Open Source HPC Hardware

John Shalf

David Donofrio, Farzad Fatollahi-Fard

Lawrence Berkeley National Laboratory

CCS/LBL Collaboration Meeting

University of Tsukuba, March 5, 2018

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY | Office of Science

Technology Scaling Trends

Exascale in 2021… and then what?

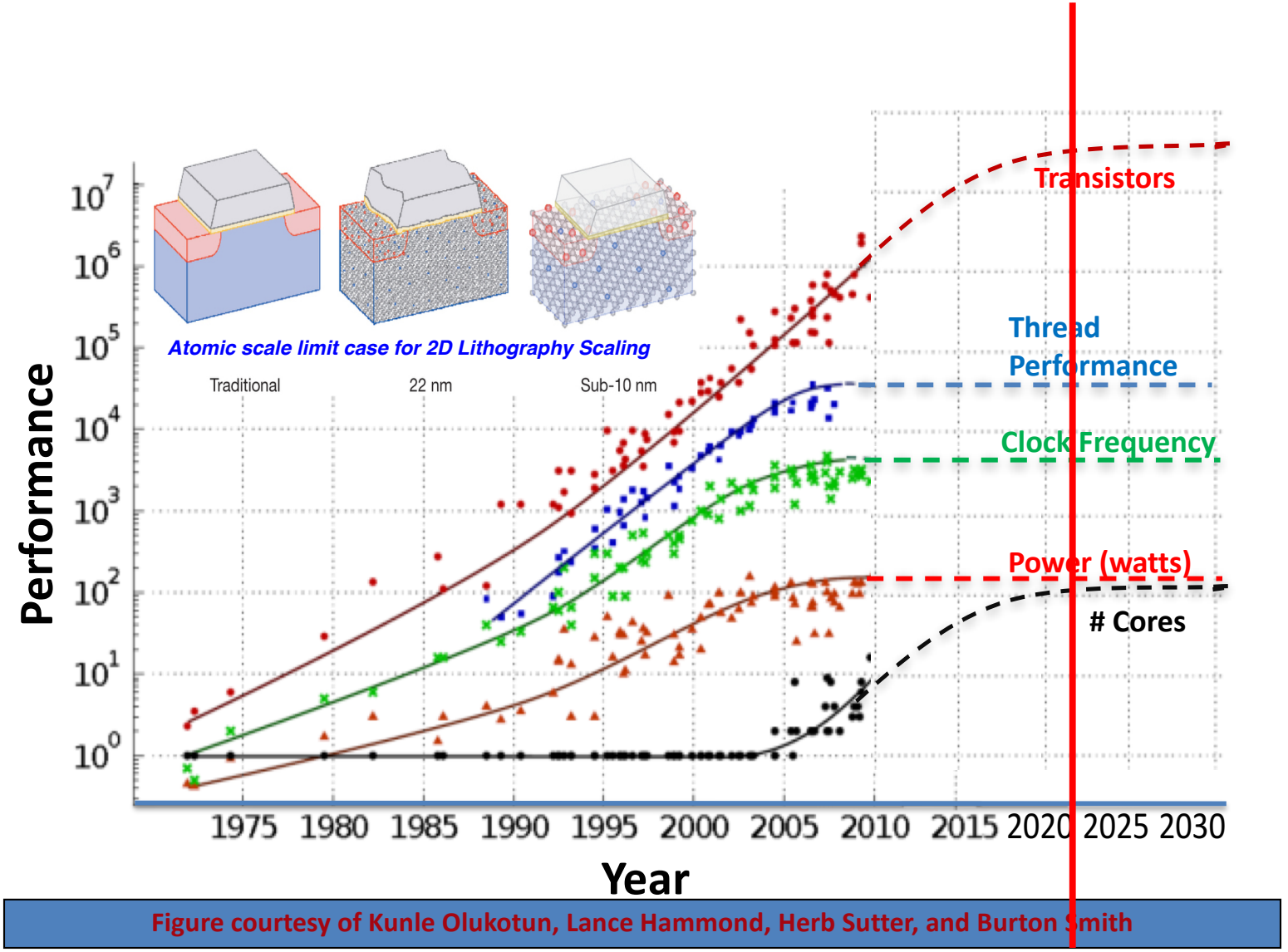
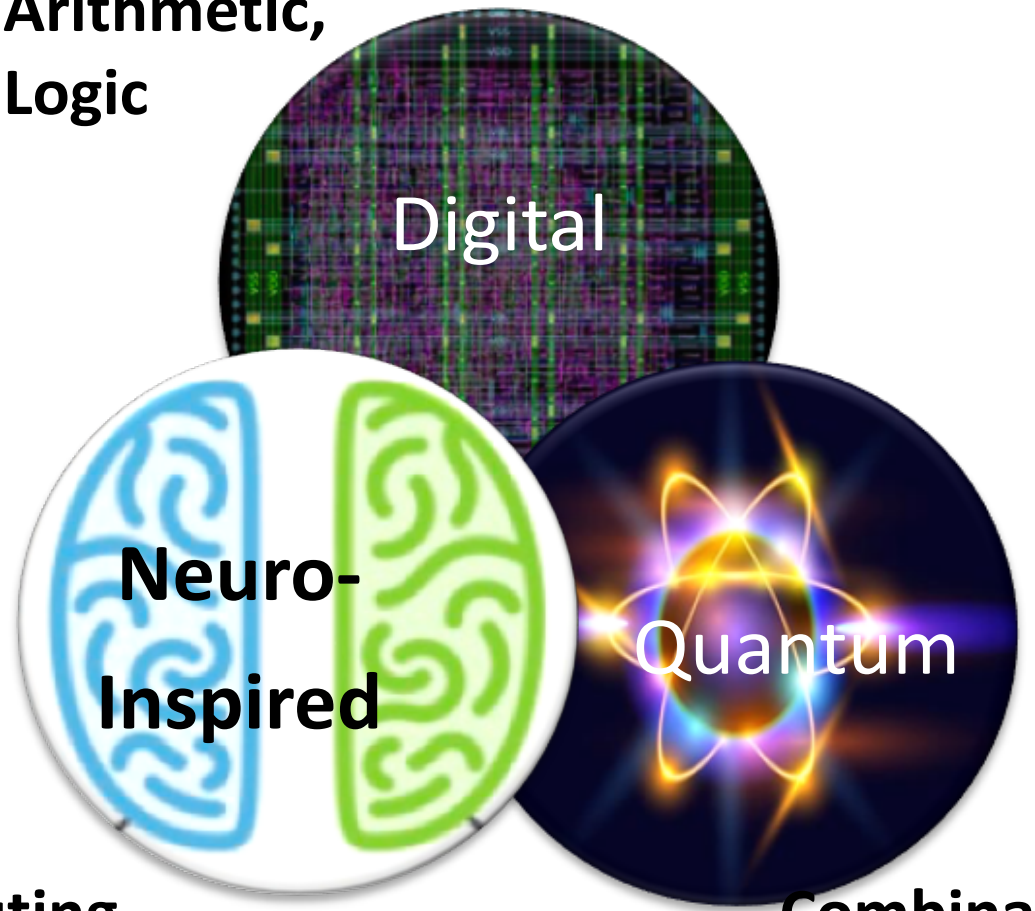


