

Amir Roth

7/26/2011

DOB: 12/12/1971

Citizenship: US

Computer and Information Science Department
University of Pennsylvania
3330 Walnut Street, 603 Levine Hall
Philadelphia, PA 19104-6389

215.573.0175
amir@cis.upenn.edu
<http://www.cis.upenn.edu/~amir/>

Education

- 2001 **Doctorate of Philosophy in Computer Science**
University of Wisconsin-Madison, Madison, WI
Thesis title: Pre-Execution via Speculative Data-Driven Multithreading.
Thesis advisor: Gurindar S. Sohi.
- 1997 **Master of Science in Computer Science**
University of Wisconsin-Madison, Madison, WI
- 1994 **Bachelor of Science in physics and Bachelor of Arts in Computer Science**
Yale University, New Haven, CT
magna cum laude

Positions Held

- 2010– **Technical Development Manager**, building performance modeling tools
US Department of Energy, Energy Efficiency and Renewable Energy, Washington, DC
- 2007– **Associate Professor**
Computer and Information Science Department, University of Pennsylvania, Philadelphia, PA
- 2001–2007 **Assistant Professor**
Computer and Information Science Department, University of Pennsylvania, Philadelphia, PA
- 2000 Intern
Microprocessor Research Lab (MRL), Intel Corporation, Haifa, Israel
- 1995–2001 **Graduate Research Assistant**
Computer Sciences Department, University of Wisconsin–Madison, Madison, WI
- 1994–1995 **Software Design Engineer**
Microsoft Corporation, Redmond, WA

Honors, Distinctions, and Awards

- 2001 Outstanding Graduate Student Research Award
Computer Sciences Department, University of Wisconsin-Madison
- 2000–2001 Intel Graduate Research Fellowship
- 1999–2000 IBM Graduate Research Fellowship
- 1998–1999 IBM Graduate Research Fellowship
- 1995–1996 Wisconsin Alumni Research Foundation Fellowship
- 1994 Phi Beta Kappa
- 1994 Yale University Physics Award
- 1994 Trumbull College (Yale University) Fellows Award

Grant Awards

- 2010 NSF SHF award CCF-1017184
“AfterBurner: Efficient Performance Scaling via Post Retirement Processing”
\$450,000
- 2010 NSF EAGER award CCF-1012008
“FIESTA: A Sound Multi-Program Workload Methodology”
\$208,000
- 2008 Intel Corporation Research Council Grant
“Unifying, Simplifying, and Amplifying Checkpoint-Based Processors”
\$40,000
- 2006–2009 NSF CPA award CCF-0541292
“ON-Core: Single-Thread Performance via Multi-Core Aggregation”
\$400,000
PI, Prof. M. Martin is Co-PI
- 2004–2006 Intel Corporation Research Council Grant
“Exploiting Dataflow Mini-Graphs in Superscalar Processors”
\$120,000
- 2003–2006 NSF CSA award CCF-0311199
“DISE: A New Hardware-Software Interface for Customizing Application Execution”
\$240,000
Co-PI, Prof. E Lewis is PI
- 2003–2008 NSF CAREER award CCF-0238203
“Correctness-Performance Partitioned (CPP) Architectures”
\$400,000

Research Interests

For the past several years, my research has had several areas of focus. One is energy-efficient microarchitectures. My students and I have developed a set of efficient interlocking mechanisms that collectively support speculative retirement and selective re-execution. These mechanisms can be used to implement any combination of load-latency tolerance, store-latency tolerance, and selective recovery from control and data mis-speculations. Latency-tolerance applications are especially attractive because they complement existing power/performance modulation techniques like dynamic voltage and frequency scaling (i.e., TurboBoost). Latency-tolerance is both more energy-efficient than TurboBoost and can improve performance in cases where TurboBoost cannot.

Another area has been workload methodologies for multi-program experiments. Multi-program execution is becoming more important as industry continues to integrate more cores onto a chip. Prevailing methodologies for multi-program experiments are unsound because they either include significant imbalance and thus don't represent multi-program execution faithfully, or they allow the experiment to influence the workload mixture. My students and I have developed an experimental framework called FIESTA that uses fixed pre-determined workloads but pre-balances them by construction.

