Andrew W. Leung

Computer Science Department
University of California, Santa Cruz
1156 High Street
Santa Cruz, CA 95064

PHONE: (805) 453-5084
EMAIL: aleung@cs.ucsc.edu
WEB: http://www.cs.ucsc.edu/~aleung

RESEARCH INTERESTS

My current research interests include search and data management for file systems, system benchmarking, security, and peer-to-peer systems.

EDUCATION

Ph. D., Computer Science, University of California, Santa Cruz, Estimated December 2009
Thesis: *Organizing, Indexing, and Searching Large-Scale File Systems*
Advisor: Ethan L Miller

M. S., Computer Science, University of California, Santa Cruz, June 2007
Thesis: *Scalable Security for High Performance, Petascale Storage*
Advisor: Ethan L Miller

B. S., Computer Science, University of California, Santa Barbara, June 2005

EMPLOYMENT HISTORY

October 2008–September 2009 — **Co-Founder**, Pergamum Systems Inc., Santa Cruz, CA. Co-founded a start-up that provides evolvable, reliable, power-efficient, disk-based archival storage.

April 2006–present — **Graduate Student Researcher**, Computer Science Department, University of California, Santa Cruz. I am researching scalable data management solutions for petabyte-scale storage systems. I am also researching file system benchmarking and security.

June 2007–September 2008 — **Intern**, Advanced Technology Group, NetApp, Inc., Sunnyvale, CA. My main work focused on integrating metadata search and indexing functionality directly into NetApp controllers. Other projects I worked on include inline data de-duplication algorithms for primary storage systems and evaluation of network file system workloads.

June 2006–September 2006 — **Senior Intern**, Yahoo! Search Technology Performance Engineering Group, Yahoo! Inc., Sunnyvale CA. I helped design and develop a tool for benchmarking the Yahoo! search engine, which was deployed and used throughout YST.

January 2006–April 2006 — **Teaching Assistant**, Computer Science Department, University of California, Santa Cruz. I lead discussion sections with students and graded assignments for CMPS 128 Distributed Systems. I covered topics including architectures, synchronization, transactions, and scalability.

June 2005–September 2005	**Information Technology Support**, Instructional Development Department, University of California, Santa Barbara. I provided Linux and Windows IT support to faculty and students.
June 2000–September 2000	**Web Page Design**, Computer Science Department, California State University, Hayward. I helped maintain the university web page.

Patents

2. Andrew W. Leung, Minglong Shao, Shankar Pasupathy, "Search and Update of Attributes in File Systems", *Confidential, pending*.

1. Minglong Shao, Andrew W. Leung, Shankar Pasupathy, Tim Bisson, "Two-Dimensional Indexes for Quick Multi-Attribute Search in a File Catalog", *Confidential, pending*.

Publications

17. Andrew W. Leung, Ian F. Adams, Ethan L. Miller, "Magellan: A Searchable Metadata Architecture for Large-Scale File Systems" Technical Report UCSC-SSRC-09-07 November, 2009.

16. Andrew W. Leung, Aleatha Parker-Wood, Ethan L. Miller, "Copernicus: A Scalable, High-Performance Semantic File System" Technical Report UCSC-SSRC-09-06 October, 2009.

15. Andrew W. Leung, Minglong Shao, Timothy Bisson, Shankar Pasupathy, Ethan L. Miller, "Spyglass: Metadata Search for Large-Scale Storage Systems", ***;login: The USENIX Magazine***, Vol. 34, No. 3, June, 2009.

14. Andrew W. Leung, Minglong Shao, Timothy Bisson, Shankar Pasupathy, Ethan L. Miller, "Spyglass: Fast, Scalable Metadata Search for Large-Scale Storage Systems", ***Proceedings of the 7th USENIX Conference on File and Storage Technologies*** (FAST '09), San Francisco, California, February, 2009.

13. Andrew W. Leung, Ethan L. Miller, "Scalable Full-Text Search for Petascale File Systems", ***Proceedings of the 2008 Petascale Data Storage Workshop*** (PDSW '08), Austin, Texas, November, 2008.

12. Andrew W. Leung, Minglong Shao, Timothy Bisson, Shankar Pasupathy, Ethan L. Miller, "High-Performance Metadata Indexing and Search in Petascale Data Storage Systems", ***Proceedings of the SciDAC 2008 Conference***, Seattle, Washington, July, 2008.

11. Andrew W. Leung, Shankar Pasupathy, Garth Goodson, Ethan L. Miller, "Measurement and Analysis of Large-Scale Network File System Workloads", ***Proceedings of the 2008 USENIX Annual Technical Conference***, Boston, Massachusetts, June, 2008.

10. Andrew W. Leung, Minglong Shao, Timothy Bisson, Shankar Pasupathy, Ethan L. Miller, "Spyglass: Fast, Scalable Metadata Search for Large-Scale Storage Systems" Technical Report UCSC-SSRC-08-01 May, 2008.

9. Minglong Shao, Andrew Leung, Shankar Pasupathy, Tim Bisson, "New Indexes for a File Metadata Catalog" ***Network Appliance Technical Journal***, Vol. 4, No. 2, 2008.

8. Sage A. Weil, Andrew W. Leung, Scott A. Brandt, Carlos Maltzahn, "RADOS: A Fast, Scalable, and Reliable Storage Service for Petabyte-scale Storage Clusters" ***Proceedings of the 2nd ACM Petascale Data Storage Workshop*** (PDSW '07), Reno, Nevada, November, 2007.

7. Jonathan Koren, Yi Zhang, Sasha Ames, Andrew Leung, Carlos Maltzahn, Ethan L. Miller, "Searching and Navigating Petabyte Scale File Systems Based on Facets" ***Proceedings of the 2nd ACM Petascale Data Storage Workshop*** (PDSW '07), Reno, Nevada, November, 2007.

