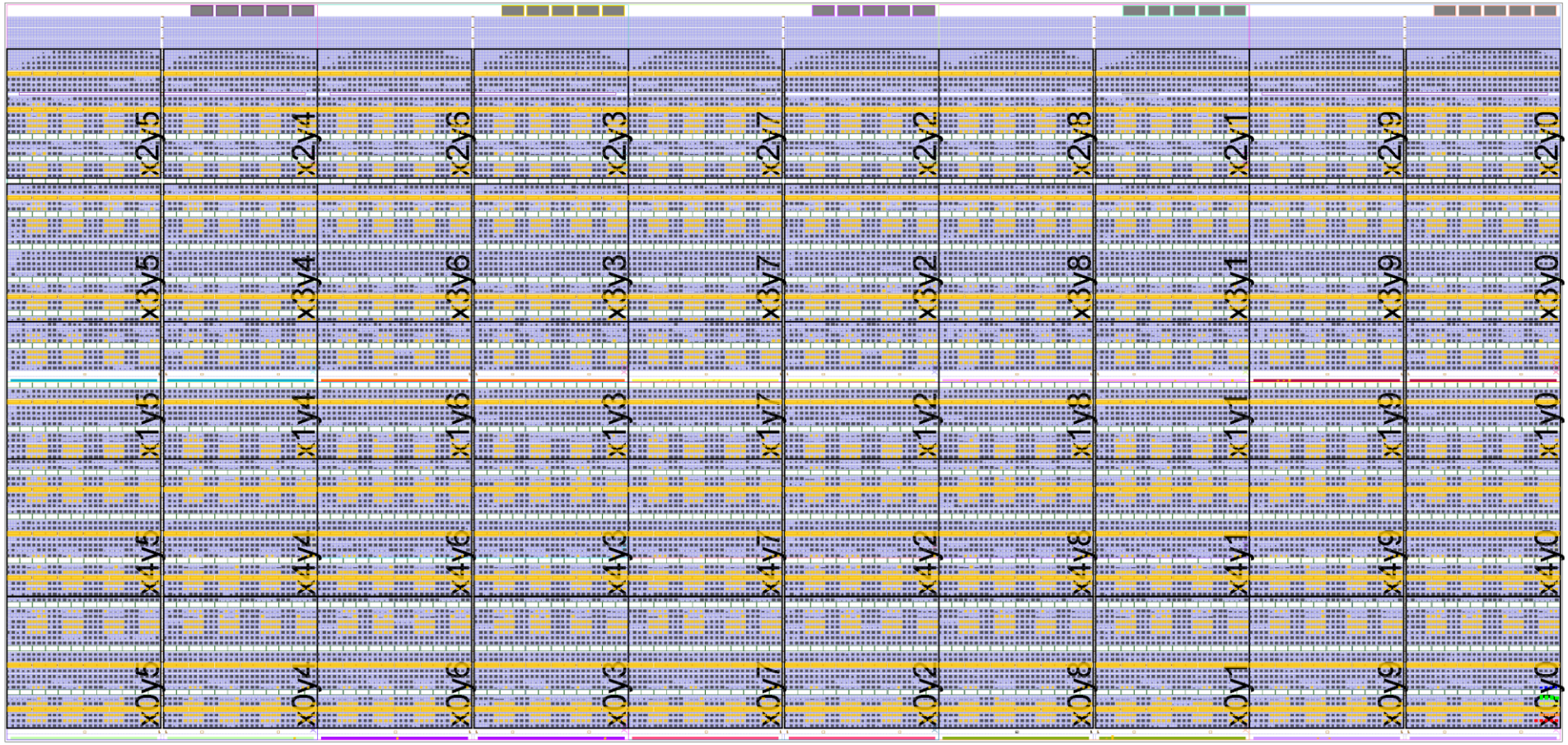
- Altera, Xilinx *optimal* area and delay
  - 1 LUT/bit of link width
  - D-FF → *(…wire…)* → 1 LUT → D-FF
  - 1b-1024b wide → 1-400 Gb/s links, easy

- Will change FPGA design, IP, tools, devices
  - Everything is connected / site doesn't matter much