One recent area of interest is memory efficient parallel optimization algorithms. Along with Prof. Peter Hahn and colleagues at Clemson University, I have been investigating a new implementation of the 3rd-level reformulation and linearization technique (RLT3) for solving the quadratic assignment problem (QAP). Our new implementation cuts memory requirements by 75%—allowing larger problems to be solved—and uses non-blocking data structures to facilitate parallelization.

Professional Societies

1998– ACM: SIGARCH, SIGMICRO
2001– IEEE: TC-MARCH

Teaching Experience

Spring 2010 CIS 371: Computer Organization and Design
Fall 2009 CIS 240: Introduction to Computer Systems
Spring 2009 CIS 372: Computer Organization and Design Lab
Fall 2008 CIS 240: Introduction to Computer Systems
Spring 2008 CIS 372: Computer Organization and Design Lab
Fall 2007 CIS 501: Introduction to Computer Architecture
Spring 2007 CSE 371: Digital Systems Organization and Design
Fall 2006 CIS 501: Introduction to Computer Architecture
Spring 2006 CSE 371: Digital Systems Organization and Design
Spring 2005 CSE 371/372: Digital Systems Organization and Design / Lab
Fall 2004 CIS 501: Introduction to Computer Architecture
Spring 2004 CSE 371/372: Digital Systems Organization and Design / Lab
Fall 2003 CIS 501: Introduction to Computer Architecture
Spring 2003 CIS 700/5: Processor Design and Implementation Techniques
Fall 2002 CIS 501: Introduction to Computer Architecture
Spring 2002 CIS 700/3: Advanced Topics in Computer Architecture
Fall 2001 CIS 501: Introduction to Computer Architecture

Theses Supervised

May 2010 Andrew Hilton, Doctorate of Philosophy
Thesis: “Energy-Efficient Latency-Tolerant Execution”
First employment: IBM Corporation

May 2009 Tingting Sha, Doctorate of Philosophy (expected)
Thesis: “Simplifying the Store-Load Datapath using Memory Dependence Prediction”
First employment: nVIDIA Corporation

Dec. 2008 Anne Bracy, Doctorate of Philosophy
Thesis title: “Mini-Graphs: High-Granularity, Application-Specific Instructions”
First employment: Intel Corporation

Apr. 2007 Vlad Petric, Doctorate of Philosophy
Thesis title: “RENO: Rename-Based Instruction Optimization”
First employment: Google, Inc.

Dec. 2009 Neeraj Eswaran, Masters of Science in Computer Science
First employment: Microsoft Corporation

May 2006 Salil Wadhavkar, Masters of Science in Engineering, Temple University
Thesis title: “Reducing the Overhead of Runahead Execution Using RENO”
First employment: Ph.D. program
Department of Electrical and Computer Engineering, North Carolina State University

Service Duties: Conference Organizing Committees

- 2009 42nd International Symposium on Microarchitecture (MICRO) Finance Chair
- 2005 32nd International Symposium on Computer Architecture (ISCA) Registration Chair
- 2004 37th International Symposium on Microarchitecture (MICRO) Travel Grant Chair
- 2002 29th International Symposium on Computer Architecture (ISCA) Registration Chair

Service Duties: Conference Program Committees

- 2010 43rd International Symposium on Microarchitecture (MICRO)
- 2010 37th International Symposium on Computer Architecture (ISCA)
- 2009 23rd International Conference on Supercomputing (ICS)
- 2008 26th International Conference on Computer Design (ICCD)
- 2007 IEEE Micro Top Picks from Architecture Conferences (Top Picks)
- 2007 40th International Symposium on Microarchitecture (MICRO)
- 2007 2007 Conference on High Performance Computing (HiPC)
- 2006 ACM Conference on Languages, Compilers, and Tools for Embedded Systems (LCTES)
- 2005 1st IEEE International Symposium on Workload Characterization (IISWC)
- 2005 38th International Symposium on Microarchitecture (MICRO)
- 2004 37th International Symposium on Microarchitecture (MICRO)
- 2003 36th International Symposium on Microarchitecture (MICRO)