Figure courtesy of Kunle Olukotun, Lance Hammond, Herb Sutter, and Burton Smith

Beyond Moore Computing Taxonomy

Symbolic Computation, Arithmetic, Logic



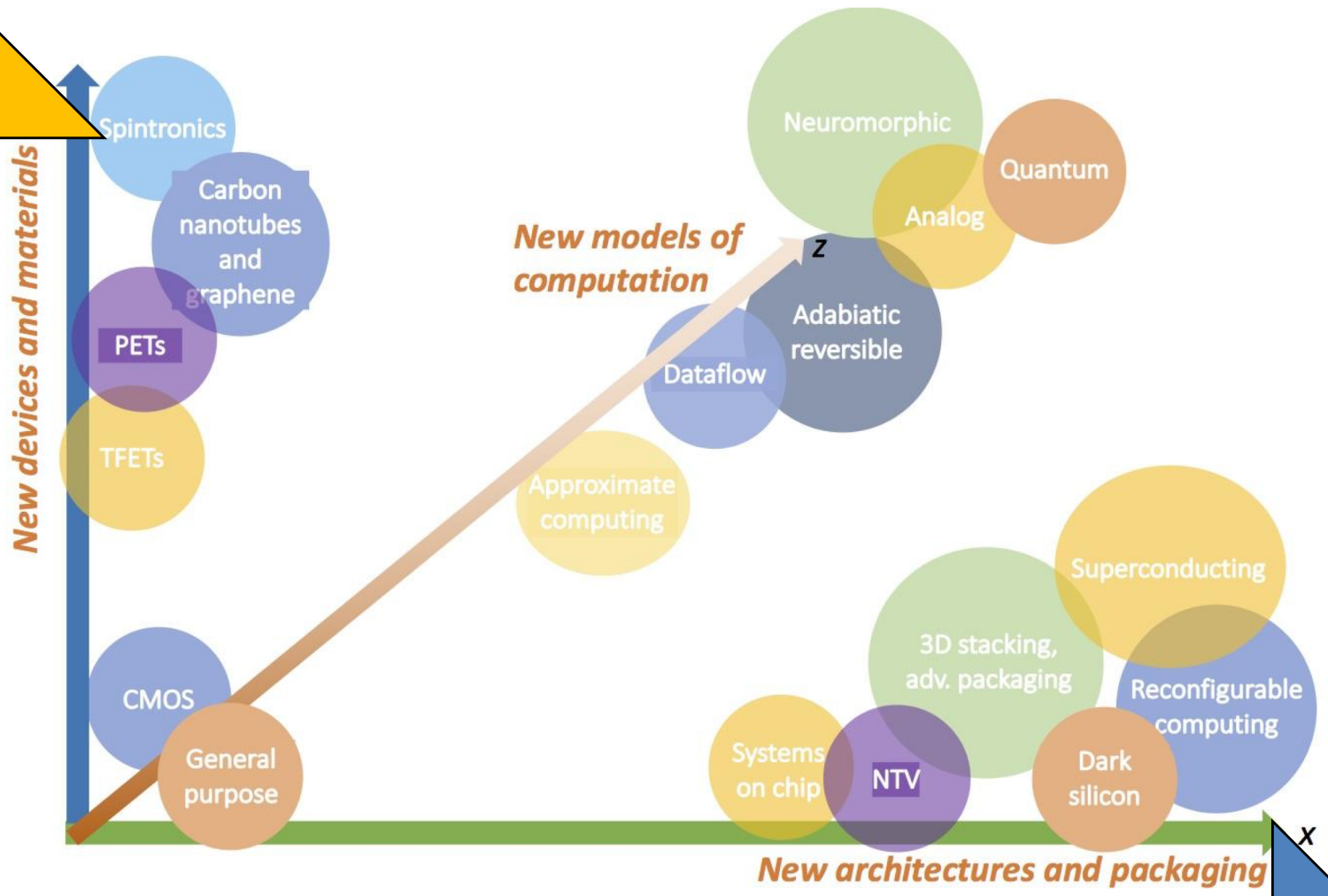
Cognitive Computing, Pattern Recognition

Combinatorial/NP, Annealing/Optimization, Simulated Atoms

COMPUTER ARCHITECTURE LABORATORY
EXASCALE DESIGN SPACE EXPLORATION

Future of Computing

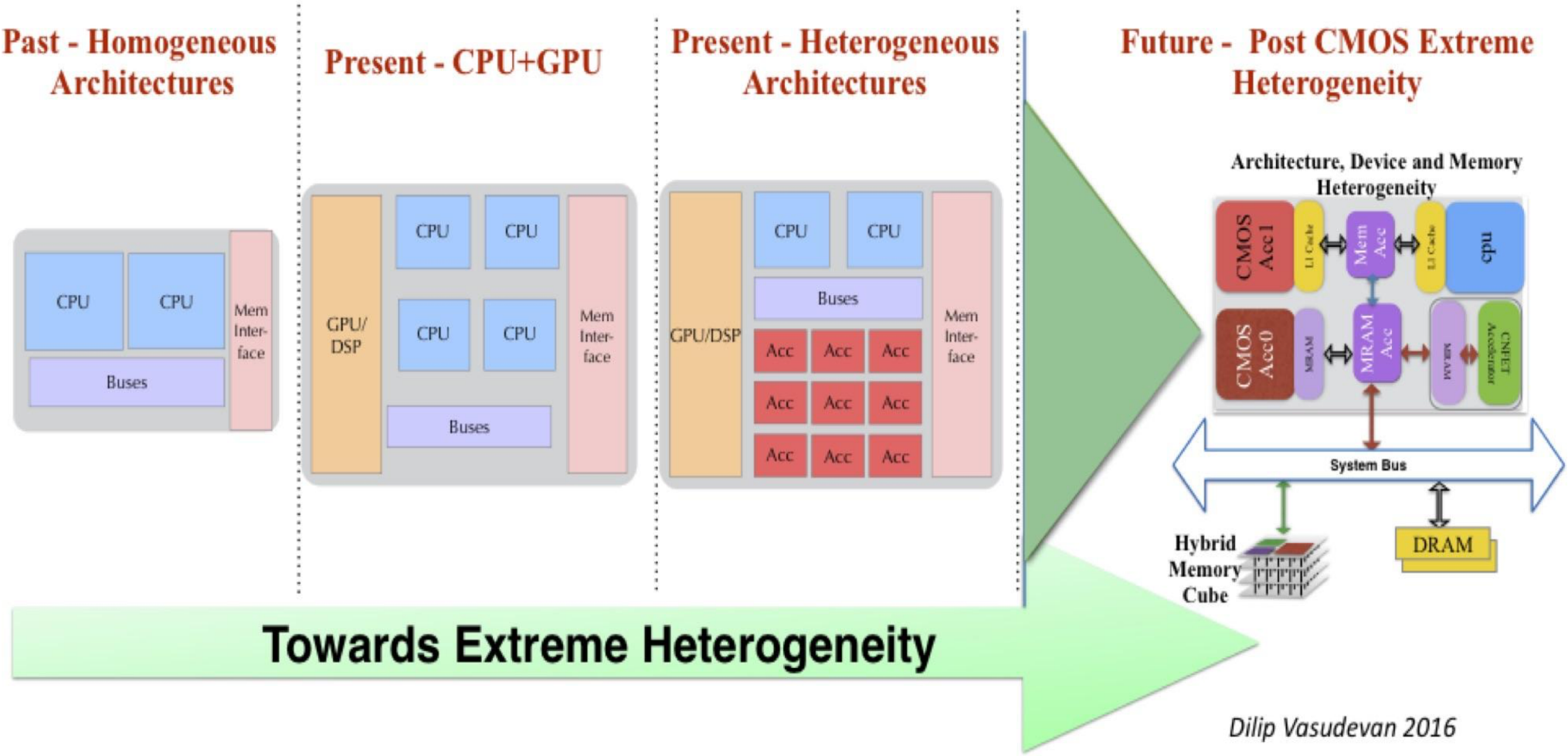
New Materials and Devices
20+ years (10 year lead time)



