Flexibility and Circuit Overheads in Reconfigurable SIMD/MIMD Systems

S. Arrabi[1], D. Moore[2], L. Wang[1], K. Skadron[1], B.H. Calhoun[1], J. Lach[1], B.H. Meyer[3]
[1]University of Virginia, [2]University of Michigan, [3]McGill University
arrabi@virginia.edu

Abstract

Dynamically reconfigurable SIMD/MIMD architectures made from simple cores have emerged to exploit diverse forms of parallelism in applications. In this work, we investigate the circuit-level overhead and flexibility tradeoffs of such architectures through the design of a custom reconfigurable SIMD/MIMD system, with a focus on the core partitioning and granularity of the reconfigurability

Summary

- Designed a 16-core processor using simple in-order core (OpenRISC)
- Explored benefits and overheads of various forms and granularities of reconfigurability using synthesis tools, circuit simulators, cache energy simulators, and architecture simulators
- Energy consumed by long wires in wide architectures can lead to surprising configurations
- Results revealed supporting 4 configurations is the optimal choice

Approach and tools

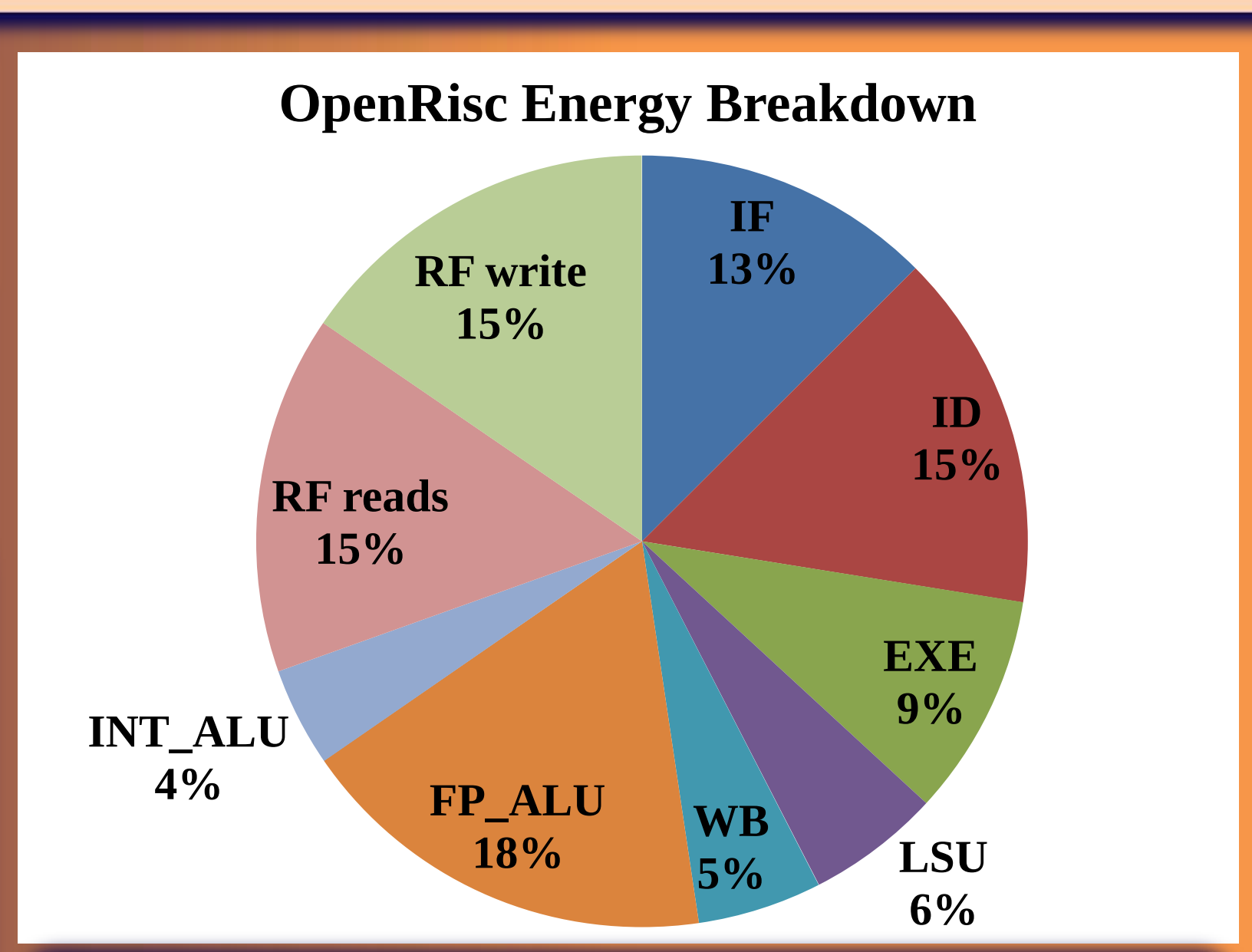
Since technology is going towards simple many cores rather than few complicated ones, we chose OpenRISC, a simple in-order processor, to be our main building block of our 16-core dynamically reconfigurable SIMD/MIMD system.