8 GRVI Cluster

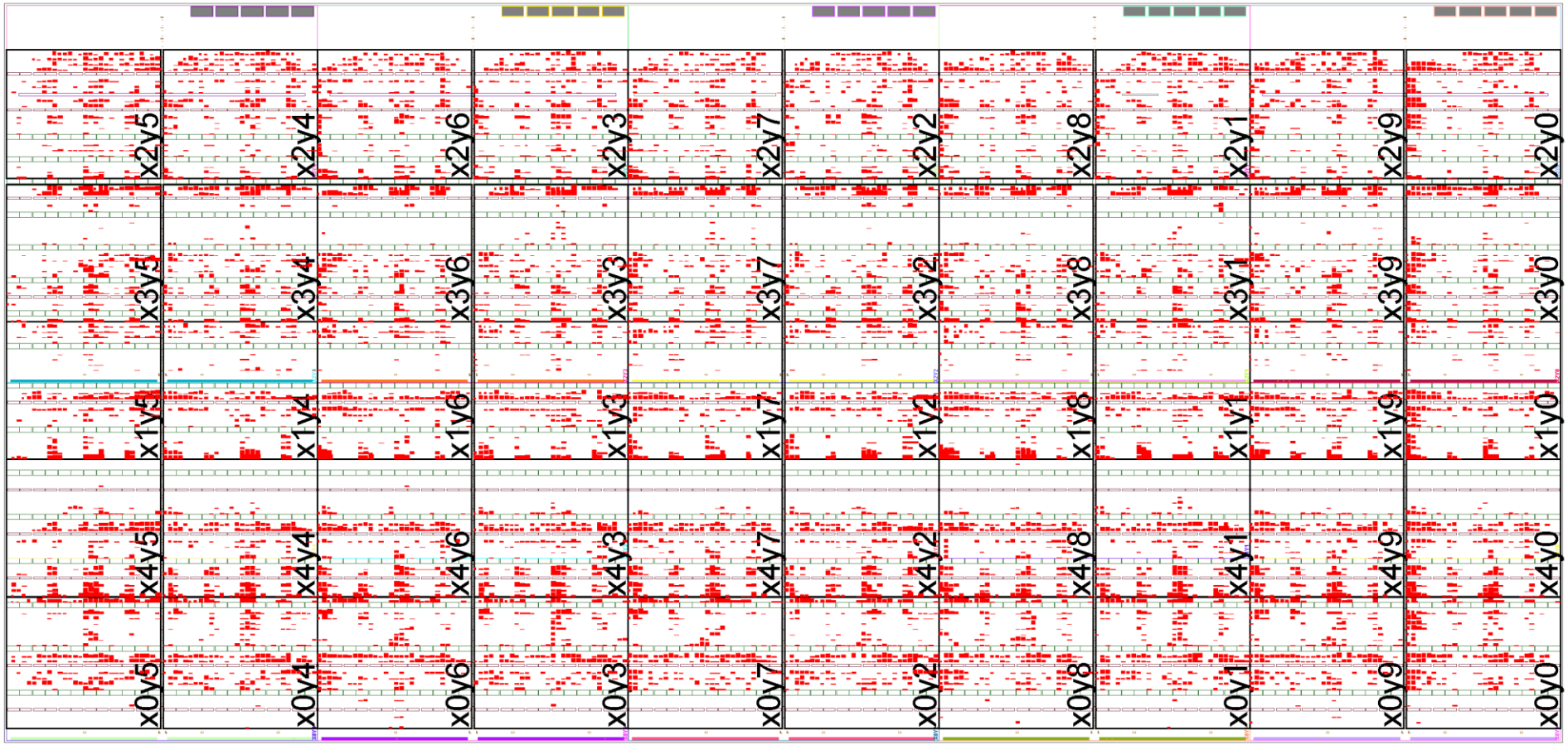


10×5×8 = 400 GRVI Phalanx (KU040)