More Efficient Architectures and Packaging
The next 10 years after exascale

BERKELEY LAB

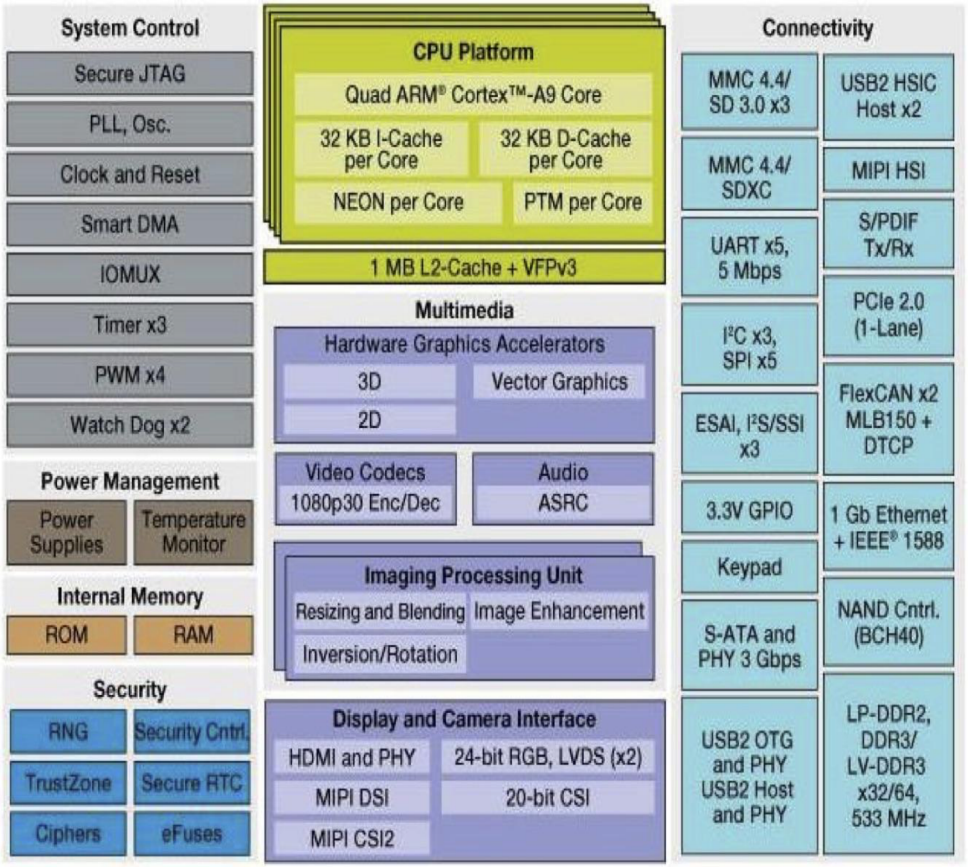
Towards Diverse Integrated Accelerators

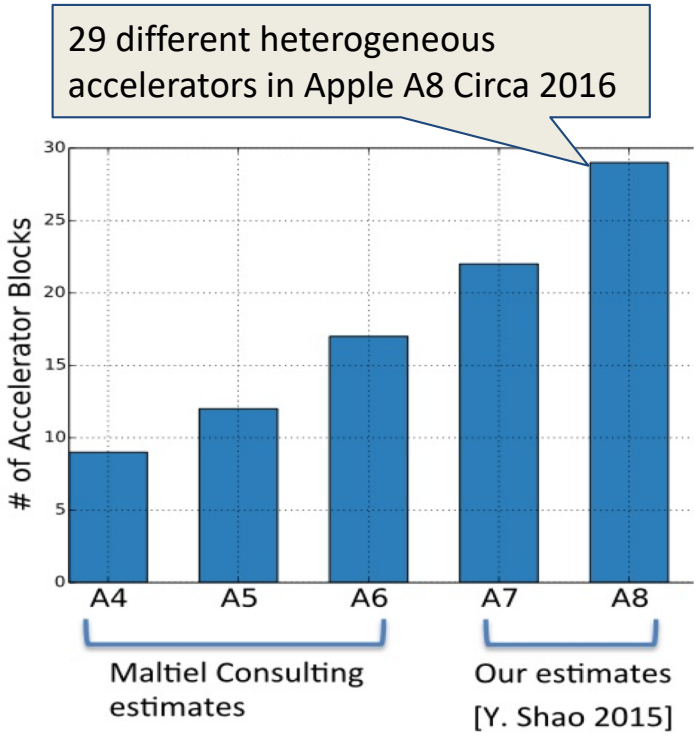


Dilip Vasudevan 2016

Extreme HW Specialization Happening Now

This trend is already well underway in broader electronics industry
Cell phones and even megadatacenters (Google TPU, Microsoft FPGA accel)
(and it will happen to HPC too… will we be ready?)





[www.anandtech.com/show/8562/chipworks-a8]

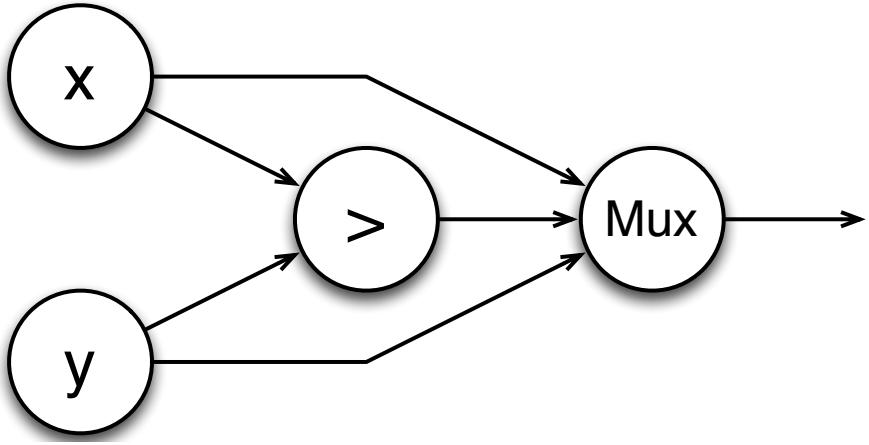
Chisel Overview

A productive, flexible language for hardware design and simulation

Not "Scala to Gates" – it is structural.

- Describe hardware functionality
- Creates graph representation of HW that is flattened as each node translated to Verilog or C++

```
Mux(x > y, x, y)
```



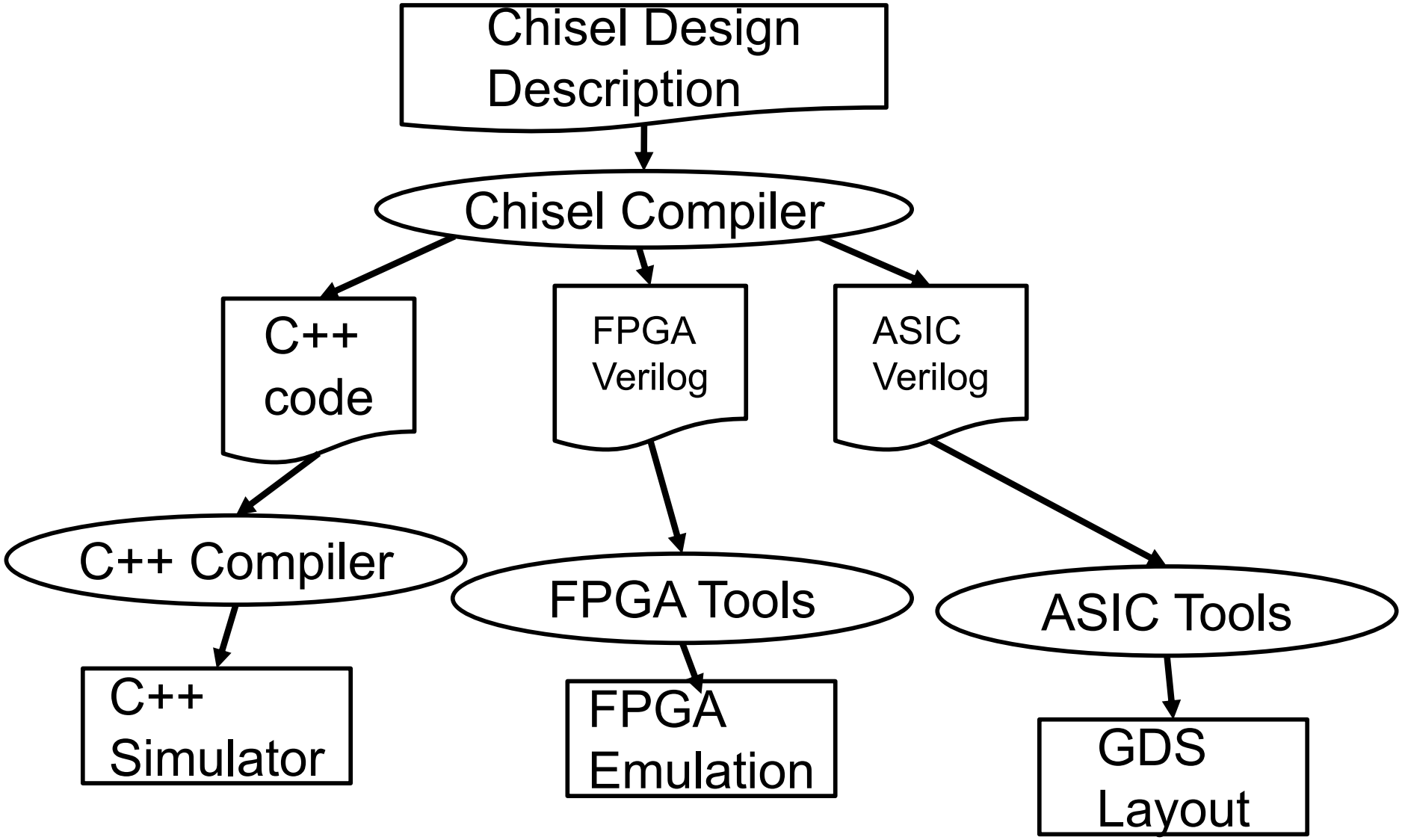
Get benefits of modern high level languages