Two main design decisions explored

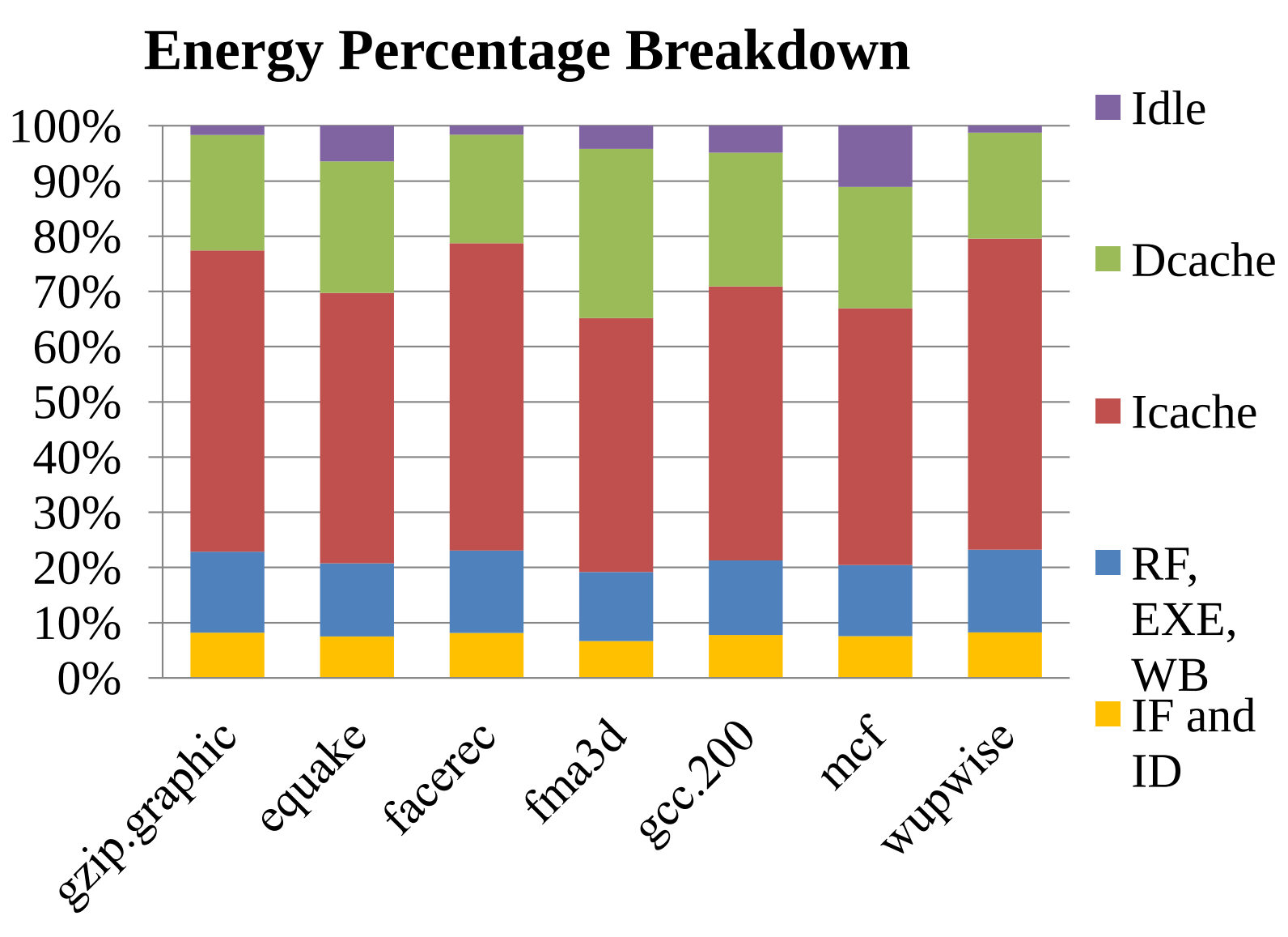
1. How to partition the processor into 2 parts (FE and PE) to support various SIMD and MIMD configurations
2. How many SIMD/MIMD configurations to support