Service Duties: Conference and Journal Paper Review

- International Symposium on Computer Architecture (ISCA)
- International Symposium on Microarchitecture (MICRO)
- Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)
- International Conference on High Performance Computer Architecture (HPCA)
- International Conference on Supercomputing (ICS)
- International Conference on Parallel Architectures and Compilation Techniques (PACT)
- International Symposium on Performance Analysis of Software Systems (ISPASS)
- International Conference on Computing Frontiers (CF)
- IEEE Transactions on Computers (TC)
- ACM Transactions on Parallel and Distributed Systems (TPDS)
- ACM Transactions on Computer Systems (TOCS)
- ACM Transactions on Architecture and Compiler Optimizations (TACO)

Service Duties: Proposal Review

- 2010 DOE ASCR Technologies and Architectures for Exascale Computing
- 2007 NSF Computer and Communications Foundations CAREER (CCF CAREER)
- 2006 NSF Computer Processes and Artifacts (CPA)
- 2004 NSF Computer Systems Architecture (CSA2)

Service Duties: Department

- 2009 – Undergraduate program chair
- 2008– Computer Engineering Program Curriculum Committee
- 2002 Written Progress Examination I (WPE I) reorganization committee
- 2001–2005 Computer systems hiring sub-committee
- 2001–2004 Computer Architecture Written Progress Examination I (WPE I) chair
- 2006 Thesis committee chair: Marc Corliss
- 2010 Thesis committee member: Gaurav Shah
- 2010 Written Progress Examination II (WPE II) chair: Santosh Nagarakatte
- 2006 Written Progress Examination II (WPE II) chair: Colin Blundell
- 2005 Written Progress Examination II (WPE II) chair: Marc Corliss
- 2006 Written Progress Examination II (WPE II) committee member: Andrew Hilton
- 2006 Written Progress Examination II (WPE II) committee member: Tingting Sha
- 2005 Written Progress Examination II (WPE II) committee member: Vlad Petric
- 2004 Written Progress Examination II (WPE II) committee member: Stephen Tse
- 2003 Written Progress Examination II (WPE II) committee member: Anne Bracy

Invited Talks

CFPm: Microarchitecture for Future Multi-Cores

Purdue University, Oct. 29, 2009.

University of Illinois at Urbana-Champaign, Oct. 30, 2009.

Resource Amplification: Efficient Scaling for Future Superscalars

Cornell University, AMD lecture series, Sept. 21, 2006.

State University of New York (SUNY)-Binghamton, Sept. 22, 2006.

Harvard University, Nov. 20, 2006.

Massachusetts Institute of Technology, Nov. 21, 2006.

Store Vulnerability Window (SVW): A Re-Execution Filter for Enhanced Load Optimization.

North Carolina State University, Oct. 19, 2004.

Dataflow Mini-Graphs: Amplifying Superscalar Capacity and Bandwidth.

Duke University, Oct. 16, 2004.

A High Bandwidth Load Store Unit for Single and Multithreaded Processors.

University of Rochester, Jul. 9, 2004.

Applying Dataflow Principles to VonNeumann Architectures.

Faculty Research Seminar, Oct. 17, 2003.

Energy Aware Pre-Execution and P-Thread Selection.

Penn-Princeton Architecture Day, Jul. 30, 2003.

Functionality and Formalism.

NSF panel on future directions in computer architecture, Jun 8, 2003.

Pre-Execution: Staying on the Performance Curve.

Interview talk given Feb. 25–May 8., 2001.

University of Pennsylvania

Princeton University

Cornell University

University of Rochester

University of Washington

Duke University

North Carolina State University

Harvard University

Brown University

University of Colorado-Boulder

University of California-Los Angeles

Dependence-Based Prefetching for Linked Data Structures.

IBM Research Lecture Series, IBM Corp. Rochester, MN. Jan. 19, 1999.

Conference Talks

BOLT: Energy-Efficient Out-of-Order Latency Tolerant Execution.

16th International Symposium on High Performance Computer Architecture (**HPCA'10**), Jan. 13, 2009.

Store Vulnerability Window (SVW): A Re-Execution Filter for Enhanced Load Optimization.

32nd International Symposium on Computer Architecture (**ISCA'05**), Jun. 8, 2005.