- Inheritance
- Complex Types
- Modularity
- Polymorphism

***Hardware Generators* are a more efficient technique for generating designs**

Chisel: A DSL for Hardware Design

Infrastructure for Synthesis with Integrated Simulation



Generator Example: Cache

Parameters

IO Connections

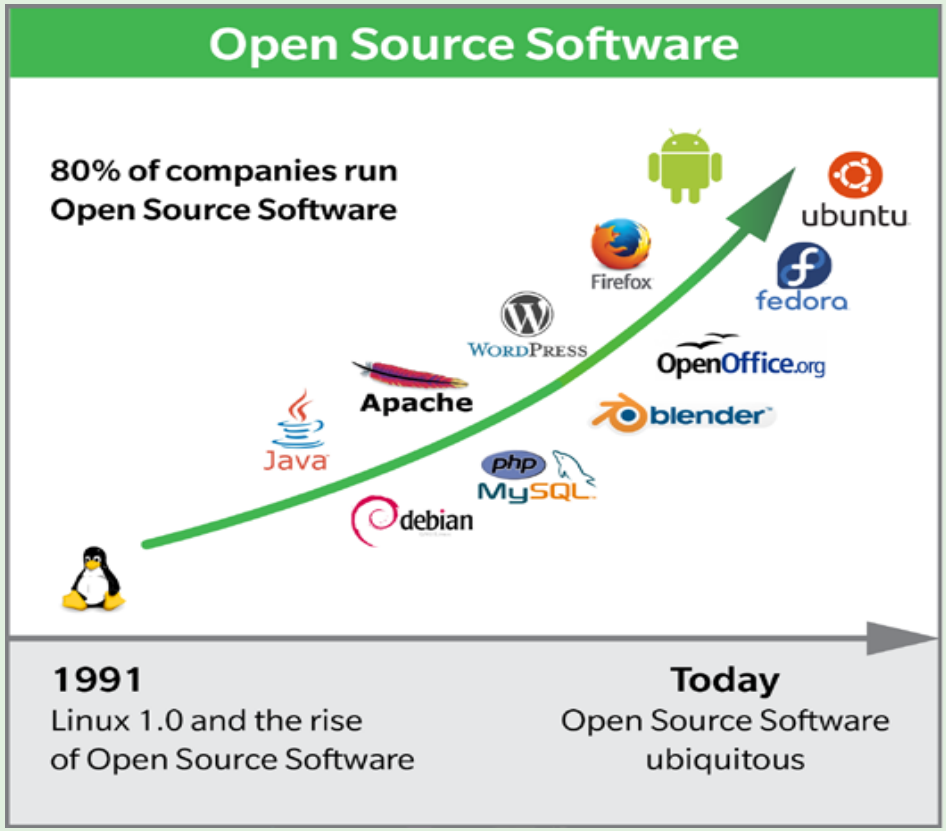
Local Parameters

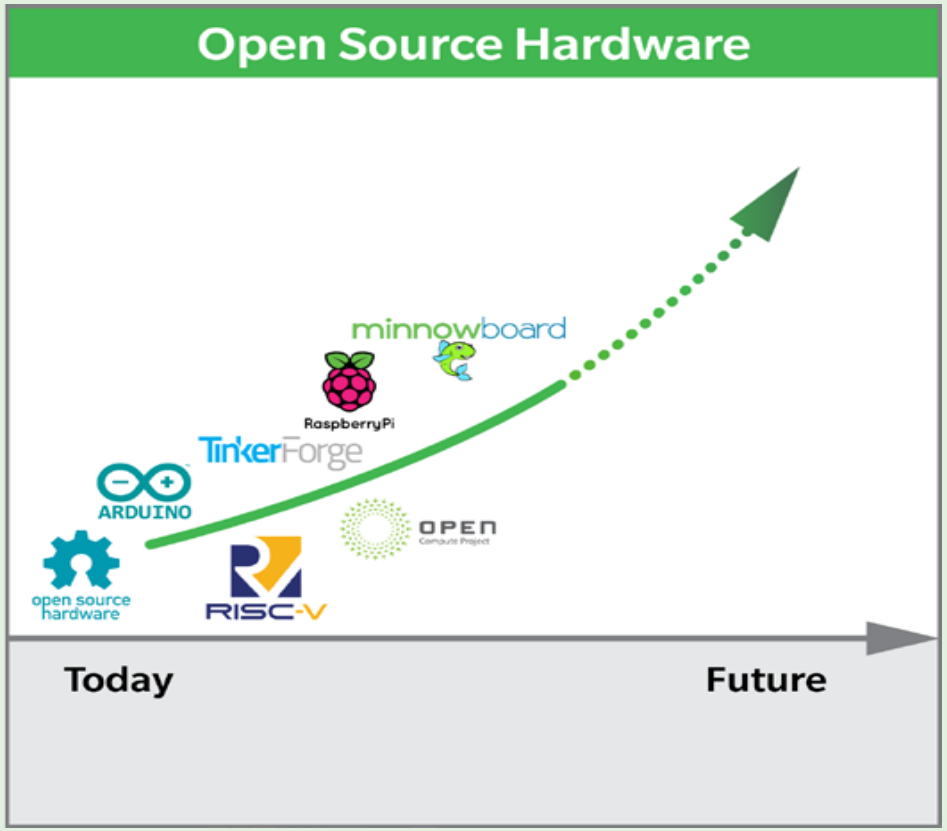
Generator Body

```
class Cache(cache_type:    Int = DIR_MAPPED,
            associativity: Int = 1,
            line_size:     Int = 128,
            cache_depth:   Int = 16,
            write_policy:  Int = WRITE_THRU
           ) extends Component {
  val io = new Bundle() {
    val cpu = new IoCacheToCPU();
    val mem = new IoCacheToMem().flip();
  }
  val addr_idx_width = log2(cache_depth).toInt
  val addr_off_width = log2(line_size/32).toInt
  val addr_tag_width = 32 - addr_idx_width -
                           addr_off_width - 2
  val log2_assoc     = log2(associativity).toInt
  ...
  if (cache_type == DIR_MAPPED)
    ...
}
```

What's Next?

The Rise of Open Source Software: Will Hardware Follow Suit?





- Rapid growth in the adoption and number of open source software projects
- More than 95% of web servers run Linux variants, approximately 85% of smartphones run Android variants
- Will open source hardware ignite the semiconductor industry? Is RISC-V the hardware industry's Linux?

Why Open Source Hardware?

- **More productive and credible hardware research**
  - Generate *real* hardware that can be measured
  - by Reducing the cost of development (Chisel)
  - by creating and sharing open hardware (RISCV, OpenSOC)
- **More innovation**
  - Don't need to be a big company to play
  - Engage academic, lab research community in DSE
- **Lower Cost / Complexity for Development**
  - Share software stack *(complete compilers, debug, Linux ports)*
  - Focus NRE and license on new/innovative IP blocks
  - Stop squeezing license costs out of items that students can implement in a summer (license *hard* stuff)

- ***Whether future is ASICs or FPGAs, you need more productive hardware generation!!!***

COMPUTER ARCHITECTURE LABORATORY
EXASCALE DESIGN SPACE EXPLORATION

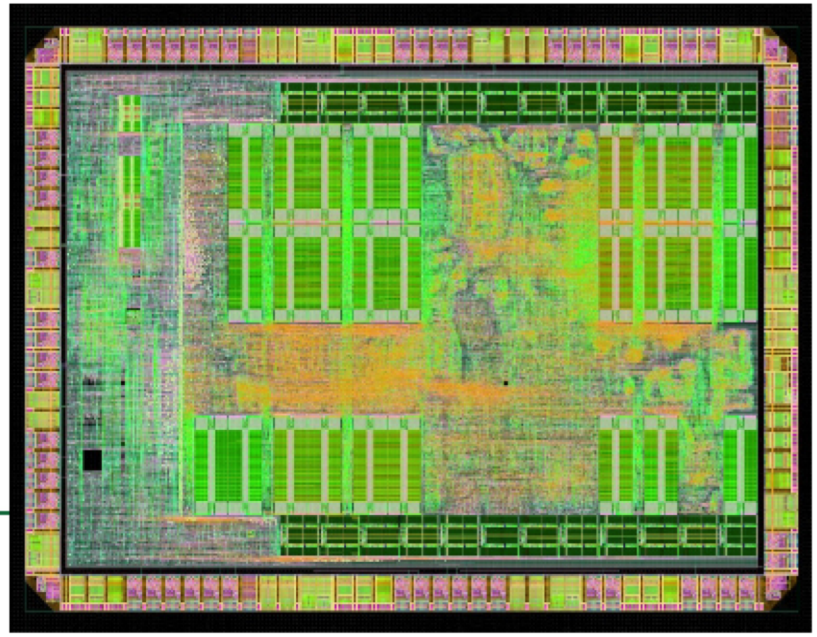
RISC-V Processors

Open source, chisel based processors based on a Open ISA

- **RISC-V Core Taxonomy**
  - Out-of-order (BOOM)
  - In-Order (Rocket)
  - IoT (Z-Scale, Sodor)

- **Can Generate for Different Addressing Widths**
  - 32, 64, 128-bit addressing
  - Double precision floating point
  - Vector accelerators

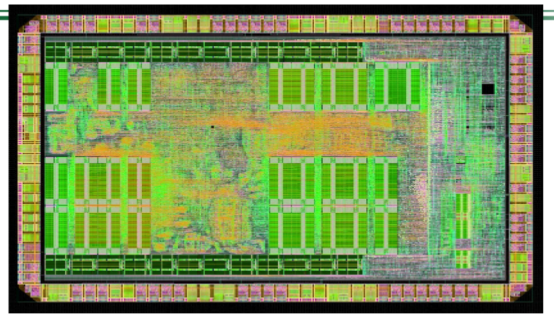
- **Complete SW stack available**
  - GNU and LLVM compilers
  - Linux implementations



"Open Source" *doesn't* mean "Low Performance"