To extract the tradeoffs of partitioning and configurability, we extracted energy and performance information using simulation at multiple design levels:

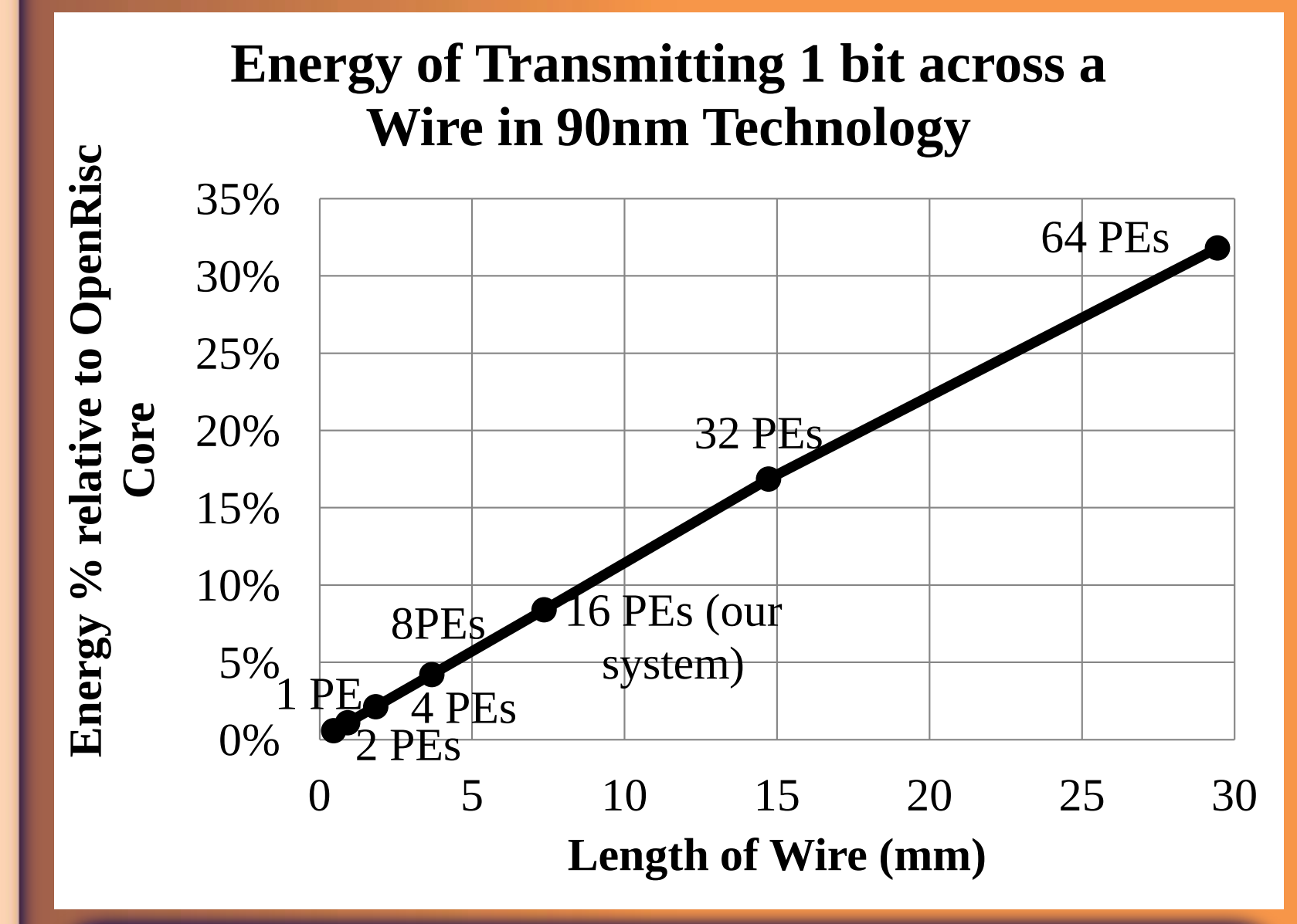
- CACTI for cache
- Cadence for wires and components
- Gem5 for benchmark utilization
- Verilog for estimating wires between components



