

## Justin Cappos

---

Computer Science Department  
Gould-Simpson Building  
1040 E. 4th St.  
University of Arizona  
Tucson, AZ 85721

Phone: (520) 626-7443  
Fax: (520) 621-4246  
[justin@cs.arizona.edu](mailto:justin@cs.arizona.edu)  
<http://www.cs.arizona.edu/people/justin/>

### Education

Ph.D. Computer Science, University of Arizona, May 2008 (Expected)  
M.S. Computer Science, University of Arizona, 2005  
B.S. Mathematics, University of Arizona, 2002, *Summa Cum Laude*  
B.S. Computer Science, University of Arizona, 2001, *Summa Cum Laude*

### Research Interests

Large Scale Distributed Systems, Security, Package Management, Virtualization

### Research Experience

**ETH Zürich** Zürich, Switzerland. Fall 2007 (Research Intern):

*Ouroboros*: Building a self-monitoring, self-deploying platform for distributed systems. Ouroboros acquires resources, monitors the application, and uses constrained logic programming to decide when to change the current resource allocation from within the distributed system.

**Microsoft Research** Cambridge, UK. Summer 2006 (Research Intern):

*Seaweed cost estimation*: Estimating the cost of materialized views in a highly distributed environment. Based upon local data update patterns on end systems in a large network, we estimate the cost of centralizing the data, replicating it in the network, or querying it in place without replication.

**University of Arizona** Tucson, AZ, USA. 2002–present (Research Associate):

#### Projects I led and founded:

*Stork*: A package management system designed for VM environments. Stork securely shares packages between VMs, provides package-based configuration management, reduces repository load, and has greatly improved security. See <http://www.cs.arizona.edu/stork> for more information.

*San Fermín*: A content aggregation system focused on completeness (number of systems returned in the answer) and scalability with respect to the aggregate data size. San Fermín uses a novel abstraction called a binomial swap tree to prevent overloading nodes or network resources.

*BACKS*: An enabling infrastructure for resource allocation. Multiple resource allocators can be concurrently evaluated on a platform, different resource types can be added without modifying resource allocators, and the user and resource allocator trust assumptions are minimized.

*DsCats*: A Data Structure visualization tool meant for CS2/3 students. DsCats also has been used in other projects to help debug data structures and algorithms as they are implemented.

*Simultaneous Embedding*: Proving the ability to simultaneously embed an outerplanar graph and a path. I proved that outerplanar graphs could have straight lines if the path is allowed to have circular arcs or 2 bends.

#### Participated in the following projects:

*Net-X*: A data-centric model of Internet services. A Distributed Hash Table (DHT) abstraction is used to indicate user interest in content. I developed security mechanisms to prevent DoS attacks by nodes in a DHT.

*DiamondTouch*: A project focused on measuring the benefits of using a collaborative touch screen tool versus a more typical mouse setup. I wrote code that allowed multiple mice to be used concurrently on the system.

*Prodeva/PSKI*: A security model where users act in a public manner and are judged by their past actions. I created collection programs that mine terabytes of data from publicly available sources.

## Justin Cappos

---

### Conference Publications

San Fermín: Aggregating Large Data Sets using Dynamic Binomial Trees. J. Cappos, J. Hartman. *The 5th USENIX Symposium on Networked Systems Design and Implementation (NSDI '08)*, San Francisco, CA, 2008. **Awarded the Graduate Research Excellence Award by the University of Arizona Computer Science Department**

Stork: Package Management for Distributed VM Environments. J. Cappos, S. Baker, J. Plichta, D. Nyugen, J. Hardies, M. Borgard, J. Johnston, J. Hartman. *21st Large Installation System Administration Conference (LISA 2007)*, Dallas, TX, 2007.

Proper: Privileged Operations in a Virtualised System Environment. S. Muir, L. Peterson, M. Fiuczynski, J. Cappos, J. Hartman. *USENIX '05 Annual Technical Conference*. Anaheim, CA, 2005.

Collaboration with DiamondTouch. S. Kobourov, K. Pavlou, J. Cappos, M. Stepp, M. Miles, A. Wixted. *The Tenth IFIP TC13 International Conference on Human-Computer Interaction*. Rome, Italy, 2005

Simultaneous Graph Embedding with Bends and Circular Arcs. J. Cappos, A. Estrella-Balderrama, J. Fowler, S. Kobourov. *14th International Symposium on Graph Drawing*. Karlsruhe, Germany, 2006

### Workshop Publications

Net-X: Unified Data-Centric Internet Services. P. Rao, J. Cappos, V. Khare, B. Moon, B. Zhang, *NetDB: Workshop On Networking Meets Databases*. Cambridge, MA, April 2007

Why It Is Hard to Build a Long Running Service on Planetlab. J. Cappos, J. Hartman, *Workshop on Real Large Distributed Systems (WORLDS 2005)*, San Francisco, CA, December 2005.

Trees on Tracks. J. Cappos, S. Kobourov. *14th Annual Fall Workshop on Computational Geometry*. MIT, Cambridge, MA, 2004.

### Other Publications

Centralized Package Management Using Stork. J. Samuel, J. Plichta, J. Cappos. *login: magazine*, February, 2008. **Awarded the Graduate Research Excellence Award by the University of Arizona Computer Science Department**

Privileged Operations in a Virtualised System Environment. S. Muir, L. Peterson, M. Fiuczynski, J. Cappos, J. Hartman. *Operating Systems Review Volume 40*, 2006.

Package Management Security. J. Cappos, J. Samuel, S. Baker, J. Hartman. *University of Arizona Tech Report 08-02*, 2008. (under submission)

A Resource Allocation Framework for Global Service-Oriented Networks. J. Cappos, J. Hartman. *University of Arizona Tech Report 05-02*, 2005.

Animating Data Structures for CS 2 and CS 3 Courses. J. Cappos, P. Homer, *University of Arizona Tech Report 01-02*, 2001.

## Justin Cappos

---

### Selected Presentations

“San Fermin: Aggregating Large Data Sets using Dynamic Binomial Trees” University of Washington, Cambridge University, Columbia University, 2007.

“Partial Replication of Highly Distributed Data” Microsoft Research Cambridge, 2006.

“Why It Is Hard To Build A Long Running Service On PlanetLab” WORLDS, 2005.

“Stork and the Benefits of Research on PlanetLab” IBM University of Arizona Forum, 2005.

“Stork” PlanetLab Spring Workshop, 2005.

“Trees on Tracks” 14th Annual Fall Workshop on Computational Geometry, 2004.

### Teaching Experience

Instructor, Computer Information Systems, Cochise College, 2001-2002 (9 classes)

Teaching Assistant, Data Structures and Algorithms, University of Arizona, Fall 2002

### Professional Activities

SOSP Shadow PC, 2007

External Reviewer NSDI 2006

Computer Science Graduate Student Committee Chair, University of Arizona, 2004-2006

Undergraduate Research Program Coordinator, University of Arizona, 2004-2008

Graduate Affairs Representative, University of Arizona, 2005-2006

Computing Committee Representative, University of Arizona, 2004-2005

Mentoring Program Coordinator, University of Arizona, 2004-2006

Wildcat Veteran’s Organization Co-founder and Webmaster, University of Arizona, 2003-2005

### Honors and Awards

Computer Science Department Outstanding Service Award, 2008

Two Graduate Research Excellence Awards by the University of Arizona Computer Science Dept.

Galileo Circle Scholar, 2006

Graduate Registration Scholarship, 2004-2008

Graduate College Fellowship, 2002

University of Arizona, Undergraduate Student Speaker 2002

2nd prize in University of Arizona Student Showcase 2000