NITIN AGRAWAL

Office: 1210 W. Dayton St, Madison, WI 53706
Email: nitina@cs.wisc.edu
Home: 1132 E. Johnson St, Apt 7, Madison, WI 53703
Cell Phone: (608) 335-7845
Webpage: http://pages.cs.wisc.edu/~nitina
Office Phone: (608) 262-2252

RESEARCH INTERESTS

Operating Systems, Distributed Systems, File and Storage Systems, SSDs

EDUCATION

UNIVERSITY OF WISCONSIN, MADISON
Ph.D. (Computer Sciences), expected May 2009.
Dissertation Title: Enabling Realistic and Practical File-System Benchmarking
Advisors: Prof. Andrea C. Arpaci-Dusseau and Prof. Remzi H. Arpaci-Dusseau
M.S. (Computer Sciences), May 2005. CGPA: 3.94/4

INSTITUTE OF TECHNOLOGY, BHU, INDIA
B.Tech. (Computer Science and Engineering), May 2003. CGPA: 3.9/4, Majors GPA: 3.97/4
Awarded the University Silver Medal

CONFERENCE PUBLICATIONS

FAST '09 — Nitin Agrawal, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. "Generating Realistic Impressions for File-System Benchmarking". *Proceedings of the Seventh Conference on File and Storage Technologies, San Francisco, CA. Feb 2009.* **Best Paper Award.**

USENIX '08 — Nitin Agrawal, Vijayan Prabhakaran, Ted Wobber, John D. Davis, Mark Manasse, Rina Panigrahy. "Design Tradeoffs for SSD Performance". *Proceedings of the USENIX Annual Technical Conference, Boston, MA. June 2008.*

HotMetrics '08 — Nitin Agrawal, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. "Towards Realistic File-System Benchmarks with CodeMRI". *Proceedings of the ACM SIGMETRICS Workshop on Hot Topics in Measurement and Modeling of Computer Systems, Annapolis, MD. June 2008.*

DSN '08 — Lakshmi Bairavasundaram, Meenali Rungta, Nitin Agrawal, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau, Michael M. Swift. "Analyzing the Effects of Disk Pointer Corruption". *Proceedings of the Conference on Dependable Systems and Networks, Anchorage, AK. June 2008.*

FAST '07 — Nitin Agrawal, William J. Bolosky, John R. Douceur, Jacob R. Lorch. "A Five-Year Study of File-System Metadata". *Proceedings of the Fifth Conference on File and Storage Technologies, San Jose, CA. Feb 2007.* **Selected as a top paper and forwarded to ACM TOS.**

SOSP '05 — Vijayan Prabhakaran, Lakshmi Bairavasundaram, Nitin Agrawal, Haryadi S. Gunawi, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. "IRON File Systems". *Proceedings of the 20th Symposium on Operating Systems Principles, Brighton, UK. Oct 2005.*

ISCA '05 — Haryadi S. Gunawi, Nitin Agrawal, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau, Jiri Schindler. "Deconstructing Commodity Storage Clusters". *Proceedings of 32nd International Symposium on Computer Architecture, Madison, WI, June 2005.*

IN PREPARATION

SUBMIT '09 — Nitin Agrawal, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. "Compressions: Enabling File System Benchmarking at Scale".

JOURNAL PUBLICATIONS

PER '08 — Nitin Agrawal, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. "Towards Realistic File-System Benchmarks with CodeMRI". *Special Issue of ACM SIGMETRICS Performance Evaluation Review, Volume 36, Issue 2, Sep 2008.*

TOS '07 — Nitin Agrawal, William J. Bolosky, John R. Douceur, Jacob R. Lorch. "A Five-Year Study of File-System Metadata". *ACM Transactions on Storage, Vol 3, Issue 3, 2007.*

OTHER PUBLICATIONS

FAST '09 — Nitin Agrawal, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. "Speedy and Scalable File-System Benchmarking with Compressions". *Seventh Conference on File and Storage Technologies (WIP Session), San Francisco, CA. Feb 2009.*

USENIX '06 — Nitin Agrawal, Haryadi S. Gunawi, Vijayan Prabhakaran, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. "Separating Policy and Mechanism for Failure Handling in Commodity File Systems". *USENIX Annual Technical Conference (Poster Session), Boston, MA. June 2006.*

UW-CS '04 — Nitin Agrawal. "Interference-Aware Scheduling in Storage Systems". *Master's Thesis*