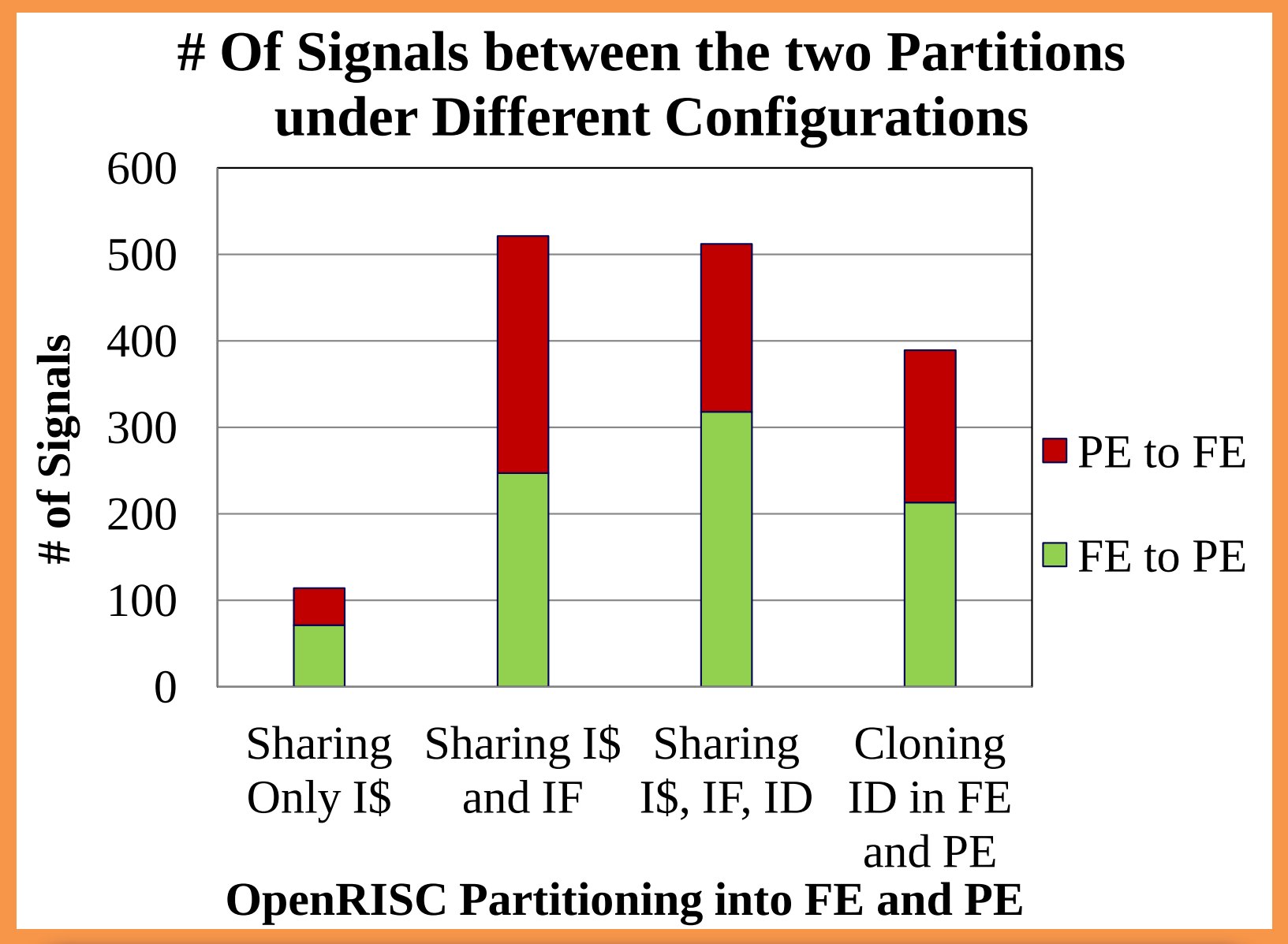
Within the CPU, energy is consumed almost equally among several components