100% of BRAMS; 73% of LUTs

GRVI Phalanx: Hoplite NOC



<40 LUTs/processor; ~6% of FPGA LUTs

GRVI Phalanx
Very Preliminary Power Data

```
Temperature = 52.75 C  Min = 51.31 C  Max = 54.17 C
-------------------------------------------------------
                                          MIN       MAX
          Power     Voltage    Current   Current   Current
  VCCINT: 11.18 W    0.95 V   11.77 A   11.21 A   11.77 A
  VCCAUX:  0.43 W    1.80 V    0.24 A    0.23 A    0.24 A
 VCCBRAM:  0.26 W    0.95 V    0.28 A    0.26 A    0.28 A
  VCC1V8:  0.00 W    1.80 V    0.00 A    0.00 A    0.01 A
VADJ_1V8:  0.05 W    1.80 V    0.03 A    0.00 A    0.03 A
  VCC1V2:  0.08 W    1.20 V    0.06 A    0.06 A    0.07 A
 MGTAVCC:  0.11 W    1.00 V    0.11 A    0.11 A    0.18 A
 MGTAVTT:  0.06 W    1.20 V    0.05 A    0.03 A    0.07 A
```

- ~250 mW/cluster = ~30 mW/core all in
- *Will go up, will come down*

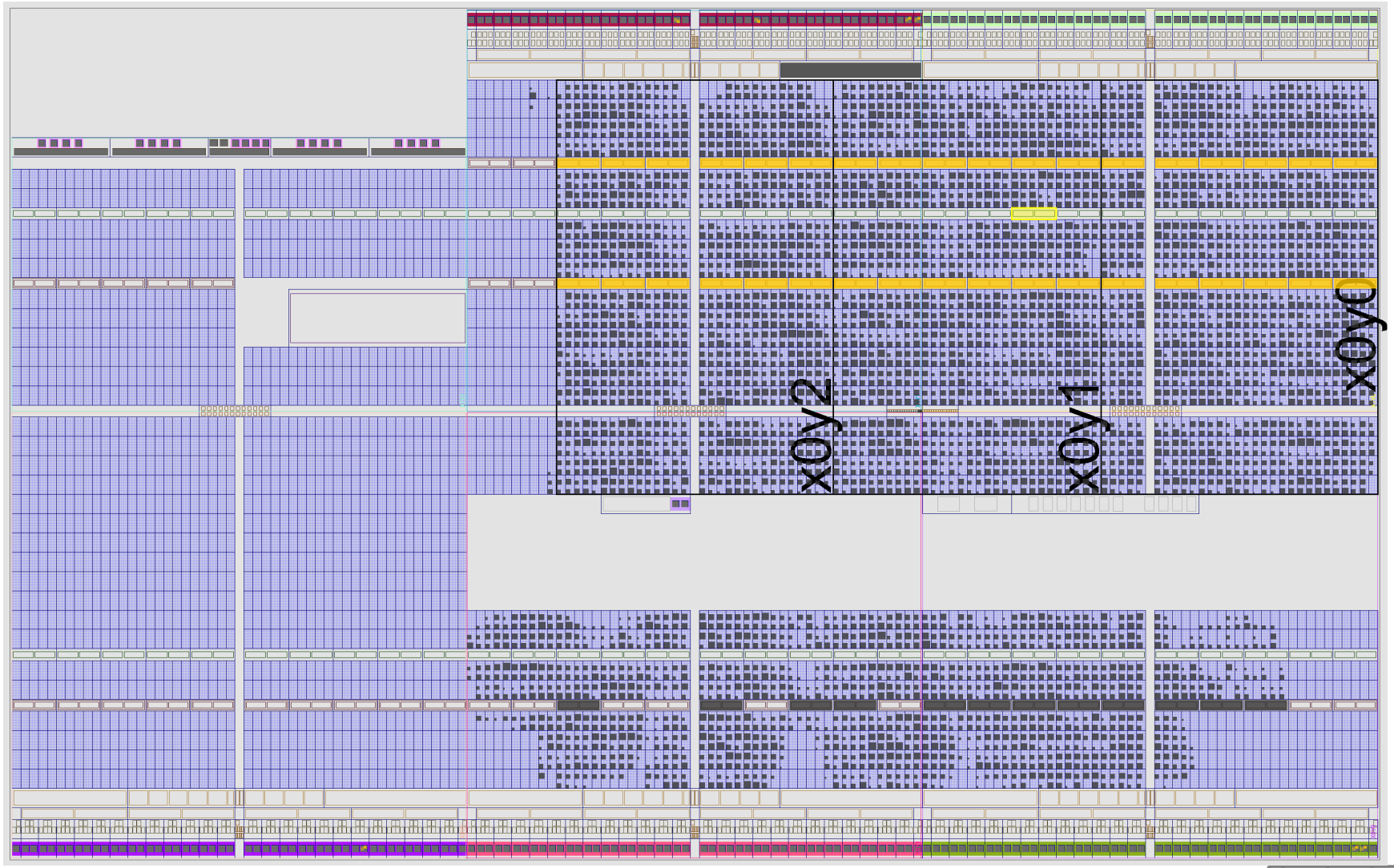
400 RISC-Vs!