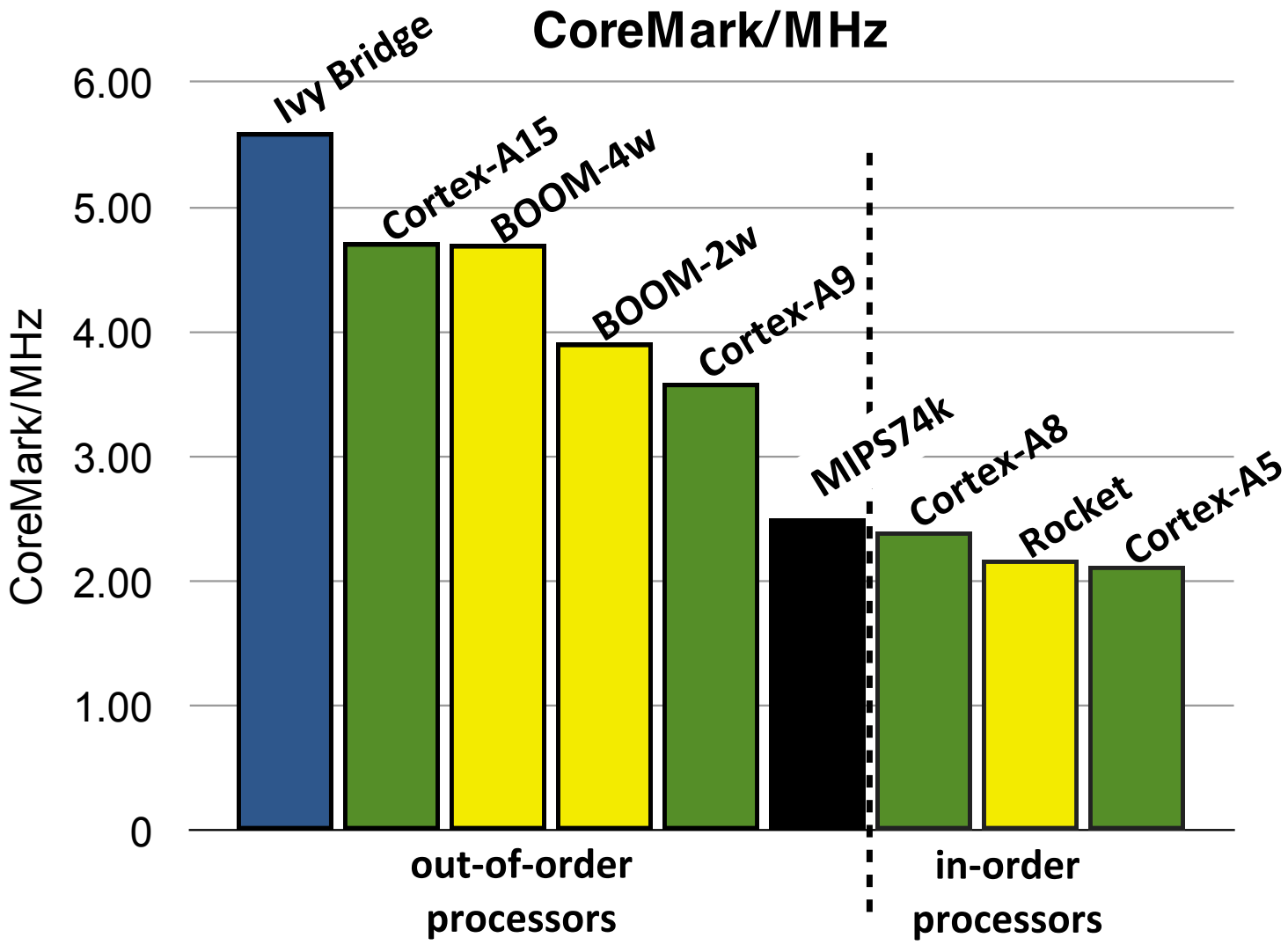
Offers a good proxy to commercial cores such as ARM

Have complete visibility inside of the core design



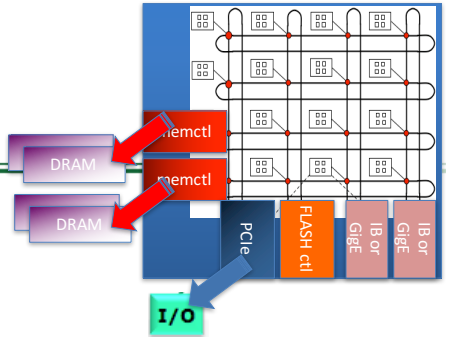
Category	ARM Cortex A5	RISC-V Rocket
ISA	32-bit ARM v7	64-bit RISC-V v2
Architecture	Single-Issue In-Order 8-stage	Single-Issue In-Order 5-stage
Performance	1.57 DMIPS/MHz	1.72 DMIPS/MHz
Process	TSMC 40GPLUS	TSMC 40GPLUS
Area w/o Caches	0.27 mm^2	0.14 mm^2
Area with 16K Caches	0.53 mm^2	0.39 mm^2
Area Efficiency	2.96 DMIPS/MHz/mm^2	4.41 DMIPS/MHz/mm^2
Frequency	>1GHz	>1GHz
Dynamic Power	<0.080 mW/MHz	0.034 mW/MHz

