

# Lakshmi N. Bairavasundaram

## Office Address:

1210 W. Dayton St, Madison, WI 53706  
Phone: (608) 262-6624  
E-mail: laksh@cs.wisc.edu

## Home Address:

1813 Keyes Ave, Madison, WI 53711  
Mobile: (608) 320-9339  
Webpage: <http://pages.cs.wisc.edu/~laksh>

---

## Research Interests

Operating systems, File systems, Storage systems, and Fault tolerance.

## Education

### Ph.D. Computer Sciences, expected May 2008

University of Wisconsin-Madison

Advisors: Prof. Andrea C. Arpaci-Dusseau and Prof. Remzi H. Arpaci-Dusseau

Dissertation title: "Characteristics, Impact, and Tolerance of Partial Disk Failures"

### M.S. Computer Sciences, May 2004

University of Wisconsin-Madison

### B.E. Computer Science and Engineering, May 2002

College of Engineering-Guindy, Anna University, Chennai, India.

## Conference Publications

- [FAST08a] Lakshmi N. Bairavasundaram, Garth R. Goodson, Bianca Schroeder, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. "An Analysis of Data Corruption in the Storage Stack". *To appear in the Proceedings of the 6th USENIX Conference on File and Storage Technologies (FAST'08)*. San Jose, CA, February, 2008.
- [FAST08b] Andrew Krioukov, Lakshmi N. Bairavasundaram, Garth R. Goodson, Kiran Srinivasan, Randy Thelen, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. "Parity Lost and Parity Regained". *To appear in the Proceedings of the 6th USENIX Conference on File and Storage Technologies (FAST'08)*. San Jose, CA, February, 2008.
- [SIGMETRICS07] Lakshmi N. Bairavasundaram, Garth R. Goodson, Shankar Pasupathy, Jiri Schindler. "An Analysis of Latent Sector Errors in Disk Drives". *Proceedings of the International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS'07)*. San Diego, CA, June, 2007. **Kenneth C. Sevcik Outstanding Student Paper Award**
- [DSN06] Lakshmi N. Bairavasundaram, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. "Dependability Analysis of Virtual Memory Systems". *Proceedings of the International Conference of Dependable Systems and Networks (DSN'06)*. Philadelphia, PA, June, 2006.
- [FAST05] Muthian Sivathanu, Lakshmi N. Bairavasundaram, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. "Database Aware Semantically-Smart Storage". *Proceedings of the 4th USENIX Conference on File and Storage Technologies (FAST'05)*. San Francisco, CA, December, 2005.
- [SOSP05] Vijayan Prabhakaran, Lakshmi N. Bairavasundaram, Nitin Agrawal, Haryadi S. Gunawi, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. "IRON File Systems". *Proceedings of the 20th ACM Symposium on Operating Systems Principles (SOSP'05)*. Brighton, UK, October, 2005.
- [OSDI04] Muthian Sivathanu, Lakshmi N. Bairavasundaram, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. "Life or Death at Block-Level". *Proceedings of the 6th Symposium on Operating Systems Design and Implementation (OSDI'04)*. San Francisco, CA, December, 2004.

- [ISCA04] Lakshmi N. Bairavasundaram, Muthian Sivathanu, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. “X-RAY: A Non-Invasive Exclusive Caching Mechanism for RAIDs”. *Proceedings of the 31st Annual International Symposium on Computer Architecture (ISCA’04)*. Munich, Germany, June, 2004.
- [HIPC02] R. Vinodh Kumar, B. Lakshmi Narayanan, R. Govindarajan. “Dynamic Path Profile Aided Re-compilation in a JAVA Just-In-Time Compiler”. *Proceedings of the International Conference on High Performance Computing (HiPC’02)*. Bangalore, India, December, 2002.

## Other Publications

- [US07] Lakshmi N. Bairavasundaram, Meenali Rungta, Nitin Agrawal, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau, Michael M. Swift. “Systematically Benchmarking the Effects of Disk Pointer Corruption”. *Under Submission*. December, 2007.
- [StorageSS06] Lakshmi N. Bairavasundaram, Meenali Rungta, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau. “Limiting Trust in the Storage Stack”. *Proceedings of the 2nd International Workshop on Storage Security and Survivability (StorageSS’06)*. Alexandria, VA, October, 2006.
- [PER06] Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau, Lakshmi N. Bairavasundaram, Timothy E. Denehy, Florentina I. Popovici, Vijayan Prabhakaran, Muthian Sivathanu. “Semantically-Smart Disk Systems: Past, Present, and Future”. *Sigmetrics Performance Evaluation Review, Volume 33, Number 4*. March, 2006.
- [HIPC01] Deepak Babu M.I, Lakshmi Narayanan B, Madhu Saravana Sibi G, Ranjani Parthasarathi. “Functional Unit Usage Based Thread Selection in a Simultaneous Multithreaded Processor”. *Poster Session of the International Conference on High Performance Computing (HiPC’01)*. Hyderabad, India, December, 2001.

## Research

### Partial Disk Failures

My research focuses on understanding how computer systems fail, and developing techniques to avoid such failures. Much of the value people place in computers stems from the data stored in them. Hence, my dissertation focuses on failures related to loss or corruption of data. Disk drive failures are the primary causes of data loss. Disk failures are typically partial failures, where some disk sectors are unavailable due to a latent sector error or some disk blocks are silently corrupted. The goals of my dissertation are to understand the characteristics of partial disk failures, analyze how these failures impact components of the storage stack, and develop solutions to tolerate such failures.

