Emulating Future HPC SoC Architectures Using RISC-V

Farzad Fatollahi-Fard, Dave Donofrio, John Shalf
Lawrence Berkeley National Lab

RISC-V Workshop
January 5, 2016 – Redwood City, CA
Should HPC Take Inspiration from the Embedded Market?

- Have most of the IP and experience with low-power technology
  - Have sophisticated tools for rapid turn-around of designs
- Vibrant commodity market in IP components
  - Change your notion of “commodity”!
  - It’s commodity IP on the chip (not the chip itself!)
- Design validation / verification dominate cost
  - Another benefit of COTS IP
Building an SoC for HPC

Is this a good idea?

- Consumer market dominates PC and server market
  - Smartphone and tablets are in control
  - Huge investments in IP, design practices, etc.

- HPC is power limited (delivered performance/watt)
  - Need better computational efficiency and lower power with greater parallelism
  - Embedded has always been driven by max performance/watt (max battery life) and minimizing cost

- HPC and embedded requirements are now aligned
  - ...and now we have a very large commodity ecosystem

- Why not leverage technologies for the embedded and consumer for HPC?
Looking back…

A previous HPC system design based on semi-custom SoCs
Green Wave
2009-2012

- Seismic imaging used extensively by oil and gas industry
  - Dominant method is RTM (Reverse Time Migration)
- RTM models acoustic wave propagation through rock strata using explicit PDE solve for elastic equation in 3D
  - High order (8th or more) stencils
  - High computational intensity
Green Wave Design Study
Seismic Imaging

Performance

Energy Efficiency

Embedded Design library

U.S. Department of Energy
Office of Science
Embedded SoC Efficiency Competitive with cutting-edge designs

Fermi without host

MPonts/s/W

NHM  Fermi  GW

8th Order

12th Order

8x  4x

7.6x  3.5x
Green Wave Chip Block Diagram

- 12 x 12 2D on-chip torus network
- 676 Compute cores (500 in compute clusters, 176 in peripheral clusters)
- 33 Supervisory cores
- 1 PCI express interface
- 8 Hybrid Memory Cube (HMC) interfaces
- 1 Flash controller
- 1 1000BaseT Ethernet controller

It is not anticipated that all cores will be utilized – some are spares for yield enhancement.
Building an SoC from IP Logic Blocks

It’s Legos with a some extra integration and verification cost

Processor Core (ARM, Tensilica, RISC-V)
With extra “options” like DP FPU, ECC

OpenSoC Fabric (on-chip network)
(currently proprietary ARM or Arteris)

DDR memory controller
(Denali/Cadence, SiCreations)
+ Phy & Programmable PLL

PCle Gen3 Root complex

Integrated FLASH Controller

10GigE or IB DDR 4x Channel

IO
Hierarchical Power Costs

Data Movement is the Dominant Power Cost

- **6 pJ**
  - Cost to move data 1 mm on-chip

- **100 pJ**
  - Typical cost of a single floating point operation

- **120 pJ**
  - Cost to move data 20 mm on chip

- **250 pJ**
  - Cost to move off-chip, but stay within the package (SMP)

- **2000 pJ**
  - Cost to move data off chip into DRAM

- **~2500 pJ**
  - Cost to move data off chip to a neighboring node
Other Parameters Impact Performance
For Example, Topology…

Network topology can greatly influence application performance

An analysis of on-chip interconnection networks for large-scale chip multiprocessors
ACM Transactions on computer architecture and code optimization (TACO), April 2010
Our SoC for HPC System
A Point Design on an FPGA

- Z-Scale processors connected in a Concentrated Mesh
- 4 Z-scale processors
- 2x2 Concentrated mesh with 2 virtual channels
- Designed for Area Efficiency on the FPGA
Z-Scale
Tiny 32-bit RISC-V System

- A tiny 32-bit 3-stage RISC-V core generator suited for microcontrollers and embedded systems
- Z-scale is designed to talk to AHB-Lite buses
- Z-scale generator also generates the interconnect between core and devices
  - Includes buses, slave muxes, and crossbars
Top-Level Z-Scale Configuration
OpenSoC Fabric
An Open-Source, Flexible, Parameterized, NoC Generator

- Written in Chisel
- Dimensions, topology, VCs all configurable
- AHB Endpoints available now
  - AXI in development
- Fast functional C++ model for functional validation
- Verilog based description for FPGA or ASIC
  - Synthesis path enables accurate power / energy modeling
Top Level Diagram
Functional Hierarchy
Top Level Modules

- Stitches routers together
- Assigns routers individual ID
- Assigns Routing Function to routers
- Connections Injection and Ejection Queues for network endpoints
OpenSoC Top Level Modules

- Created and connected by Topology module
- Instantiates and connects:
  - Routing Function
  - Allocators
  - Switch
- Pipelined
  - 3 stage pipeline for Wormhole
  - 4 stage pipeline for VCs
  - Includes state storage for each sub-module
- Connects to Injection / Ejection Queues
OpenSoC Top Level Modules

VC Router

Routing Functions
- Function
- Consume Head Flits
- Assign range of VCs
- Update Register file with result

Register Files per VC for Head Flits

Head Flits passed to allocator

Input FIFO per VC

Input / VC Control

VC Allocator
- Arb
- Assigns VC from range specified by routing function to a single head flit

Switch Allocator
- Arb
- Per VC Credit passed back to allocator

Routing Function
- Consume Head Flits
- Assign range of VCs
- Update Register file with result

Switch Control

Allocation

Router State Machine
- Router
- Concentration
- Router N

Single Output Reg

Concentration

Router

Per VC Credit passed back to allocator
Configuration options
A few of the current run time configuration parameters

- **Network Parameters**
  - Dimension
  - Routers per dimension
  - Concentration
  - Virtual Channels
  - Topology
  - Queue depths
  - Routing Function

- **Packet / Flit Parameters**
  - Flit widths
  - Packet types / lengths

- **Testing Parameters**
  - Pattern
    - Neighbor, random, tornado, etc
  - Injection Rate

Highly modular architecture supports FUB replacement
Demo Setup

HPC SoC Design Versus Commodity HPC Design
Demo Results
Exchanging of Cache Lines Versus Message Passing

Inter-Thread Latency

Cycles

Remote Exchange

Local Exchange

RISCV-SoC
x86
More Information

http://www.codexhpc.org

http://opensocfabric.org