Performance Comparison Across RISC-V Processor Families and ARM Equivalents



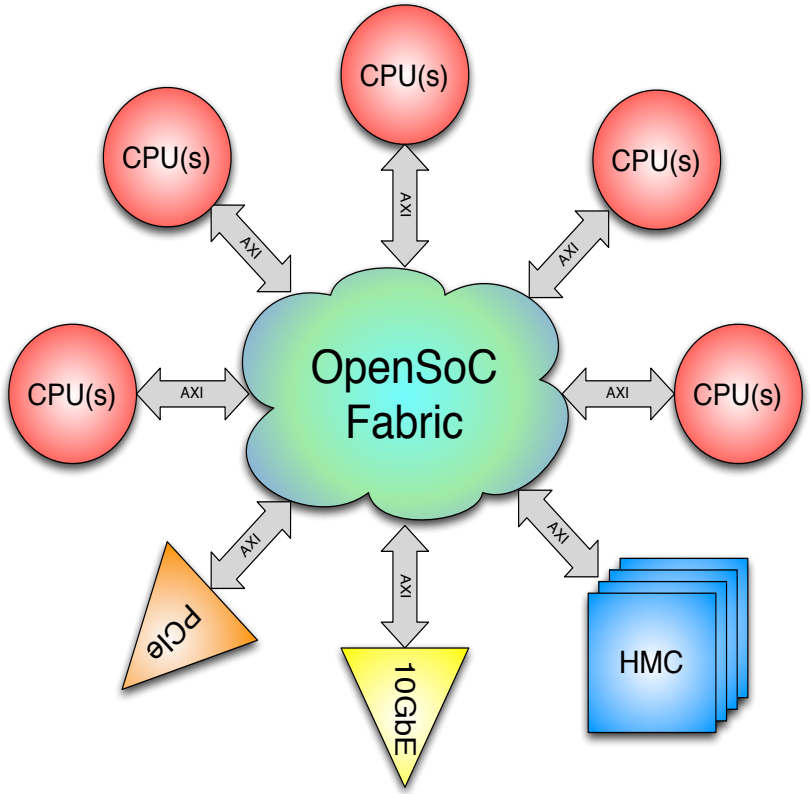
Y. Lee UCB @ Hotchips 2016

OpenSoC Fabric

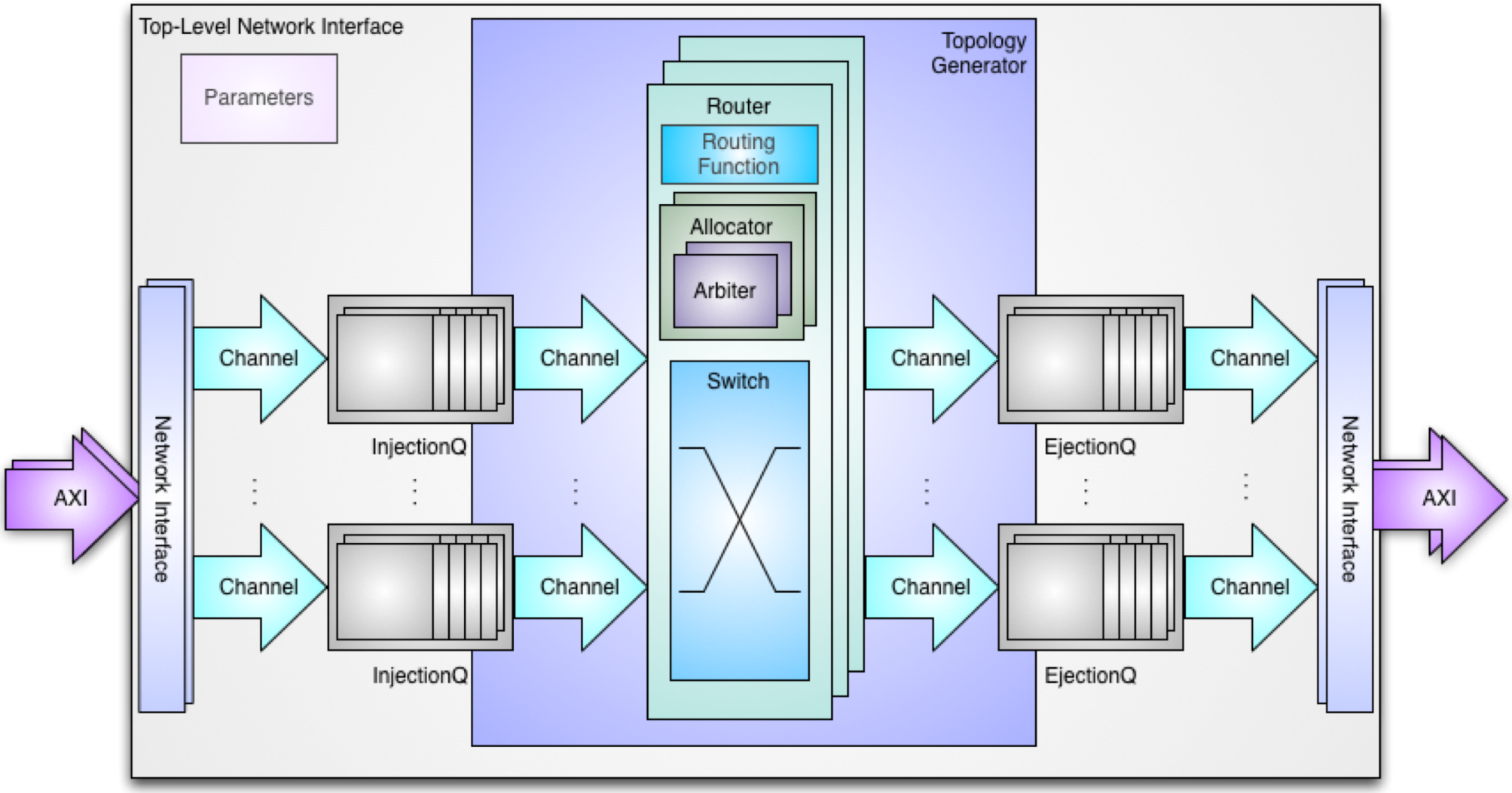


An Open-Source, Flexible, Parameterized, NoC Generator

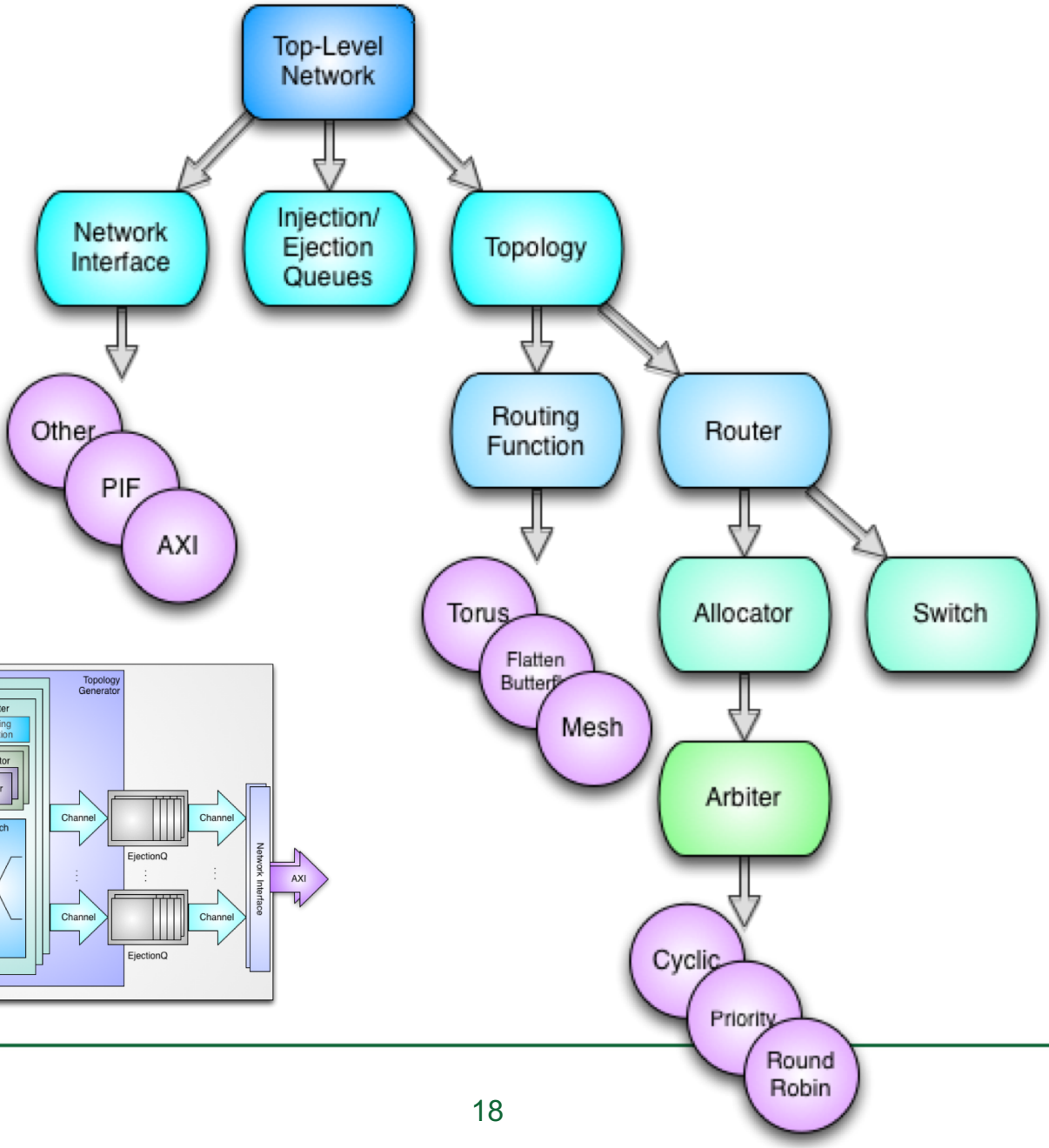
- **Part of the CoDEx tool suite**
- **Written in Chisel**
- **Dimensions, topology, VCs all configurable**
- **Fast functional C++ model for functional validation**
- **Verilog based description for FPGA or ASIC**
  - Synthesis path enables accurate power / energy modeling