Across all benchmarks, the majority of energy is consumed in caches accesses

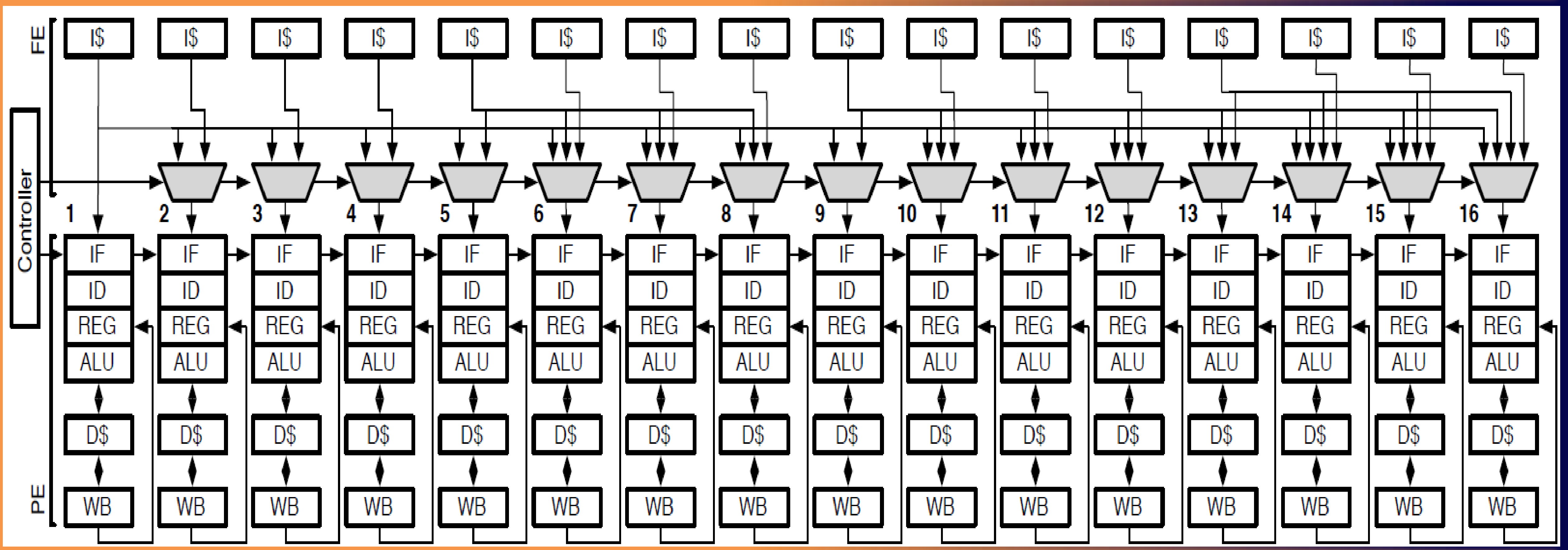


Energy consumed in wires is