- 50×32 KB = 1.6 MB total CRAM
- ≤ 100,000 MIPS, 600 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!

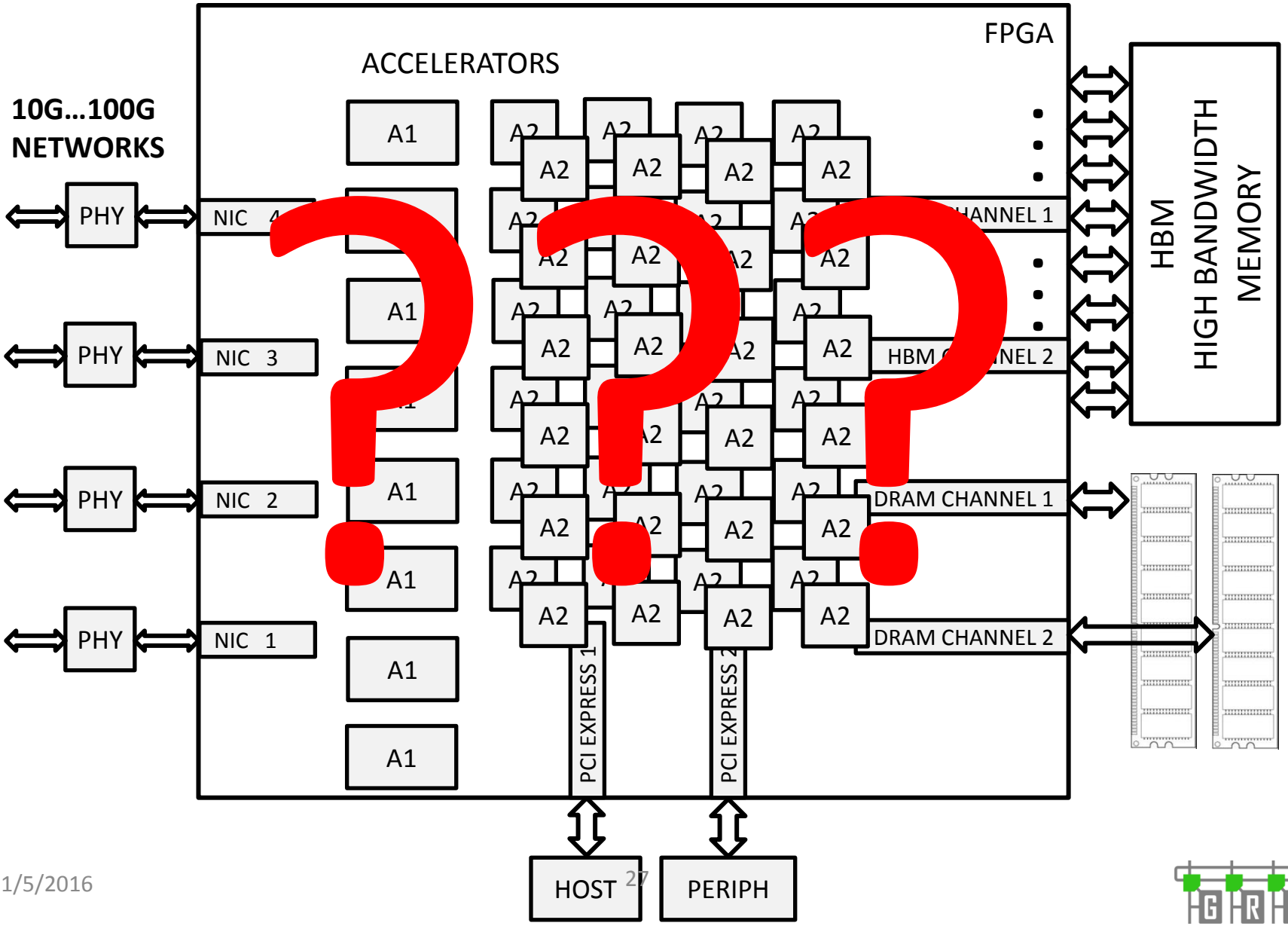
Q&A?