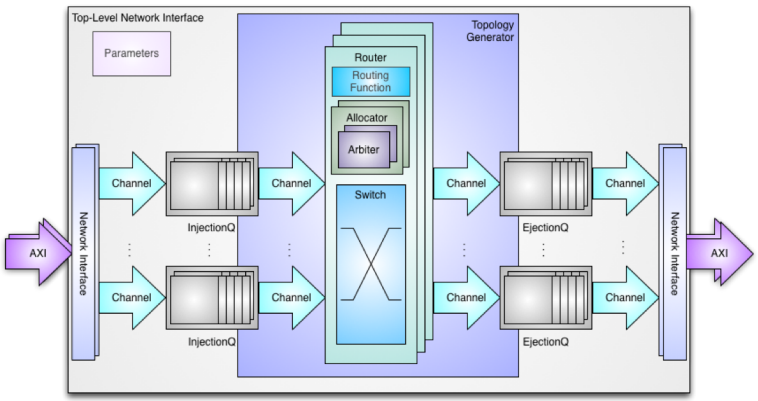


Top Level Router Diagram



OpenSOC: Functional Hierarchy

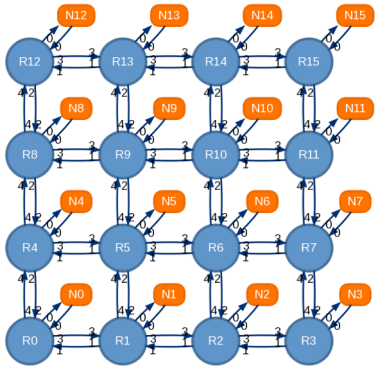


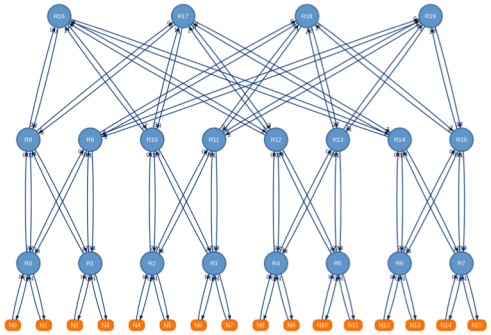


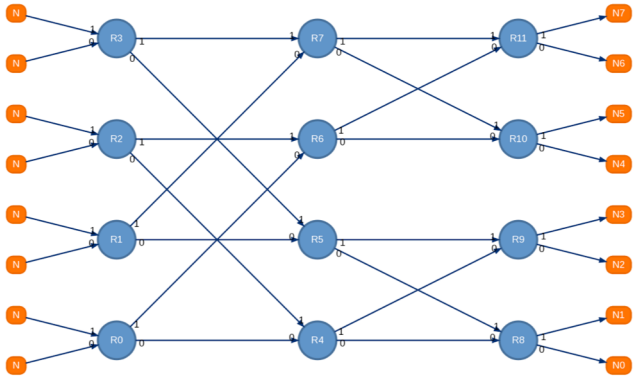
COMPUTER ARCHITECTURE LABORATORY
EXASCALE DESIGN SPACE EXPLORATION

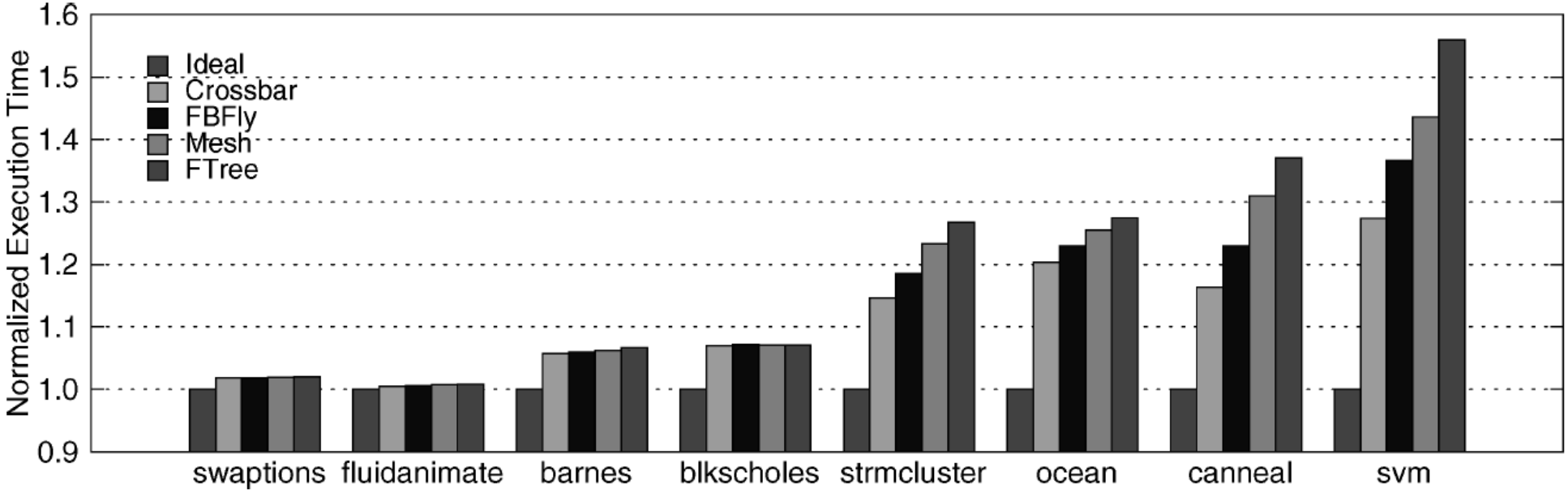
Design Space Exploration of NoC Hardware
(all aspects of NOC parameterized, including topology)

Some common topologies









BERKELEY LAB

An analysis of on-chip interconnection networks for large-scale chip multiprocessors
ACM Transactions on computer architecture and code optimization (TACO), April 2010

SC16 Demonstration of Scalable SoC Emulation on FPGAs

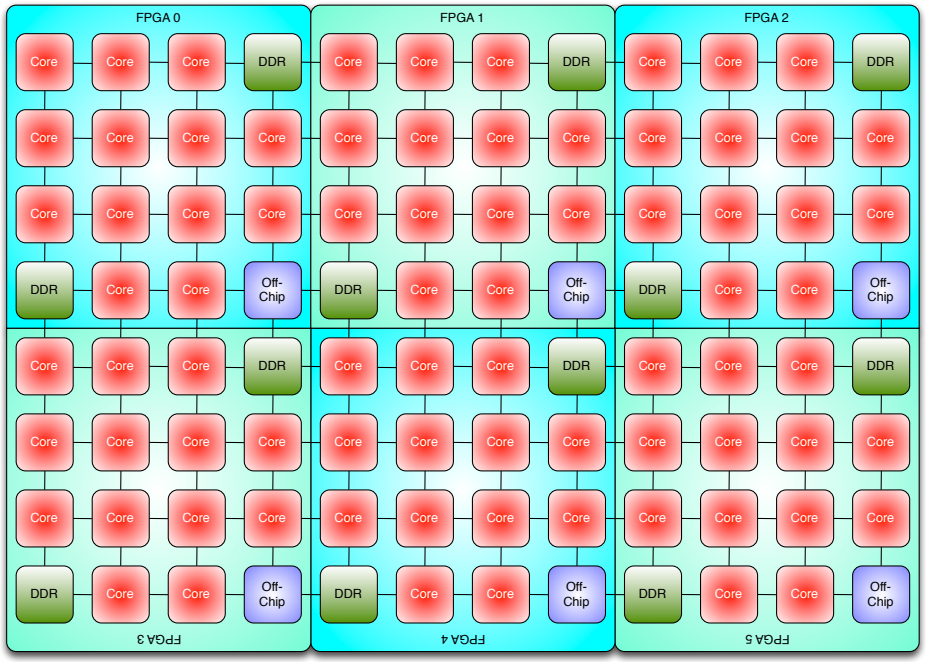
A 96 core SoC with local-store and HBC Memory Subsystem

Full RTL implementation of SOC

The Demo: 96 Core SoC Design for HPC

Shockingly (*but accidentally*) similar to Sunway node architecture

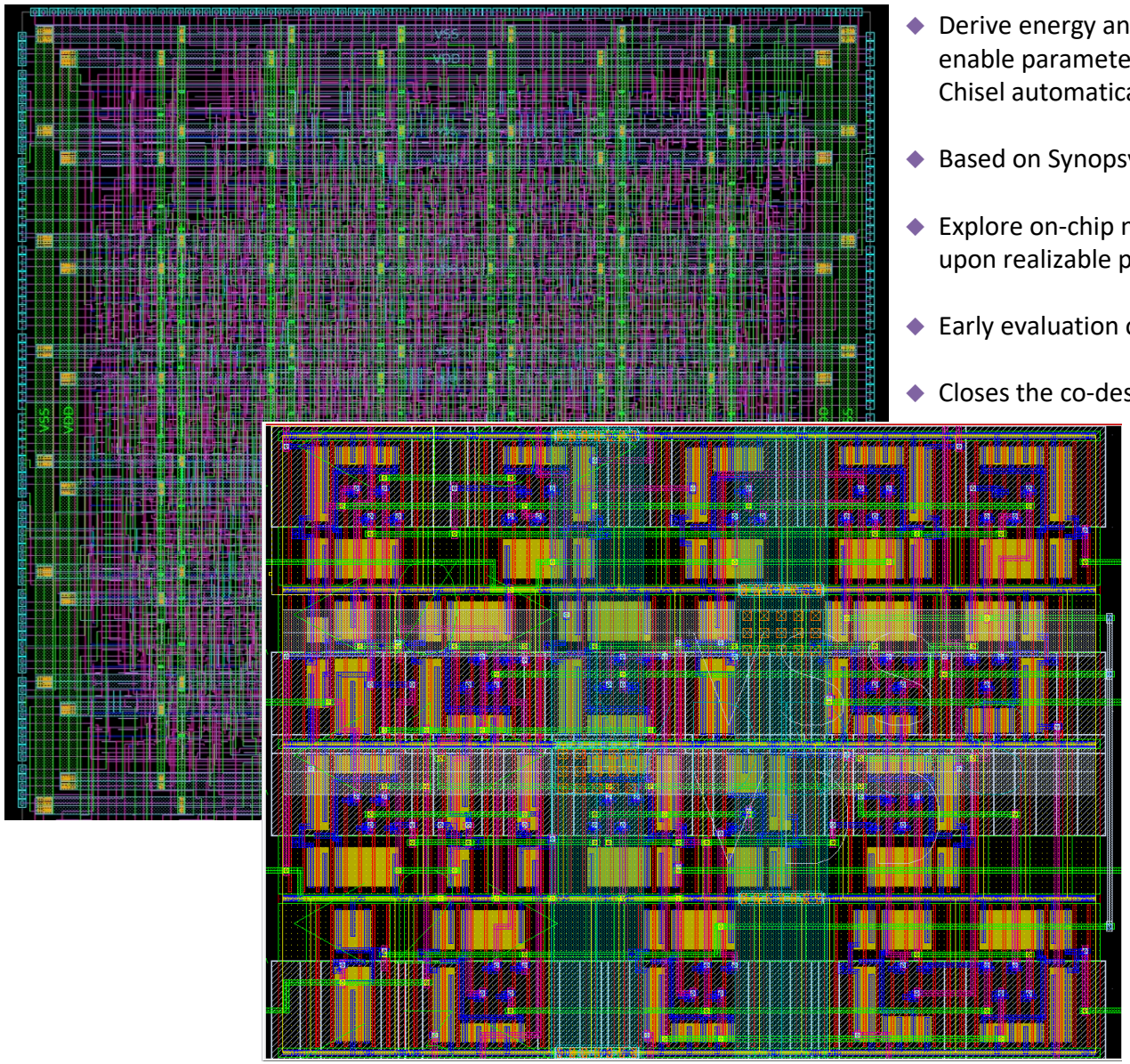
2 people spent 2 months to create



- **Z-Scale processors connected in a Concentrated Mesh**
- **4 Z-scale processors**
- **2x2 Concentrated mesh with 2 virtual channels**
- **Micron HMC Memory**