significant, especially in wide multi-core systems



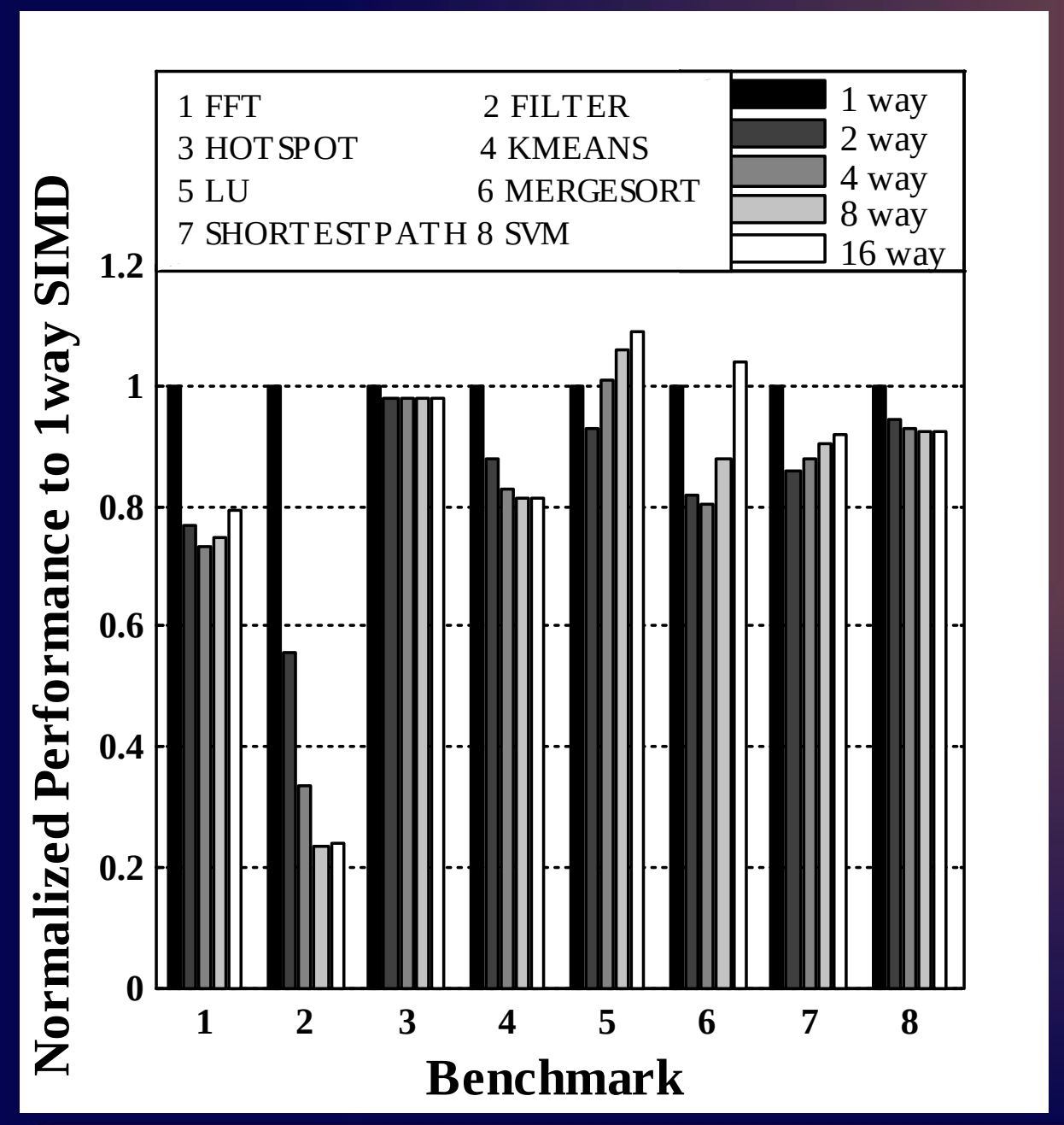
There are many more signals between CPU stages than between I-cache and CPU