DISE: A Programmable Macro-Engine for Customizing Applications.

30th International Symposium on Computer Architecture (**ISCA'03**), Jun. 11, 2003.

A Quantitative Framework for Pre-Execution Thread Selection.

35th International Symposium on Microarchitecture (**MICRO'02**). Nov. 22, 2002.

Speculative Data-Driven Multithreading.

7th International Symposium on High Performance Computer Architecture (**HPCA'01**). Jan. 20, 2001.

Register Integration: A Simple and Efficient Implementation of Squash Reuse.

33rd International Symposium on Microarchitecture (**MICRO'00**). Dec. 12, 2000.

Improving Virtual Function Target Prediction via Dependence Based Pre-Computation.

1999 ACM International Conference on Supercomputing (**ICS'99**). June 25, 1999.

Effective Jump Pointer Prefetching for Linked Data Structures.

26th International Symposium on Computer Architecture (**ISCA'99**). May 3, 1999.

Microarchitectural Miss/Execute Decoupling.

2nd Workshop on Memory Access Decoupling for Superscalar and Multiple Issue Architectures (**MEDEA'00**).

Oct. 19, 2000.

New Techniques for Exploiting Program Structure and Behavior in Computer Architecture.

2nd International Workshop on Innovative Architecture (**IWIA'98**). Oct. 26, 1998.

Dependence-Based Prefetching for Linked Data Structures.

8th International Conference on Architectural Support for Programming Languages and Operating Systems (**ASPLOS'98**). Oct. 5, 1998.

Exploiting Dead Value Information.

30th International Symposium on Microarchitecture (**MICRO'97**). Dec. 2, 1997.

Refereed Journal Publications

SMT-Directory: Efficient Load-Load Ordering for SMT

Andrew Hilton and Amir Roth

Computer Architecture Letters (**CAL'10**), Volume 12, Number 2, May 2010.

iCFP: Tolerating All-Level Cache Misses in In-Order Processors

Andrew Hilton, Santosh Nagarakatte, and Amir Roth.

IEEE MICRO (**IEEE MICRO'10**). Volume 30, Number 1, Jan/Feb. 2010.

Physical Register Reference Counting

Amir Roth

Computer Architecture Letters (**CAL'08**). Volume 8, Number 1, Jan. 2008.

NoSQ: Store-Load Communication without a Store Queue

Tingting Sha, Milo M. K. Martin, and Amir Roth.

IEEE MICRO (**IEEE MICRO'07**). Volume 27, Number 1, Jan/Feb. 2007.

Store Vulnerability Window (SVW): A Filter and Potential Replacement for Load Re-Execution

Amir Roth.

Journal of Instruction Level Parallelism (**JILP'06**). Volume 8, Sep. 2006. (<http://www.jilp.org/vol8>)

The Implementation and Evaluation of Dynamic Code Decompression Using DISE

Marc Corliss, E Christopher Lewis, and Amir Roth.

Invited to ACM Transactions on Embedded Computer Systems (**TECS'03**). Oct. 2003.

Squash Reuse via a Simplified Implementation of Register Integration.

Amir Roth and Gurindar S. Sohi.

Journal of Instruction Level Parallelism (**JILP'02**). Volume 3, 2002. (<http://www.jilp.org/vol2/>)

Dynamic Methods for Load and Load-Dependent Scheduling.

Amir Roth, Avi Mendelson, and Ronny Ronen.

Invited to IEEE Special Issue on Microprocessor Architecture and Compiler Technology. Nov., 2001.

Speculative Multithreaded Processors.

Amir Roth and Gurindar S. Sohi.

In IEEE Computer (**IEEE COMPUTER'01**). Volume 34, number 4, Apr. 2001.

Refereed Conference Publications

BOLT: Energy-Efficient Out-of-Order Latency-Tolerant Execution

Andrew Hilton and Amir Roth

Proc. 16th International Symposium on High-Performance Computer Architecture (**HPCA'10**), pages 343–354, Jan. 12–14, 2010.

CPROB: Checkpoint Processing with Opportunistic Minimal Recovery

Andrew Hilton, Neeraj Eswaran and Amir Roth.