6. Andrew W. Leung, Ethan L. Miller, Stephanie Jones, "Scalable Security for Petascale Parallel File Systems" ***Proceedings of the ACM/IEEE International Conference on Supercomputing*** (SC '07), Reno, Nevada, November 2007.

5. Andrew W. Leung, Eric Lalonde, Jacob Telleen, James Davis, Carlos Maltzahn, "Using Comprehensive Analysis for Performance Debugging in Distributed Storage Systems" ***Proceedings of the 24rd IEEE / 15th NASA Goddard Conference on Mass Storage Systems and Technologies*** (MSST '07), San Diego, California, September, 2007.

4. Andrew W. Leung, "Scalable Security for High Performance, Petascale Storage" Technical Report UCSC-SSRC-07-07 June, 2007.

3. Andrew W. Leung, Eric Lalonde, Jacob Telleen, James, Davis, Carlos Maltzahn, "Using Comprehensive Analysis for Performance Debugging in Distributed Storage Systems" Technical Report UCSC-SSRC-07-05 May, 2007.

2. Andrew W. Leung, Ethan L. Miller, "Scaling Security for Big, Parallel File Systems" Work in Progress. ***Proceedings of the 5th USENIX Conference on File and Storage Technologies*** (FAST 2007), San Jose, California, February 2007.

1. Andrew W. Leung, Ethan L. Miller, "Scalable Security for Large, High Performance Storage Systems" ***Proceedings of the 2nd ACM Workshop on Storage Security and Survivability*** (StorageSS 2006), Alexandria, Virginia, October 2006.

Selected Research Projects

6. ***Organizing, Indexing, and Searching Large-Scale Storage Systems*** I am investigating scalable data management techniques for improving how users find and manage files. This includes analyzing user file access patterns and designing new file system and index structures for scalable file search and retrieval.
   Advisor: Prof. Ethan L. Miller, September 2007–present.

5. ***Scalable Security for Large Scale, High Performance Storage Systems*** I am exploring scalable security solutions for use in petabyte scale, high-performance storage systems. My goals are to introduce strong security techniques without compromising system performance.
   Advisor: Prof. Ethan L. Miller, April 2006–present.

4. ***Sparrow: An Infrastructure for Managing Data Across Heterogeneous Storage Devices***
   Advisor: Prof. Scott A. Brandt, January 2006–April 2006.

3. ***Towards Format Independent Streaming for Multimedia***
   Advisor: Prof. Roberto Manduchi, January 2006–April 2006.

2. ***An API for Easy Deployment of Peer-to-Peer Overlay Networks***
   Advisor: Prof. Ethan Miller, September 2005–December 2005.

1. ***Methods of Transparent Host Communication in Peer-to-Peer Overlay Networks***
   Advisor: Prof. Ben Y. Zhao, September 2004–March 2005.

Talks

13. "Spyglass: Fast, Scalable Metadata Search for Large-Scale Storage Systems", Seventh Annual Storage System Research Center Retreat, Santa Cruz, CA, May, 2009.

12. "Spyglass: Fast, Scalable Metadata Search for Large-Scale Storage Systems", 7th USENIX Conference on File and Storage Technologies (FAST '09), San Francisco, CA, February, 2009.

11. "Scalable Full-Text Search for Petascale File Systems", 2008 Petascale Data Storage Workshop (PDSW '08), Austin, TX, November, 2008.

10. "Measurement and Analysis of Large-Scale Network File System Workloads", 2008 USENIX Annual Technical Conference, Boston, MA, June, 2008.

9. "Measurement and Analysis of Large-Scale Network File System Workloads", Sixth Annual Storage System Research Center Retreat, Santa Cruz, CA, May, 2008.

8. "Measurement and Analysis of Large-Scale Network File System Workloads", NetApp, Inc., Sunnyvale, CA, June, 2008.

7. "Ceph: Petabyte-scale, High-Performance Distributed Storage", Cal State University, East Bay, Hayward, CA, February, 2008.

6. "Scalable Security in the Ceph Parallel File Systems", Agami Systems, Sunnyvale, CA, January, 2008.

5. "Scalable Security for Petascale Parallel File Systems", ACM/IEEE International Conference on Supercomputing (SC 07), Reno, NV, November, 2007.

4. "Using Comprehensive Analysis for Performance Debugging in Distributed Storage Systems", 24th IEEE/15th NASA Goddard Conference on Mass Storage Systems and Technologies (MSST 07), San Diego, CA, September, 2007.

3. "Scalable Security for High Performance, Petascale Storage", Fifth Annual Storage Systems Research Center Retreat, Santa Cruz, CA, June, 2007.

2. "Scalable Security for Large, High Performance Storage Systems", Second ACM International Workshop on Storage Security and Survivability (StorageSS) 2006, Alexandria, VA, October, 2006.

1. "Scalable Security for Large, High Performance Storage Systems", Fourth Annual Storage Systems Research Center Retreat, Santa Cruz, CA, June, 2006.

PROFESSIONAL ACTIVITIES

Membership in Professional Societies

- **Member**, USENIX

Awards

- *USENIX Student Grant.* (5th USENIX Conference on File and Storage Technologies, 2007)

External Reviewer

- *The IEEE Transactions on Parallel and Distributed Systems Journal*
- *The Journal of Computer Science and Technology*
- *The 2009 IEEE International Conference on Networking, Architecture, and Storage* (NAS 2009)
- *The 6th USENIX Conference on File and Storage Technologies* (FAST 2008)
- *The 2nd International ACM Workshop on Storage Security and Survivability* (StorageSS 2006)