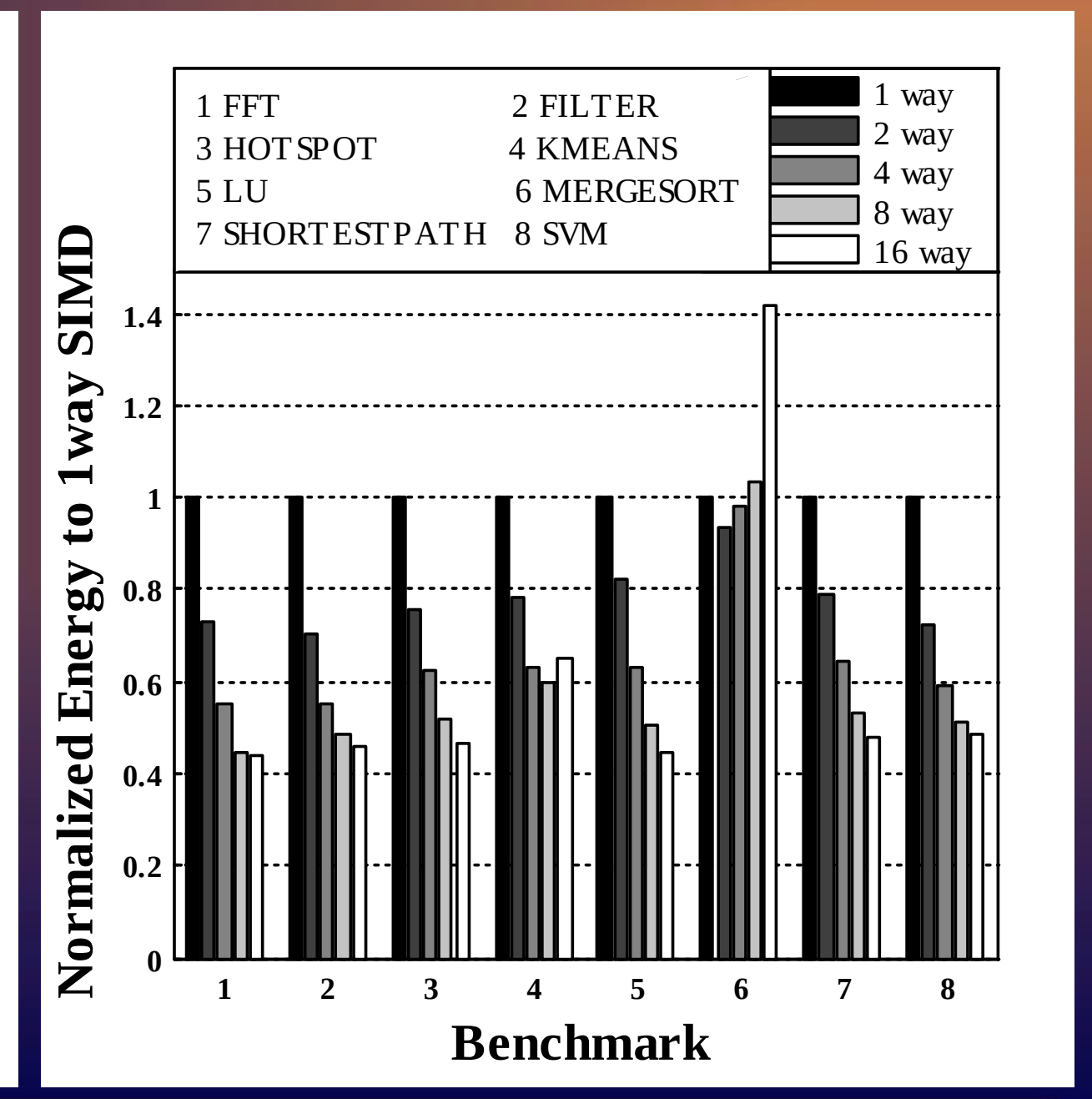
Final design



Features

- 16 cores based on OpenRISC architecture
- Interconnect after I-Cache (rather after Decode) achieves lowest energy consumption
- Supports 4 possible configurations (16-way, 2 8-way, 4 4-way SIMD and 16-way MIMD)
- Caches are divided into banks to enable clock gating in SIMD modes
- Achieves 32% energy savings on average relative to 16-way MIMD
- Achieves 1.7x speedup on average relative to single core