http://www.codexhpc.org/?p=367

Energy/Power Estimates for Co-design Toolflow



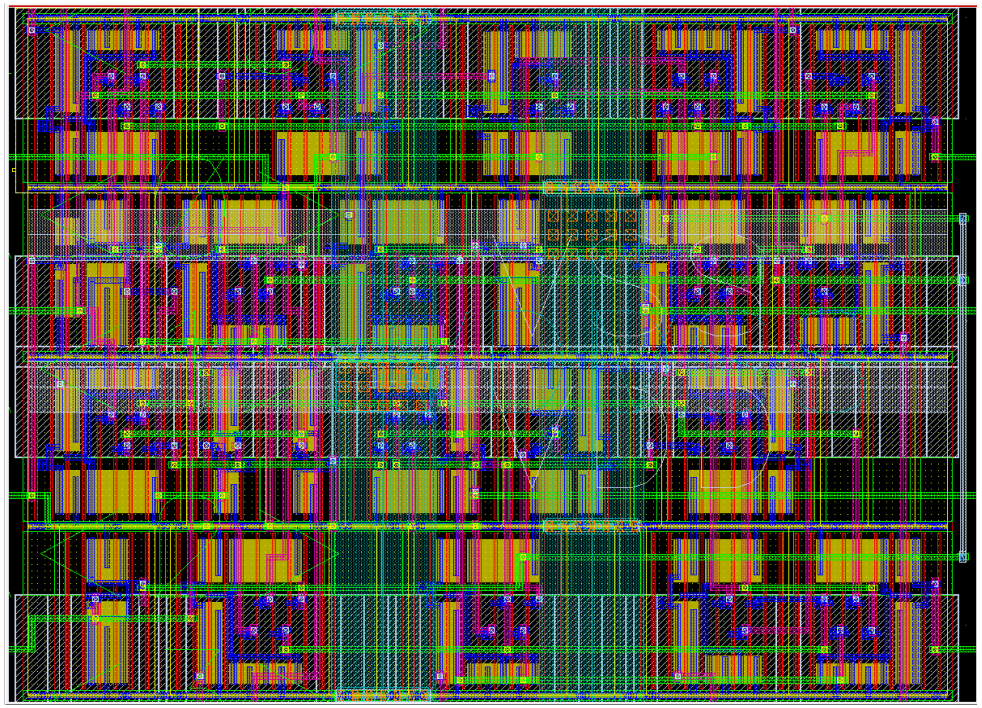
- Derive energy and power estimates automatically through scripted builds to enable parameter-sweep-driven optimization. Input files obtained from Chisel automatically generated Verilog output.
- Based on Synopsys DC synthesis, 32nm SAEG libraries.
- Explore on-chip network topologies useful for HPC and characterize based upon realizable physical parameters for cost extraction.
- Early evaluation of OpenSOC components (8x8 crossbar switch illustrated).
- Closes the co-design loop with physical artifacts from analytical input data.

FPGA vs. ASIC

FPGA

Cost for first FPGA (NRE):	$1,200-$3,000
Cost for 20,000th :	$1,200-$3,000
Clock Rate:	0.1-0.3Ghz

ASIC



Cost for first ASIC (NRE):	$2M-$5M
Cost for 20,000th :	$100
Clock Rate:	1-2 Ghz (10x)
Area Efficiency:	10x FPGA
Energy Efficiency :	10x-100x FPGA

What are the NRE and MFR Costs?

sample "quote" from a design firm (probably a bit low, but gives you an idea of rough cost model) (again: anonymized)

Parameter	Value
Foundry	TSMC
Process Technology	40nm 1P10M (m1 7x2y) , RDL, Bump, VTCp, ESD
Die size estimate	16mm x 16mm
IP	See IP Summary Slide
Package	1156 FCBGA, 3:2:3-layer, 35x35 body size, 1 mm ball pitch
Tester Platform	Agilent-93K-640-300MHz
Test Time	Wafer Sort: 10s Final Test: 10s

IP Description
PCIe Gen 2 PHY $200k (could be IB)
PCIe Gen 2 End point controller $80K
DDR3 PHY $338K
DDR3 Controller $100K

Component	Amount
Manufacturing NRE Masks, 12 Prototype Char Wafers, 150 Prototypes, Process Eng, Product Eng, Project Management	$1,512,970
Package NRE Package Tooling and Package Engineering	$23,460
Test Development NRE Test Engineering and Tester Rental Time	$82,800
IP NRE IP Licensing Fees, Support and Maintenance	$918,000
Characterization NRE Char units, Tester rental, Test Engineering, Process Engineering, Char report	$52,000
Qualification NRE Q&R Engineering, HTOL, TMCL, HTSL, UHAST, ESD & LU	$225,000
Total NRE	$2,814,230

First 2.5 Ku	Next 2.5 Ku	Next 5 Ku	Next 10 Ku	Additional
$121.60	$119.76	$117.97	$87.78	$74.80

With Marty Deneroff

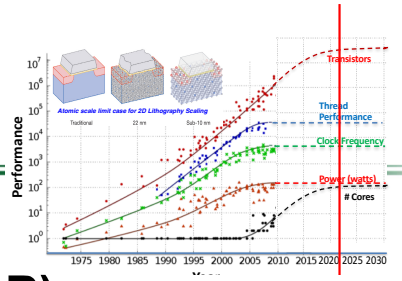
See http://www.socforhpc.org/

Multi-agency + industry and academia workshop
to flesh out this approach to HPC.

Accelerate how to exploit hardware Customization
Beyond Exascale

BERKELEY LAB

Future of Computing



- **Exponentially Increasing Parallelism (central challenge for ECP)**
  - **Trend**: *End of exponential clock frequency scaling (end of Dennard scaling)*
  - **Consequence:** *Exponentially increasing parallelism*

- **End of Lithography as Primary Driver for Technology Improvements**
  - **Trend:** *Tapering of lithography Scaling*
  - **Consequence:** *Many forms of heterogeneous acceleration (not just GPGPUs)*

- **Data Movement Heterogeneity and Increasingly Hierarchical Machine**
  - **Trend**: *Moving data operands costs more than computation performed on them*
  - **Consequence**: *More heterogeneity in data movement performance and energy*

- **Performance Heterogeneity**
  - **Trend**: *Heterogeneous execution rates from contention and power management*
  - **Consequence**: *Extreme variability and heterogeneity in execution rates*

- **Diversity of Emerging Memory and Storage Technologies**
  - **Trend**: *Emerging memory technologies and stall in performance improvements*
  - **Consequence**: *Disruptive changes to our storage environment*

Are we content to target 2021?

Open Source Hardware

- **More credible hardware research**
  - Generate *real* hardware that can be measured
  - by Reducing the cost of development (Chisel)
  - by creating and sharing open hardware (RISCV, OpenSOC)

- **More innovation**
  - Don't need to be a big company to play
  - Engage academic, lab research community in DSE

- **Lower Cost / Complexity for Development**
  - Focus NRE and license on new/innovative IP blocks
  - Stop squeezing license costs out of items that students can implement in a summer (license *hard* stuff)

The End

For more information go to
http://www.cal-design.org/

OpenSOC: http://www.opensocfabric.org/
SOCforHPC: http://www.socforhpc.org/

How is Architecture Research Done Today?

International Symposium on Computer Architecture (ISCA) 2010, 44 accepted papers (18% accept rate)

Types of papers:

- **Proposals (no/few numbers): 2 papers**
- **Modeling techniques: 4 papers**
- **Real machines (analyzed, or used for eval.): 6 papers**
- **New device technology (photonics, MTRAM): 4 papers**
- **Outer memory system (LLC/DRAM/NVM): 9 papers**
- **Processor or inner-cache mechanisms: 19 papers**