IT-BHU '03 — Nitin Agrawal. "Symbolic Rule-Extraction from Artificial Neural Networks". *Bachelor's Thesis*

RESEARCH

Enabling Realistic and Practical File-System Benchmarking (Dissertation Research): The file and storage systems that store and retrieve our important data play an essential role in our computer systems. In spite of tremendous advances in file system design, the approaches for benchmarking them still lag far behind. My dissertation research bridges this gap with four contributions:

- **Characterizing Real File System Images:** I performed a large scale analysis of file-system metadata snapshots collected over a five-year period. The metadata snapshots were used to study temporal changes in file size, file age, namespace structure etc. We gave consequent lessons for designers of file systems and related software, and presented a generative model that explained the namespace structure and the distribution of directory sizes [FAST '07, TOS '07].

- **Generating Realistic File-System Images with Impressions:** Unfortunately, system designers continue to rely on ad hoc assumptions and (often inaccurate) rules of thumb when generating file-system images. To remedy these problems, I developed Impressions, a framework to generate statistically accurate file-system images with realistic metadata and file content. Impressions gives the user flexibility to specify one or more parameters from a detailed list of file system parameters (file-system size, number of files, distribution of file sizes, etc.). Impressions incorporates statistical techniques (automatic curve-fitting, resolving multiple constraints, interpolation and extrapolation, etc.) and uses statistical tests for goodness-of-fit to ensure the accuracy of the image [FAST '09].

- **Generating Realistic File-System Benchmarks with CodeMRI:** Synthetic file system benchmarks are widely used, but largely based on the benchmark writer's interpretation of the real workload. This approximation is insufficient since even a simple operation through the API may end up exercising the file system in very different ways due to caching and prefetching. I took first steps in creating "realistic synthetic" benchmarks by building a system called CodeMRI. It uses file-system domain knowledge and a small amount of system profiling to better understand how the benchmark is stressing the system and to deconstruct its workload [HotMetrics '08, PER '08].

- **Scalable File-System Benchmarking with Compressions:** The final part of my thesis addresses the problem of scalable benchmarking. In order to benchmark large disk partitions, the set up required is often cumbersome and the benchmark takes a huge amount of time to finish. To address this challenge, I am developing *Compressions*, a system that scales down the requirements of the benchmark by requiring only a small fraction (1 to 10%) of the original partition size to run the workload. Through this transformation, Compressions also reduces the total runtime to a fraction of the original. In order to deliver accurate benchmarking results, Compressions maintains a statistical timing model to account for the remapping and data squashing [SUBMIT '09].

Solid-State Storage Devices: Flash-based SSDs have the potential to change the storage landscape. As an intern at Microsoft Research, I worked on design tradeoffs that are relevant to NAND-flash solid-state storage. Ted Wobber, Vijayan Prabhakaran and I analyzed several of these tradeoffs using a trace based disk simulator that we built to characterize different SSD organizations. We developed algorithms for cleaning and wear-leveling flash-media to make it viable for use in environments with high I/O rates, along with improvements in the performance of random writes. Our analysis was driven by various traces captured from running systems such as a full-scale TPC-C benchmark, an Exchange server workload, and various standard file system benchmarks. From our analysis, we found that SSD performance and lifetime is highly workload-sensitive, and that complex systems problems that normally appear higher in the storage stack, or even in distributed systems, are relevant to device firmware [USENIX '08].

Reliability in the Storage Stack: Hardware constituents of the storage stack (storage devices, interconnects, etc) fail and software (device firmware, drivers, file systems) exhibits bugs and other inconsistencies, leading to data loss and corruption. I have worked on techniques to help understand causes of failures in the storage stack:

- Analyzed how commodity file systems handle disk failures. I built a type-aware fault injection framework for the Reiser file system and performed analysis of its fault handling towards partial disk failures. I found that the most prominent aspect of the failure handling policy of ReiserFS was its tendency to `panic` the system upon detection of virtually any write failure, while paying close attention to error codes across reads and writes. It also exhibited several inconsistencies and bugs [SOSP '05].
- File systems demonstrate inconsistent and inadequate handling of latent sector errors and other partial disk failures. In order to identify the root causes of the observed inconsistencies, I developed *Differential Failure Analysis*, a combination of static and run-time analysis to achieve a thorough understanding of the failure handling characteristics of file-system source code [USENIX '06].
- Developed and applied type-aware corruption to understand the effects of disk-pointer corruption on file-system reliability, using Windows NTFS and Linux Ext3 as case studies. I analyzed the Ext3 file system and found that it does not recover from most corruptions, including many scenarios for which sufficient redundant information is available leading to further corruption, crashes, and unmountable file systems [DSN '08].
- Evaluated failure handling of SCSI drivers by injecting faults at the lowest level of the tiered SCSI architecture and observing the detection and recovery mechanisms employed by the upper driver levels.