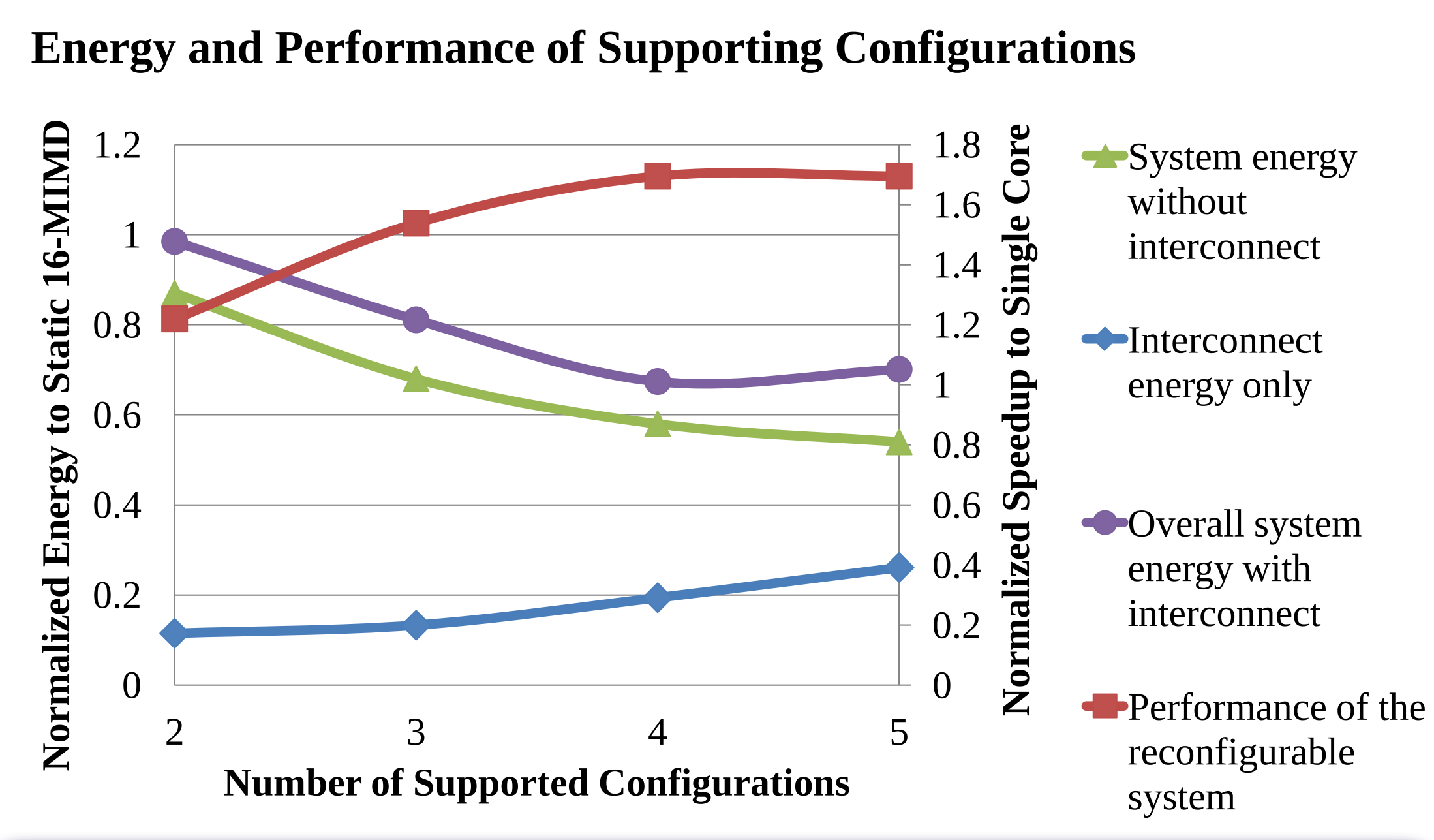




Different benchmarks have different optimal SIMD width for performance and energy consumption

Supported Configurations	16-way SIMD and 16-way MIMD	16-way, 2 8-way SIMD and 16-way MIMD	16-way, 2 8-way, 4 4-way SIMD and 16-way MIMD	16-way, 2 8-way, 4 4-way, 8 2-way SIMD and 16-way MIMD
Multiplexers	15 2to1	8 2to1, 7 3to1	4 2to1, 8 3to1, 3 4to1	4 2to1, 6 3to1, 4 4to1, 1 5to1
Extra wires in the interconnect	1 16x, 15 1x	1 16x, 1 8x, 14 1x	1 16x, 1 8x, 2 4x, 12 1x	1 16x, 18x, 2 4x, 4 2x, 8 1x

Supporting more configurations increases the energy overhead of the interconnect. However, the area impact is insignificant



Supporting reasonable amount of configurations achieves the optimal energy and performance results

Future work

- How to control (and predict) the right configuration at runtime
- Explore the costs/benefits of reconfiguring during a single application due to phase changes
- Explore cache hierarchy systems
- Explore different suites of benchmarks

This work is supported in part by NSF Grant #CCF-0903471 and the Semiconductor Research Corporation