BACKUP MATERIAL

The Software Problem

- Evolving application → ??? → FPGA?
- OpenCL!
  - Software defined data, compute, synch
- OpenCL?
  - Much parallel software is not OpenCL
  - FPGA is specialized to a kernel
  - Hours per design spin?

Hardware: Physical Implementation



See Also

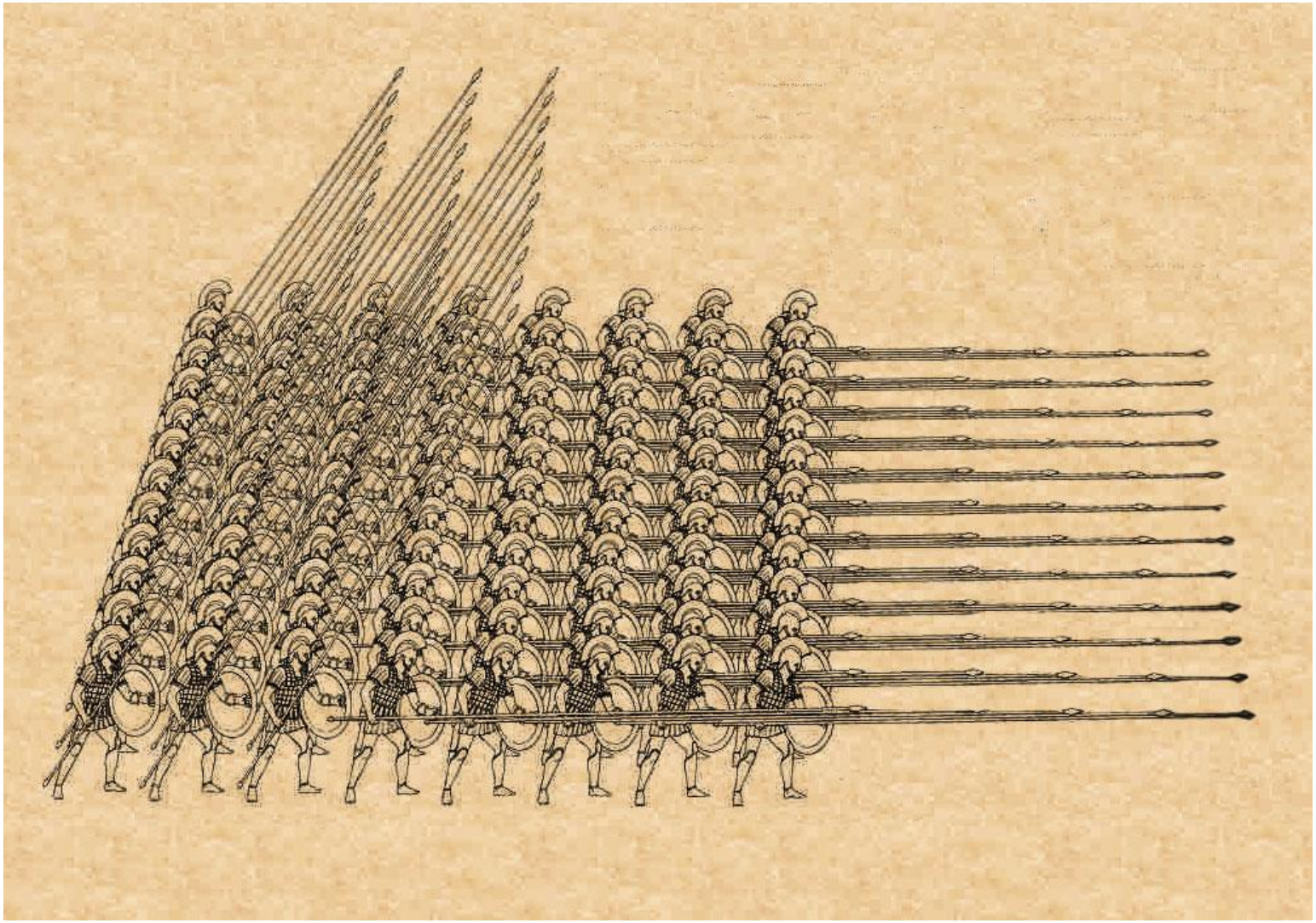
- Jan's Razor ('02)
  http://www.fpgacpu.org/log/mar02.html#020305
- Mapping CMPs to FPGAs ('05)
  http://ramp.eecs.berkeley.edu/Publications/Mapping%20CMPs%20to%20Xilinx%20FPGAs.ppt
- The Past and Future of FPGA Soft Processors
  http://fpga.org/2014/12/31/the-past-and-future-of-fpga-soft-processors
- 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

