

Bianca Schroeder

Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213-3891
bianca@cs.cmu.edu; (412) 608-4675
www.cs.cmu.edu/~bianca

RESEARCH INTERESTS

My research focuses on the design and implementation of computer systems. The methods I am using in my work are inspired by a broad array of disciplines, including performance modeling and analysis, workload and fault characterization, machine learning, and scheduling and queueing theory. My work spans a number of different areas in computer systems, including high-performance computing systems, web servers, computer networks, database systems and storage systems.

EDUCATION

- 2005-present Carnegie Mellon University, Pittsburgh.
Post-doctoral researcher in the Computer Science Department.
Topic: *Empirical system reliability*
Advisor: Garth A. Gibson.
- 1999-2005 Carnegie Mellon University, Pittsburgh.
Doctoral Candidate in the Department of Computer Science.
Thesis: *Improving the performance of static and dynamic requests at a busy web site*
Advisor: Mor Harchol-Balter.
- 1996-1997 University of Limerick, Limerick, Ireland.
Exchange student in the Computer Science Program.
- 1993-1999 Universitaet des Saarlandes, Saarbruecken, Germany.
M.S. in Computer Science, August 1999
Thesis: *Upper and lower bounds for online scheduling*
Advisors: Kurt Mehlhorn and Susanne Albers

HONORS

- 2007 Best paper award 5th Usenix Conference on File and Storage Technologies (FAST'07)
- 2003-2005 Two-time winner of IBM Phd fellowship
- 2003 Best paper award 18th International Teletraffic Congress (ITC 2003)
- 1996-1997 Erasmus fellowship for studies abroad

PUBLICATIONS

- [2007] Bianca Schroeder, Garth A. Gibson. “Disk failures in the real world: What does an MTTF of 1,000,000 hours mean to you?” *5th Usenix Conference on File and Storage Technologies (FAST 2007)*. **Winner of best paper award.**
- [2007] Ernst Biersack, Bianca Schroeder. “Scheduling in Practice.” Invited to *Special Issue of the ACM Sigmetrics PER (Performance Evaluation Review)* on “New Perspectives in Scheduling”.
- [2006] Bianca Schroeder, Garth A. Gibson. “A large scale study of failures in high-performance-computing systems.” *International Symposium on Dependable Systems and Networks (DSN 2006)*.
As **one of the best papers of DSN’06** invited to *IEEE Transactions on Dependable and Secure Computing (TDSC)*.
- [2006] Bianca Schroeder, Adam Wierman and Mor Harchol-Balter. “Open vs. closed: A cautionary tale.” *3rd Symposium on Networked System Design and Implementation (NSDI 2006)*.
- [2006] Bianca Schroeder, Mor Harchol-Balter, Arun Iyengar, Erich Nahum. “Achieving class-based QoS for transactional workloads.” *22th International Conference on Data Engineering (ICDE 2006)*.
- [2006] Bianca Schroeder, Mor Harchol-Balter, Arun Iyengar, Erich Nahum. “How to determine a good multi-programming level for external scheduling.” *22th International Conference on Data Engineering (ICDE 2006)*.
- [2005] David T. McWherter, Bianca Schroeder, Anastassia Ailamaki and Mor Harchol-Balter. “Improving Preemptive Prioritization via Statistical Characterization of OLTP Locking.” *21th International Conference on Data Engineering (ICDE 2005)*.
- [2004] David T. McWherter, Bianca Schroeder, Anastassia Ailamaki and Mor Harchol-Balter. “Priority Mechanisms for OLTP and Transactional Web Applications.” *20th International Conference on Data Engineering (ICDE 2004)*. **Has been cited 30+ times.**
- [2003] Bianca Schroeder and Mor Harchol-Balter. “Web servers under overload: How scheduling can help.” *18th International Teletraffic Congress (ITC 2003)*. **Winner of student paper award.**
Extended version appeared in *ACM Transactions on Internet Technologies (TOIT)*, 6(1): 20-52, 2006.
- [2003] A. Nucci, B. Schroeder, S. Bhattacharyya, N. Taft, C. Diot. “IS-IS Link Weight Assignment for Transient Link Failures.” *18th International Teletraffic Congress (ITC 2003)*. Berlin, Germany, 2003.
- [2003] Mor Harchol-Balter, Bianca Schroeder, Nikhil Bansal, Mukesh Agrawal. “Size-based Scheduling to Improve Web Performance.” *ACM Transactions on Computer Systems (TOCS)*, 21(2): 207-233.
Has been cited 100+ times.
- [2001] Mor Harchol-Balter, Nikhil Bansal, Bianca Schroeder. “Implementation of SRPT Scheduling in Web Servers.” *Job Scheduling Strategies for Parallel Processing, 7th International Workshop (JSSPP 2001)*.
- [2000] Bianca Schroeder, Mor Harchol-Balter. “Evaluation of Task Assignment Policies for Supercomputing Servers: The Case for Load Unbalancing and Fairness.” *9th IEEE Symposium on High Performance Distributed Computing (HPDC’00)*.
As **one of the best papers of HPDC’00** invited to *Cluster Computing* 7(2): 151-161 (2004).