*Failure characteristics:* As part of my dissertation, I have analyzed the occurrence and characteristics of latent sector errors and data corruption in a population of 1.53 million disk drives [SIGMETRICS07, FAST08a]:

- The study of latent sector errors found that as many as 20% of nearline (SATA) disk drives may be affected in 2 years of use, and that latent sector errors show high spatial and temporal locality.
- The analysis of data corruption identified very interesting corruption trends including the tendency of consecutive disk blocks to become corrupt, and the non-independence of corruption instances within the same disk and across different disk drives in the same storage system.

*Impact on storage stack:* Given that partial failures could affect a significant percentage of disks, it is important to understand their impact on different elements of the storage stack:

- I have analyzed the impact of such errors on modern commodity file systems (IBM JFS, ext3, ReiserFS, and Windows NTFS) using “type-aware” fault injection techniques [SOSP05, StorageSS06, US07]. The analyses found that even widely-used file systems have bugs in failure handling code, use illogically-inconsistent policies, and do not implement fault tolerance techniques like type-checking and replication effectively.

- I have also investigated the mechanisms used by virtual memory systems (of Linux, FreeBSD, and Windows XP) to tolerate disk errors [DSN06], and found that they are inconsistent and ineffective as well.
- I have applied model checking to data protection schemes used in real parity-based RAID systems and found that they are vulnerable to data loss or corruption due to disk errors such as “lost” writes [FAST08b].

*Tolerance:* The studies above show that even widely-used systems have bugs and cannot be trusted to handle partial disk failures. Therefore, it is important to lower the trust we place in any single system. I am currently developing a file system architecture based on N-version programming principles that reduces the need to trust any one file system.

## Semantically-smart Disk Systems

Prior to my dissertation research, I was involved in designing semantically-smart disk systems (SDS). These systems utilize knowledge of file system data structures inside storage arrays to improve performance, reliability and security. I have used SDS techniques to design an exclusive caching mechanism for multi-level caches involving file system and storage array caches [ISCA04]. I have extended the mechanism to work with a database cache instead of a file system cache [FAST05]. Finally, I have investigated the use of block liveness information inside SDS [OSDI04].

## Work Experience

- Research Assistant.** *Computer Sciences Department, University of Wisconsin-Madison.* Aug 2003 - present  
(Details provided in previous section).
- Summer Intern.** *Advanced Development Group, Network Appliance, Sunnyvale, CA.* Jun - Aug, 2007  
I analyzed the occurrence of silent data corruption in disk drives (details in previous section).
- Summer Intern.** *Advanced Development Group, Network Appliance, Sunnyvale, CA.* Jun - Aug, 2006  
I analyzed data on latent sector errors in disk drives (details in previous section).
- Summer Intern.** *Core Virtualization Research Group, Intel Corporation, Hillsboro, OR.* May - Aug, 2005  
I developed techniques for hypervisor-based fault injection and applied these techniques to study operating system behavior when memory and disk errors occur.
- Summer Intern.** *IBM T.J. Watson Research Center, Hawthorne, NY.* May - Aug, 2004  
I designed and implemented a scalable distributed cache architecture for an enterprise storage system. The architecture uses remote memory (on machines over a high-speed LAN) for caching disk blocks.
- Summer Fellow.** *Jawaharlal Nehru Centre for Advanced Scientific Research, India.* May - Jun, 2001  
I developed an instruction scheduler for a Java Just-in-Time compiler. The scheduler was built to utilize path profile information.

## Selected Projects at UW-Madison

**Understanding the EMC Centera:** In this project, we deconstructed the policies and mechanisms used inside the EMC Centera, a content-addressable storage system. We used “gray-box” techniques such as timing storage system operations and observing network traffic between nodes of the storage array. We identified (i) the synchronous nature of operations within the storage array, (ii) the cache hierarchy inside the array, and (iii) the maximum size of stored blobs. (part of a 2-person team, for the Advanced Computer Networks course, Spring 2004)

**Characterizing multiprocessor memory degree of use:** In this project, we characterized the “degree of use” of memory values (i.e. the number of processors that load a value produced by a processor) in a shared-memory multiprocessor system for popular workloads. Important findings include the observation that the degree of use of memory values correlated with the memory location at which the value was stored and also with the program counter of the instruction producing the value. We used this characterization to propose mechanisms to track the degree of use of values, with the goal of using it for more efficient cache coherence mechanisms. (part of a 3-person team, for the Advanced Computer Architecture II course, Spring 2003)

## Honors and Achievements

- Awarded the Kenneth C. Sevcik Outstanding Student Paper Award at SIGMETRICS 2007.
- Awarded the University Fellowship by the University of Wisconsin-Madison (2002 - 2003).
- Ranked 2nd (by GPA) among the May 2002 Computer Science graduates of Anna University.
- Awarded the Summer Research Fellowship by the Jawaharlal Nehru Centre for Advanced Scientific Research (JNCASR), India (May-June 2001).

## Technical Presentations

- “An Analysis of Data Corruption in the Storage Stack”  
at *Network Appliance*, and *FAST’08* (Upcoming talk)
- “An Analysis of Latent Sector Errors in Disk Drives”  
at *Network Appliance*, *Microsoft Research Silicon Valley*, and *SIGMETRICS’07*
- “Limiting Trust in the Storage Stack”  
at *StorageSS’06*
- “Dependability Analysis of Virtual Memory Systems”  
at *DSN’06*
- “X-RAY: A Non-Invasive Exclusive Caching Mechanism for RAIDs”  
at *IBM T.J Watson Research Center*, and *ISCA’04*

## Professional Activities

- Student member of the ACM.
- Have reviewed papers for the following conferences and journals: *IBM Journal of R&D*, *FAST’07*, *ASPLOS’06*, *USENIX’05*, *ICDCS’04*, *USENIX’04*.

## References

*Available upon request.*

## Miscellaneous

*Country of Citizenship:* India, *Visa Status:* F1

*Personal Interests:* Traveling, Hiking, Ballroom dancing