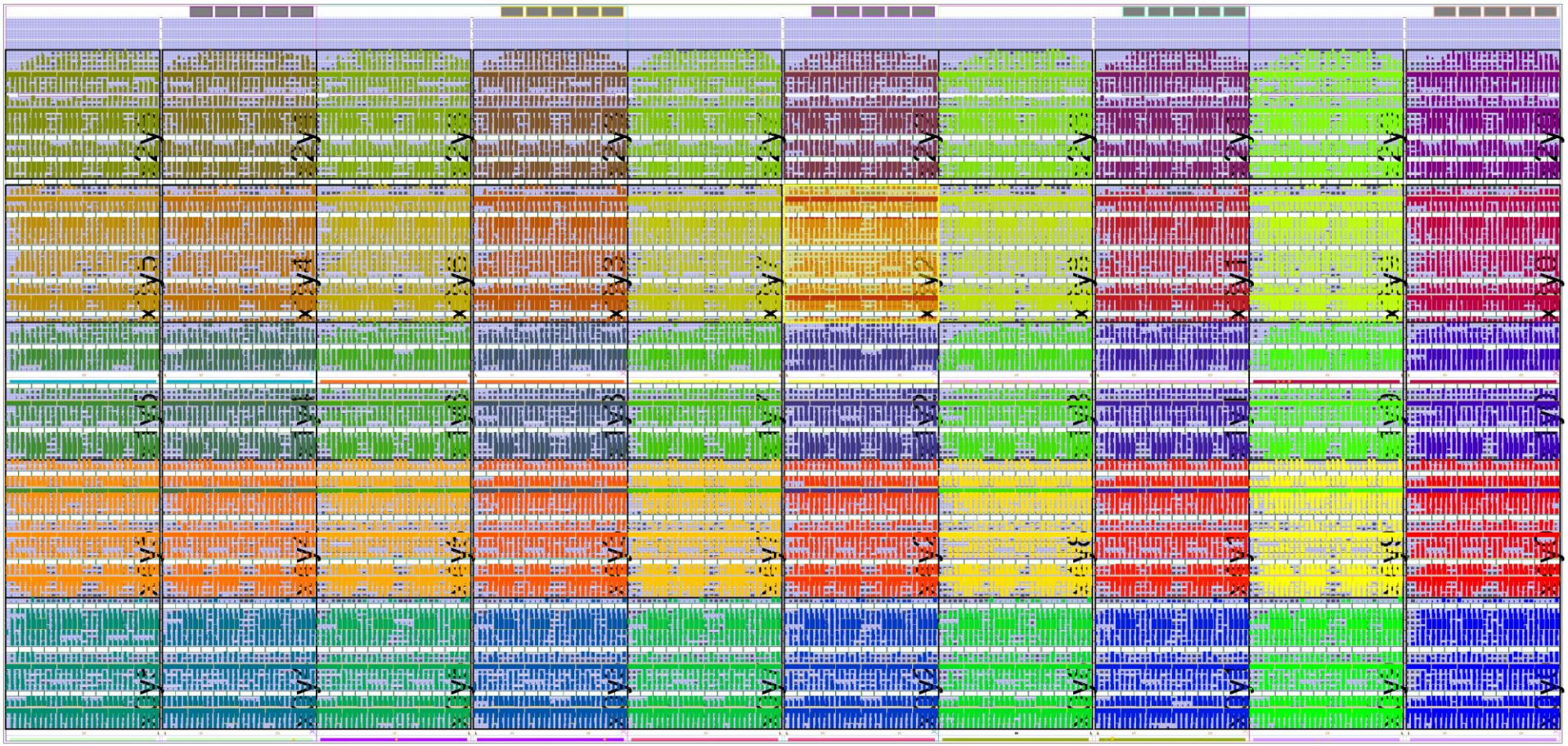
- Branch/jump delay slots. *No, really!*