- [2000] Susanne Albers, Bianca Schroeder. “An Experimental Study of Online Scheduling Algorithms.” *4th International Workshop on Algorithm Engineering (WAE’00)*.
As **one of the best papers of WAE’00** selected to appear in *ACM Journal of Experimental Algorithms* 7(3) (2002).

BOOK CHAPTERS

- [2005] Arun Iyengar, Lakshmi Ramaswamy, and Bianca Schroeder. “Techniques for efficiently serving and caching dynamic web content.” In “Recent Advances on Web Data Delivery” by S. Chanson, X. Tang, J. Xu. Kluwer Academic Publisher, 2005.
- [2005] Anastassia Ailamaki, Sailesh Krishnamurthy, Spiros Papadimitriou, and Bianca Schroeder. “The PostgreSQL Open Source DBMS.” In “Database System Concepts” by Abraham Silberschatz, Henry F. Korth, S. Sudarshan, 5th Edition. McGraw-Hill Book Company, 2005.

CONFERENCE TALKS

- November 2006 Workshop on Petascale Data Storage at SC’06
“Learning to live with our failures”.
- November 2006 The 7th Usenix Symposium on Operating Systems Design and Implementation (OSDI 2006)
Work-in-Progress session. “Failures in the real world.”
- June 2006 The International Conference on Dependable Systems and Networks (DSN 2006)
“ A Large-Scale Study of Failures in High-Performance-Computing Systems.”
- May 2006 3rd Symposium on Networked System Design and Implementation (NSDI 2006)
“Closed versus open system models: Understanding their impact on performance evaluation and system design”.
- April 2006 22th International Conference on Data Engineering (ICDE 2006)
“How to determine a good multi-programming level for external scheduling.”
- November 2005 Workshop on Dependability Benchmarking at the 16th IEEE International Symposium on Software Reliability Engineering (ISSRE 2005)
“Analyzing failure data from large HPC clusters”.
- October 2004 Grace Hopper conference for women in computing
“Improving the performance of static and dynamic requests at a busy web server.”
- May 2004 CORS/INFORMS joint conference 2004
“Scheduling web servers.”
- April 2004 20th International Conference on Data Engineering (ICDE 2004)
“Priority Mechanisms for OLTP and Transactional Web Applications.”
- September 2003 18th International Teletraffic Congress (ITC 2003)
“Web servers under overload: How scheduling can help.”
- September 2000 4th International Workshop on Algorithm Engineering (WAE’00)
“An Experimental Study of Online Scheduling Algorithms.”

August 2000 9th IEEE Symposium on High Performance Distributed Computing (HPDC'00)
"Evaluation of Task Assignment Policies for Supercomputing Servers:
The Case for Load Unbalancing and Fairness."

INVITED TALKS

November 2006 University of Washington. Host: Hank Levy.
"Failures in the real world".

November 2006 Invited talk in the ASC booth at SC'06.
(ASC is a program of the DOE's National Nuclear Security Administration (NNSA)).
"Failures in the real world: Collecting, sharing, and analyzing failure data".

October 2006 Google Inc., Mountain View, CA. Hosts: Luiz Barroso, Eduardo Pinheiro.
"Failures in the real world: Collecting, sharing, and analyzing failure data".

October 2006 IBM Almaden, San Jose, CA. Host: Frank Schmuck.
"Failures in the real world: Collecting, sharing, and analyzing failure data".

August 2006 HEC-IWG File Systems and I/O R&D Workshop, Washington D.C. 2006
"The failure data usage project". Host: Gary Grider

June 2006 University of California, Berkeley. Host: Armando Fox.
"Understanding failure at scale".

June 2006 Hewlett Packard Laboratories, Palo Alto, CA. Host: Kim Keeton.
"Understanding failure at scale".

June 2006 Microsoft Research, Mountain View, CA. Host: Chandu Thekkath.
"Understanding failure at scale".

May 2004 University of Calgary. Host: Carey Williamson.
"Scheduling web servers: Theory and practice".