Commodity Storage Clusters: High-end storage systems are increasingly being built using commodity components. We designed techniques for characterizing complex storage clusters in the context of the EMC Centera storage system. By correlating disk and network traffic with the running workload using *observation* and *delay*, we inferred the structure of the software system as well as its policies (e.g., how it performs caching, replication, load-balancing) without any access to the source code [ISCA '05].

Regulatory Compliant Storage: The introduction of federal regulations such as Sarbanes-Oxley and HIPAA mandate stricter enforcing of data retention, access and tampering guidelines. I worked on an auditing framework to enforce regulatory compliance on archival storage. The focus was to provide continuous verification of system state and support feature-rich querying. As part of a larger project on compliant storage at IBM Almaden, I designed and built a prototype for the auditing engine.

RESEARCH DELIVERABLES

- **Impressions Framework.** I will be releasing the *Impressions* software that generates statistically accurate file-system images. Please visit *http://www.cs.wisc.eu/adsl/impressions/* for the latest status.
- **Five Year Metadata Dataset.** As an intern at Microsoft Research, I contributed towards the release of a large scale file-system metadata dataset along with Bill Bolosky, John Douceur and Jay Lorch. The dataset was collected over a span of five years from more than 63,000 distinct Windows file systems. These systems contained 4 billion files totaling 700TB of file data. Available at SNIA's IOTTA repository. *http://iotta.snia.org/traces/tracesStaticSnapshot*.
- **SSD Simulator for Disksim.** As an intern at Microsoft Research, I contributed towards the design of a simulator for solid state devices, along with Ted Wobber and Vijayan Prabhakaran. It is now available as an add-on to the Disksim Simulation Environment. *http://research.microsoft.com/research/downloads/*.

WORK EXPERIENCE

• **Research Assistant, UW Madison**, Computer Sciences Department, Aug 2004-present: I work as part of the Advanced Systems Lab (ADSL).

• **Research Intern, Microsoft Research, Silicon Valley Lab**, June-Aug 2007: Worked on flash-based solid state storage devices with Ted Wobber, Andrew Birrell and Chuck Thacker.

• **Research Intern, Microsoft Research, Redmond** (Systems and Networking Group), May-Aug 2005: Performed a large scale analysis of file system metadata with Bill Bolosky, John Douceur and Jay Lorch.

• **Research Intern, IBM Almaden Research Center, San Jose** (Storage Systems Group), May-Aug 2004: Worked on an auditing framework for regulatory compliant archival storage with Ying Chen.

• **Teaching Assistant, UW Madison**, Computer Sciences Department, Aug 2003-May 2004: Designed assignments and examinations, and led discussions for the course on Machine Organization and Programming.

• **Research Intern, University of Dortmund, Germany** (Artificial Intelligence Group), May-July 2002: As part of the MiningMart project, designed a chain of preprocessing operators for a data mining case-base with SVM as the learning step. The work resulted in the first successful MiningMart case and is part of a downloadable case-base.

• **Summer Intern, Tata Consultancy Services (TCS), India**, June-July 2001: The project was based in Java, Oracle and JDBC. It involved developing an automated and interactive system for project management review.

SELECTED PRESENTATIONS

• *Generating Realistic Impressions for File-System Benchmarking*
FAST '09 (Feb 2009)

• *Design Tradeoffs for SSD Performance*
USENIX '08 (June 2008); OS Seminar, UW-Madison (June 2008)

• *Towards Realistic File-System Benchmarks with CodeMRI*
ACM HotMetrics '08 (June 2008)

• *Content-Centric File-System Evaluation*
Microsoft Research India (January 2008)

• *High-Performance Flash-Drives for I/O Intensive Workloads*
Microsoft Research Silicon Valley Lab (August 2007)

• *Code-Aware Benchmarking*
OS Seminar, UW-Madison (April 2007)

• *A Five-Year Study of File-System Metadata*
FAST '07 (Feb 2007); OS Seminar, UW-Madison (Feb 2007)

• *IRON File Systems*
Microsoft Research Redmond (July 2005)

• *Unified Auditing Framework for Regulatory Compliant Archival Storage*
IBM Almaden (August 2004)