Proc. 18th International Conference on Parallel Architectures and Compilation Techniques (**PACT'09**), pages 159–168, Sep. 14–16, 2009.

Decoupled Store Completion/Silent Deterministic Replay: Enabling Scalable Data Memory for CPR/CFP Processors

Andrew Hilton and Amir Roth.

Proc. 36th International Symposium on Computer Architecture (**ISCA'09**), pages 245–254, Jun. 22–24, 2009.

iCFP: Tolerating All-Level Cache Misses in In-Order Processors

Andrew Hilton and Amir Roth.

Proc. 15th International Symposium on High Performance Computer Architecture (**HPCA'09**), pages 431–442, Feb. 16–18, 2009.

Ginger: Control Independence Using Tag Rewriting

Andrew Hilton and Amir Roth.

Proc. 34th International Symposium on Computer Architecture (**ISCA'07**), pages 436–447, Jun. 11–13, 2007.
46 out of 204 submissions accepted (23%)

NoSQ: Store-Load Forwarding without a Store Queue.

Tingting Sha, Milo M. K. Martin, and Amir Roth.

Proc. 39th International Symposium on Microarchitecture (**MICRO'06**), pages 285–296, Dec. 11–13, 2006.
42 of 174 submissions accepted (24%).

Serialization Aware Mini-Graph Selection.

Anne Bracy and Amir Roth.

Proc. 39th International Symposium on Microarchitecture (**MICRO'06**), pages 171–182, Dec. 11–13, 2006.
42 of 174 submissions accepted (24%).

Scalable Store-Load Forwarding via Store Queue Index Prediction.

Tingting Sha, Milo M. K. Martin, and Amir Roth.

Proc. 38th International Symposium on Microarchitecture (**MICRO'05**), pages 159–170, Nov. 12–16, 2005.
29 of 147 submissions accepted (20%).

RENO: A Rename Based Instruction Optimizer.

Vlad Petric, Tingting Sha, and Amir Roth.

Proc. 32nd International Symposium on Computer Architecture (**ISCA'05**), pages 98–102, Jun 4–8, 2005.
45 of 194 submissions accepted (23%).

One of only nine people in 33 years of ISCA with three papers in one conference.

Energy Aspects of Pre-Execution and Energy Aware P-Thread Selection.

Vlad Petric and Amir Roth.

Proc. 32nd International Symposium on Computer Architecture (**ISCA'05**), pages 322–333, Jun 4–8, 2005.
45 of 194 submissions accepted (23%).

Store Vulnerability Window (SVW): Re-Execution Filtering for Enhanced Load Optimization.

Amir Roth.

Proc. 32nd International Symposium on Computer Architecture (**ISCA'05**), pages 458–468, Jun 4–8, 2005.
45 of 194 submissions accepted (23%).

Low Overhead Interactive Debugging via Dynamic Instrumentation with DISE.

Marc Corliss, E Christopher Lewis, and Amir Roth.

Proc. 11th International Symposium on High-Performance Computer Architecture (**HPCA'05**), pages 303–314,
Feb. 12–16, 2005.

28 of 181 submissions accepted (15%)

Dataflow Minigraphs: Amplifying Superscalar Capacity and Bandwidth.

Anne Bracy, Prashant Prahlaad, and Amir Roth.

Proc. 37th International Symposium on Microarchitecture (**MICRO'04**), pages 18–29, Dec. 4–8, 2004.
29 of 158 submissions accepted (18%).

A DISE Implementation of Dynamic Code Decompression.

Marc L. Corliss, E Christopher Lewis, and Amir Roth.

Proc. 2003 Symposium on Languages, Compilers and Tools for Embedded Systems (**LCTES'03**), pages 232–
243, Jun. 11–13, 2003.

29 of 128 submissions accepted (23%).

DISE: A Programmable Macro-Engine for Customizing Applications.

Marc L. Corliss, E Christopher Lewis, and Amir Roth.

Proc. 30th International Symposium on Computer Architecture (**ISCA'03**), pages 362–373, Jun. 9–11, 2003.
37 of 184 submissions accepted (20%).

Three Extensions to Register Integration.

Vlad Petric, Anne Bracy, and Amir Roth.