8×12 = 96 Hoplite Phalanx ☺

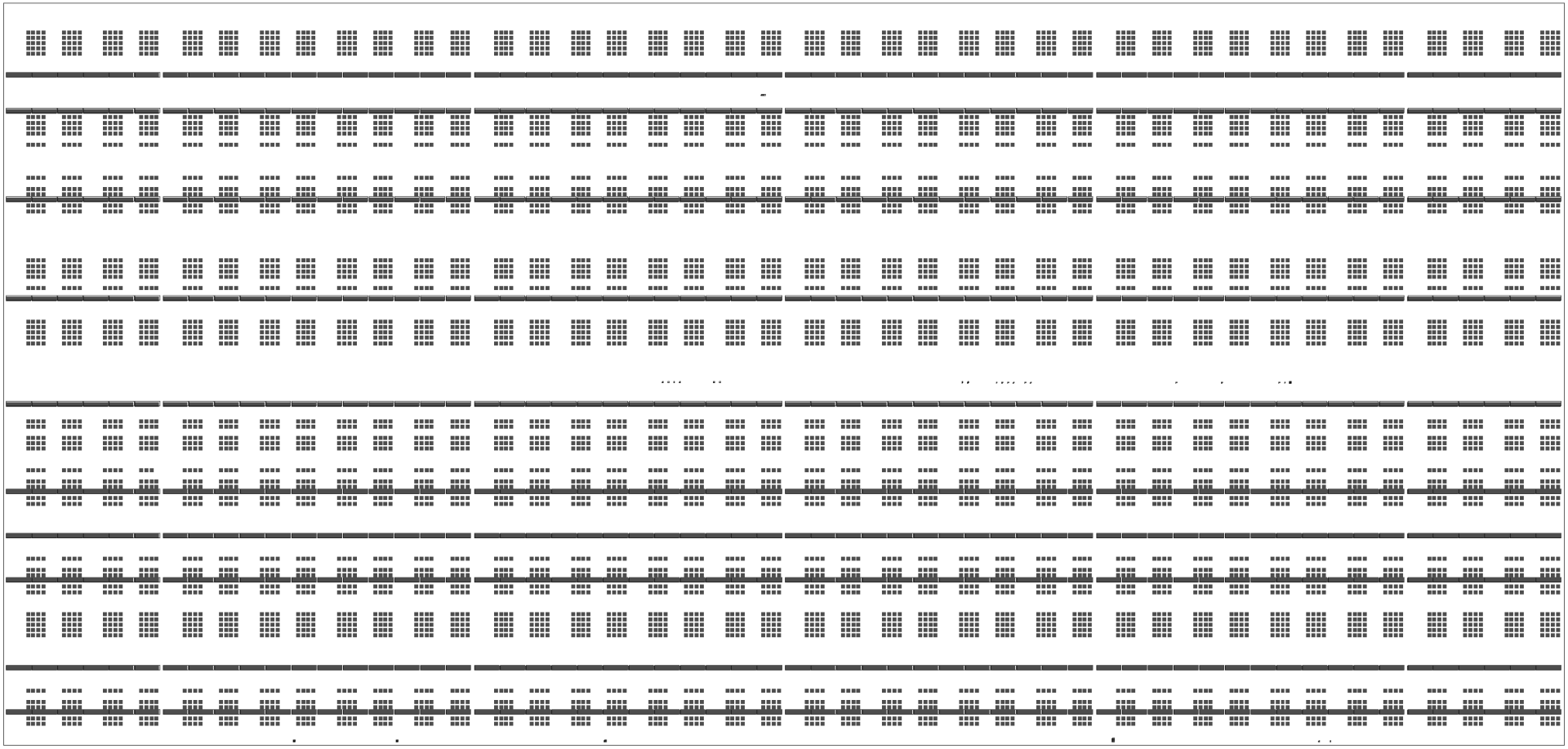


http://www.militaer-wissen.de/griechische-phalanx/

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