April 2004 Selected as one of two PhD students to give research presentation at
CMU open house for prospective students. "QoS for databases".

March 2004 Boston University, Networks seminar.
"QoS for online shopping".

August 2003 IBM TJ Watson Research Center. Host: Arun Iyengar.
"Priority Mechanisms for OLTP and Transactional Web Applications".

July 2001 Sprint Advanced Technology Laboratories, Burlingame, CA. Host: Christophe Diot.
"Improving Performance of Web Servers under Overload."

June 2001 Stanford University. Hosts: Nick McKeown and Balaji Prabhakar.
"Size-based Scheduling to Improve Web Performance".

PATENTS

- 2004 IBM patent filed: A. Iyengar, E. Nahum, and B. Schroeder.
“Method for Dynamically Scheduling Requests”.
- 2003 Sprint patent filed: S. Bhattacharyya, A. Nucci, N. Taft, B. Schroeder and C. Diot.
“Method for Assigning Link Weights in a Communications Network”.

PROFESSIONAL SERVICE

- 2007 Program committee member: WWW 2007.
- 2006 – pres. Head of a steering committee of researchers from seven organizations with the goal of creating a Usenix hosted, public failure data repository.
- 2002 – 2006 Referee for several journals and conferences, including: TOCS, FAST, SIGMETRICS, VLDB, ASPLOS, SC, OSDI, WWW, EuroSys, TWEB.

GRANT SUPPORT

- 2006 – 2007 Co-PI on SciDAC grant awarded by DOE to a collaboration of researchers at three universities and five national laboratories, led by CMU. Amount: \$11,000,000
“The Petascale Data Storage Institute (PDSI)”.
- 2005 – 2006 Co-PI on grant from the Technology Cooperative of Pittsburgh. Amount: \$188,550
“External QoS Management Systems for Backend Database Servers”.

TEACHING EXPERIENCE

- Spring 2006 Instructor, Carnegie Mellon University
Designed and co-taught a brand-new graduate course on
“Manageability and Dependability of IT systems”.
Instructor reviews averaged 5/5. Course reviews averaged 4.5/5.
- Spring 2002 Teaching Assistant, Carnegie Mellon University
Graduate level “Computer Networking” class
Taught by: Srinu Seshan
Reviews averaged 4.6/5.
- Fall 2000 Teaching Assistant, Carnegie Mellon University
Undergraduate level “Introduction to Computer Systems” class
Taught by: David O’Halloran and Randy Bryant
Reviews averaged 4.5/5.
- Fall 1997 Teaching Assistant, Max-Planck-Institute, Germany
“Combinatorial Optimization” class for computer science Masters students
Taught by: Torben Hagerup

RESEARCH EMPLOYMENT

- Summer 2004 Internship IBM TJ Watson Research Center.
Project: *Analysis of Web workload parameters and their impact on performance and capacity planning.*
Hosted by: Arun Iyengar and Erich Nahum
- Summer 2003 Internship IBM TJ Watson Research Center.
Project: *Developed external scheduler for providing QoS for backend servers.*
Hosted by: Arun Iyengar and Erich Nahum
- Summer 2001 Internship Sprint Advanced Technology Laboratories, Burlingame, CA.
Project: *Using techniques from combinatorial optimization to optimize OSPF and IS-IS link weights for robustness in the case of link failure.*
Hosted by: Christophe Diot, Nina Taft and Supratik Bhattacharyya
- Spring 1995 Research Assistant, Universitaet des Saarlandes, Germany.
Project: *Developed indexing tool for medieval Latin texts in collaboration with the German Academy for Art and Literature.*
Hosted by: Guenther Hotz

PERSONAL

Citizenship: German.
Visa status: J-1 (no home-stay requirement).

REFERENCES

Prof. Mor Harchol-Balter
Computer Science Department
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3891
harchol@cs.cmu.edu; (412) 268-7893

Prof. Garth A. Gibson
Computer Science Department
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3891
garth.gibson@cs.cmu.edu; (412) 268-5890

Prof. Greg Ganger
Electrical & Computer Engineering
Carnegie Mellon University
Collaborative Innovation Center 2208
5000 Forbes Ave.
Pittsburgh, PA 15213
ganger@ece.cmu.edu; (412) 268-1297

Prof. Todd C. Mowry
Computer Science Department
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3891
tcm@cs.cmu.edu
(412) 268-3725 or (412) 297-4141

Dr. John Wilkes
Hewlett Packard Laboratories
1501 Page Mill Road, MS 1134
Palo Alto, CA 94304-1100
john.wilkes@hp.com; (650) 857-3568