Proc. 35th International Symposium on Microarchitecture (**MICRO'02**), pages 37–47, Nov. 18–22, 2002.

36 of 150 submissions accepted (24%)

Anne Bracy wins best student presentation award.

A Quantitative Framework for Pre-Execution Thread Selection.

Amir Roth and Gurindar S. Sohi.

Proc. 35th International Symposium on Microarchitecture (**MICRO'02**), pages 430–441, Nov. 18–22, 2002.

36 of 150 submissions accepted (24%)

Speculative Data-Driven Multithreading.

Amir Roth and Gurindar S. Sohi.

Proc. 7th International Symposium on High Performance Computer Architecture (**HPCA'01**), pages 37–48, Jan. 20–24, 2001.

26 of 110 submissions accepted (24%).

Speculative Multithreaded Processors.

Gurindar S. Sohi and Amir Roth.

Proc. 7th International Conference on High Performance Computing (**HiPC'00**). Dec. 17–20, 2000.

Register Integration: A Simple and Efficient Implementation of Squash Reuse.

Amir Roth and Gurindar S. Sohi.

Proc. 33rd International Symposium on Microarchitecture (**MICRO'00**), pages 223–234, Dec. 10–13, 2000.

31 of 110 submissions accepted (28%).

Improving Virtual Function Target Prediction via Dependence Based Pre-Computation.

Amir Roth, Andreas Moshovos and Gurindar S. Sohi.

Proc. 1999 ACM International Conference on Supercomputing (**ICS'99**), pages 356–364, June 21–25, 1999.

57 of 180 submissions accepted (32%).

Effective Jump Pointer Prefetching for Linked Data Structures.

Amir Roth and Gurindar S. Sohi.

Proc. 26th International Symposium on Computer Architecture (**ISCA'99**), pages 111–121, May 2–4, 1999.

26 of 135 submissions accepted (19%).

Dependence-Based Prefetching for Linked Data Structures.

Amir Roth, Andreas Moshovos and Gurindar S. Sohi.

Proc. 8th International Conference on Architectural Support for Programming Languages and Operating Systems (**ASPLOS'98**), pages 115–126, Oct. 4–7, 1998.

28 of 123 submissions accepted (23%).

Exploiting Dead Value Information.

Milo M. Martin, Amir Roth and Charles N. Fischer.

Proc. 30th International Symposium on Microarchitecture (**MICRO'97**), pages 125–135, Dec. 1–3, 1997.

35 of 103 submissions accepted (34%).

Refereed Workshop Publications

FIESTA: A Sample-Balanced Multi-Program Workload Methodology

Andrew Hilton, Neeraj Eswaran, and Amir Roth.

Proc. 5th Workshop on Modeling, Benchmarking, and Simulation (**MoBS'09**), Jun 21, 2009.

Using DISE to Protect Return Addresses from Attack.

Marc Corliss, E Christopher Lewis, and Amir Roth.

Proc. 1st Workshop on Architectural Support for Security and Anti-Virus (**WASSA'04**), Oct. 9, 2004.

Microarchitectural Miss/Execute Decoupling.

Amir Roth, Craig B. Zilles and Gurindar S. Sohi.

Proc. 2nd Workshop on Memory Access Decoupling for Superscalar and Multiple Issue Architectures (MEDEA'00). Oct. 19, 2000.

New Techniques for Exploiting Program Structure and Behavior in Computer Architecture.

Gurindar S. Sohi and Amir Roth.

Proc. 2nd International Workshop on Innovative Architecture (IWIA'98). Oct. 26–28, 1998.

Technical Reports (not published elsewhere)

Encoding Mini-Graphs using Handle Prefix Outlining.

Anne Bracy and Amir Roth.

University of Pennsylvania Technical Report #MS-CIS-08-27, Aug. 2008.

A High-Bandwidth Load-Store Unit for Single- and Multi- Threaded Processors.

Amir Roth.

University of Pennsylvania Technical Report #MS-CIS-04-09, Jun. 2004.

A Static, Profile-Driven Approach to Energy-Efficient Register Integration.

Anne Bracy, Vlad Petric and Amir Roth.

University of Pennsylvania Technical Report #MS-CIS-03-35, Dec. 2003.