HONORS

Best Paper Award at FAST 2009 for Impressions.
"A Five-Year Study of File-System Metadata" forwarded to ACM TOS as one of the top papers in FAST '07.
ACM SIGMETRICS Student Travel Award to attend SIGMETRICS 2008.
USENIX Student Scholarship for OSDI '08, USENIX '08, FAST '08, FAST '07, USENIX '06, NSDI '05, FAST '04.
ACM SIGOPS Student Travel Award to attend SOSP 2005.
ACM SIGARCH Student Award to attend ISCA 2005.
Nominated for the Microsoft Research Fellowship by the University of Wisconsin-Madison, 2005.
BHU University Silver Medal and Merit Award for standing *second* in the graduating batch of 2003.
BHU Merit Scholarship for meritorious performance in B.Tech., 2000-01 and 2001-02.

U.P. State Government Scholarship for merit position in B.Tech. Part I, 2000.
Government of India National Scholarship Certificate for meritorious performance in AISSE (class X)-1997.
Central Board of Secondary Education's merit certificate for being top 0.1% in Mathematics in AISSE 1997.
All India Talent Search Scholarship awarded by INCEF, 1994.

PROFESSIONAL ACTIVITIES

Guest lecturer, Introduction to Operating Systems course, University of Wisconsin Madison, 2009
External Reviewer: FAST '07, ASPLOS '06, Supercomputing '05, USENIX '05
Student Member of ACM and USENIX
Member of the Graduate Mentoring Committee, CS Department, University of Wisconsin Madison, 2006-07
Member of the Graduate Admissions Committee, CS Department, University of Wisconsin Madison, 2003-04
Chairman (2006-07) and Advisor (2007-08), Wisconsin International Students Association.
Invited Member, ISS Student Advisory Committee, University of Wisconsin Madison, 2007-08
Vice Chairman of IEEE (CS) Student's Chapter IT-BHU, 2002-03.
General Secretary of IEE Student's Chapter IT-BHU, 2002-03.
Convener, IEEE SPAVes and Quest, symposia organized by IEEE students branch BHU (sponsored by IEEE USA).

SELECTED PROJECTS

- *SPEAD: Script Programming Environment for Active Disks:* Proposed and evaluated a script programming environment for active disks. SPEAD enables file system and database system to transfer block level disk operations encoded as scripts to the disk (with Prof. Remzi Arpaci-Dusseau).
- *Static Optimization of Multiple Conjunctive Queries with Sliding Windows:* Studied query optimization in the context of streaming information sources with multiple queries. Suggested methods to model the cost of a multi-query plan and chose the best plan accordingly (with Prof. Jeff Naughton and Prof. Raghu Ramakrishnan).
- *I/O Request Scheduling:* Designed Interference-Aware Smallest-Cost-First policy, a request scheduler for storage systems. While existing scheduling policies are oblivious of the location of blocks, and do not distinguish between blocks on-disk or in-cache, iSCF not only takes into account the location of blocks servicing the request but is also aware of the in-cache position of the blocks (Master's Thesis).
- *Technology Development Plan for Neuropeptides:* Proposed a technology development plan for use of Neuropeptides in therapeutics (Strategic Management in Life and Engineering Sciences course with Prof. Anne Miner).
- *IEEE Computer Society's International Design Contest:* Member of a four-person undergraduate team which participated in CSIDC 2003 and reached the pre-final stage. The team designed and implemented a fully functional prototype of an "Intelligent Mailbox" which provided virtual presence of an individual or organization.

REFERENCES

1. Professor Andrea C. Arpaci-Dusseau
Associate Professor
Department of Computer Sciences
University of Wisconsin-Madison
Email: dusseau@cs.wisc.edu
Phone: (608) 265-6013

2. Professor Remzi H. Arpaci-Dusseau
Associate Professor
Department of Computer Sciences
University of Wisconsin-Madison
Email: remzi@cs.wisc.edu
Phone: (608) 263-7764

3. John Douceur
Researcher and Manager
Distributed Systems Group
Microsoft Research, Redmond
Email: johndo@microsoft.com
Phone: (425) 706-8827

4. Professor Michael M. Swift
Assistant Professor
Department of Computer Sciences
University of Wisconsin-Madison
Email: swift@cs.wisc.edu
Phone: (608) 890-0131

MISCELLANEOUS

Visa Status: F1, Country of Citizenship: India Personal Interests: Poetry, Sailing, Tennis