Jieming Yin

4-192 Keller Hall, 200 Union Street SE, Minneapolis, MN 55455
Cell phone: 1-612-423-7296 Office: 1-612-625-7876
jyin@cs.umn.edu

EDUCATION

Ph.D in Computer Science

University of Minnesota, Twin Cities, Sep. 2009–current
Thesis Research: Power-efficient Heterogeneous Multi-processor Systems

B.E. in Electrical and Computer Engineering

Harbin Institute of Technology, China, July 2008
Thesis: Point-to-point Scalar Operand Networks for Clustered Superscalar Processor

RESEARCH EXPERIENCE

Research Assistant in University of Minnesota, Twin Cities Sep. 2009–current

Energy Efficient Network Design for Heterogeneous Multicore Systems:

With the number of applications that can benefit from massive parallel processing increasing, specialized graphics processors units (GPU) with hundreds of small cores capable of running thousands of threads simultaneously have become popular. To fully exploit the diversity of calculation ability, the fusion of CPU and GPU onto the same die becomes attractive and brings with it interesting challenges for optimization of on-chip interconnection network. In this heterogeneous system, the interconnection network needs to handle both latency-sensitive and throughput-sensitive traffic. On the other hand, the network can become a significant source of power consumption and a major performance bottleneck. This ongoing project aims to design an energy efficient on-chip interconnection network while guaranteeing the overall performance for heterogeneous systems.

In this project, I am responsible for:

1. Building up a heterogeneous system that contains both CPU and GPU.
2. Exploit design space and optimizations for on-chip interconnection network.
3. Energy and performance analysis for the proposed heterogeneous system.

Integrated Layout and Architectural Design for Multicore Platforms:

In emerging multicore system, on-chip communication must be facilitated by an interconnection network that connects together various modules within the system. Traditional tile-based multicore architecture that provides each core with a switch is not only inefficient in link utilization but also increases the average hop for message transmission; moreover, redundant buffers and links in symmetric network topologies may show suboptimal energy efficiency. This project aims to develop customized on-chip interconnection architectures that can potentially provide large benefits for building high-performance and energy efficient multicore systems.

In this research project, I am primarily responsible for:

1. The performance evaluation of the simulation infrastructure.
2. Designing power efficient router structure.
3. Developing the optimization engine and schemes on dynamic thread scheduling and data allocation.

Parallelization of Single-Threaded Programs Using Speculative Slicing

People are paying more and more attention to the emerging multicore systems, aiming to achieve high throughput. However, improving the performance of sequential programs is still important. An attractive way of utilizing on-chip

resources and achieving high performance for a single-threaded program is to execute the program in parallel by speculatively slicing the program. In other words, divide the program into multiple slices that can be executed simultaneously by grouping dependent instructions into a single slice. The goal of the project is to design and evaluate slicing algorithms that can be implemented in multicore systems.

In this project, I work with two groups that develop compilation and hardware support for speculative program slicing. In particular, I focus on:

1. Compiler algorithms that generate program dependent graphs and collects program runtime profile.
2. Speculatively parallelize single thread programs using slicing algorithms based on dependent graph and profile.
3. Generate traces that provide profile information for simulation.

Research Assistant in Harbin Institute of Technology **Jan. 2009–June. 2009**

Instruction Criticality Analysis and Implementation for Clustered Superscalar Processor

Predict the performance of clustered superscalar processors using instruction criticality and instruction dependent distance information, carrying out hardware optimizations.

Undergraduate Thesis **Nov. 2007-July. 2008**

Research on Point-to-point Scalar Operand Networks for Clustered Superscalar Processor

Research on High Performance Clustered Superscalar Processors. Specifically, explore novel structures for low-latency and area-efficient on-chip inter-connections, carrying out system level modeling and evaluation.

PUBLICATION

Conference Paper:

[1] **Jieming Yin**, Pingqiang Zhou, Anup P. Holey, Sachin S. Sapatnekar, Antonia Zhai. *Energy Efficient Non-Minimal Path On-chip Interconnection Network for Heterogeneous Multicore Systems*. The International Symposium on Low Power Electronics and Design (ISLPED), Redondo Beach, USA, Aug. 2012.

[2] Pingqiang Zhou, **Jieming Yin**, Antonia Zhai, Sachin S. Sapatnekar. *NoC Frequency Scaling with Flexible-Pipeline Routers*. The International Symposium on Low Power Electronics and Design (ISLPED), Fukuoka, Japan, Aug. 2011.

[3] Bing Yang, Zhigang Mao, **Jieming Yin**, Xiao Chen. *A Point-to-Point Inter-Cluster Communication Network in Clustered Superscalar Processor*. IEEE The 9th International Conference on Solid-State and Integrated-Circuit Technology. Beijing, China, Oct. 2008.

Manuscript in Progress:

[1] *Energy-Efficient Time-Division Multiplexed Hybrid NoC*

Poster:

[1] *Energy-efficient NoC Design for CMP System*. 15th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS 2010). Pittsburgh, PA, March, 2010

COMPUTER SKILLS

Programming Languages	C/C++ and SystemC, CUDA, Standard ML, Perl, VerilogHDL, and assembly languages for x86, IA64 and SPARC.
Operating Systems	LINUX, UNIX and Windows.
Tools	Pin Tools, Matlab, ModelSim, SimpleScalar Toolkit and Simics/GEMS, gpgpu-sim Simulator, Wattch and Orion Power model.