

# Anshul Gandhi

Assistant Professor, Computer Science, Stony Brook University  
347, Computer Science, Stony Brook, NY 11790  
Phone: (631) 632-8475  
Email: anshul@cs.stonybrook.edu  
Web: www.cs.stonybrook.edu/~anshul

## Research Interests

- Performance Modeling and Design of Distributed Systems.

## Employment

- **Stony Brook University** (August 2014 - present)  
Assistant Professor, Department of Computer Science  
Affiliate Faculty, Department of Applied Mathematics and Statistics  
Stony Brook, NY, USA
- **IBM T.J. Watson Research Center** (August 2013 - July 2014)  
Post Doctoral Researcher, Cloud Management and Analytics  
Hosts: Andrew Kochut and Li Zhang  
Yorktown Heights, NY, USA

## Education

- **Carnegie Mellon University** (2007-2013)  
Ph.D., Computer Science Department  
Thesis: Dynamic Server Provisioning for Data Center Power Management  
Adviser: Prof. Mor Harchol-Balter  
Committee Members: David Andersen (CMU), Mor Harchol-Balter (CMU), Jeff Kephart (IBM T.J. Watson), Alan Scheller-Wolf (Tepper School of Business, CMU), Karsten Schwan (Georgia Tech)  
Pittsburgh, PA, USA
- **Indian Institute of Technology, Kanpur** (2003-2007)  
B.Tech, Computer Science and Engineering  
CGPA: 3.9/4.0 (Departmental Rank: 1)  
Thesis: Indecomposable Graphs  
Adviser: Prof. Shashank K. Mehta  
Kanpur, UP, India

## Work Experience

- **HP Labs** (Summer 2010)  
Summer Intern  
Hosts: Yuan Chen and Cullen Bash  
Palo Alto, CA, USA
- **Intel Research Pittsburgh** (Summer 2009)  
Summer Intern  
Hosts: Michael Kozuch and Jason Campbell  
Pittsburgh, PA, USA
- **IBM T.J. Watson Research Center** (Summer 2008)  
Summer Intern  
Hosts: Rajarshi Das and Jeff Kephart  
Hawthorne, NY, USA
- **Institute of Theoretical Computer Science, ETH** (Summer 2006)  
Summer Intern  
Hosts: Uli Wagner and Emo Welzl  
Zurich, Switzerland

## Teaching Experience

- Instructor (Spring 2018, Stony Brook University)  
CSE 544: Probability and Statistics for Data Science (Graduate Course)
- Instructor (Fall 2017, Stony Brook University)  
CSE 390: Special Topics: Probability and Statistics for Data Science (Undergraduate Course)
- Instructor (Spring 2017, Stony Brook University)  
CSE 544: Probability and Statistics for Data Science (Graduate Course)
- Instructor (Spring 2016, Stony Brook University)  
CSE 591: Energy-Efficient Computing (Graduate Course)
- Instructor (Spring 2015, Stony Brook University)  
CSE 531: Performance Analysis of Systems (Graduate Course)
- Instructor (Fall 2014, Stony Brook University)  
CSE 691: Energy-Efficient Computing (Graduate Course)

## Awards & Honors

- NSF CAREER Award, 2018.
- General Co-Chair, Middleware 2017.
- Program Vice Chair, ICDCS 2017 (Area Co-Chair of Cloud Computing and Data Centers track).
- Microsoft Azure Research Award, 2016.
- Google Research Award, 2015.
- IBM Faculty Award, 2015.
- SPEC Distinguished Dissertation Award, 2013.
- Best Paper Award (IGCC 2011 paper), IEEE International Green Computing Conference, 2011.
- Travel grant for USENIX Federated Conferences Week 2012
- Travel grant for IFIP Performance 2010.
- Pick of the Month (IGCC 2011 paper) for March 2012 in the IEEE STC on Sustainable Computing.
- Mentored the CMU Senior Thesis Project that won the 2011 Award for Undergraduate Excellence.
- Semi-finalist at the MIT Clean Energy Prize 2011 competition.
- Winner of the CMU Cross Campus New Venture Competition, 2011.
- Kauffman Foundation Commercialization prize, 2010.
- Intel/CMU Summer Fellowship, 2009.
- General Proficiency Medal for Best Academic Performance in the Department of Computer Science and Engineering (IIT Kanpur).
- Proficiency Medal for the Best Project Work (Indecomposable Graphs) in the Department of Computer Science and Engineering (IIT Kanpur).
- Pratibha Scholarship from the state government for academic excellence (IIT Kanpur).

## Publications

*Total Citations: 1698      h-index: 18*

- **Providing Performance Guarantees for Cloud-deployed Applications**  
Anshul Gandhi, Parijat Dube, Alexei Karve, Andrzej Kochut, and Li Zhang.  
*Transactions on Cloud Computing*, 2017.
- **Realizing an Elastic Memcached via Cached Data Migration**  
Ubaid Ullah Hafeez, Deepthi Male, Sharath Kumar Naeni, Muhammad Wajahat, and Anshul Gandhi.  
*Middleware* (poster), Las Vegas, NV, USA, December 2017.

- **Improving server utilization via resource-adaptive batch VMs**  
Seyyed Ahmad Javadi, Piyush Shyam Banginwar, Vaishali Chanana, Rashmi Narvekar, Mitesh Kumar Savita, and Anshul Gandhi.  
*Middleware* (poster), Las Vegas, NV, USA, December 2017.
- **MLscale: A Machine Learning based Application-Agnostic Autoscaler**  
Muhammad Wajahat, Alexei Karve, Andrzej Kochut, and Anshul Gandhi.  
*Sustainable Computing, Informatics and Systems*, 2017.
- **Lessons Learnt from Software Tuning of a Memcached-Backed, Multi-Tier, Web Cloud Application**  
Muhammad Wajahat, Salman Masood, Abhinav Sau, and Anshul Gandhi.  
*International Green and Sustainable Computing Conference*, Orlando, FL, USA, October 2017.
- **Rethinking TCP Throughput and Latency Modeling**  
Yi Cao, Aruna Balasubramanian, and Anshul Gandhi.  
*SIGCOMM* (poster), Los Angeles, CA, USA, August 2017.
- **Modeling and Analysis of Performance under Interference in the Cloud**  
Scott Votke, Seyyed Ahmad Javadi, and Anshul Gandhi.  
*MASCOTS*, Banff, Canada, September 2017.
- **DIAL: Reducing Tail Latencies for Cloud Applications via Dynamic Interference-aware Load Balancing**  
Seyyed Ahmad Javadi and Anshul Gandhi.  
*ICAC*, Columbus, OH, USA, July 2017.
- **Deconstructing the Energy Consumption of the Mobile Page Load**  
Yi Cao, Javad Nejati, Muhammad Wajahat, Aruna Balasubramanian, and Anshul Gandhi.  
*Sigmetrics*, Urbana-Champaign, IL, USA, June 2017.
- **Model-driven optimal resource scaling in cloud**  
Anshul Gandhi, Parijat Dube, Alexei Karve, Andrzej Kochut, and Li Zhang.  
*Software and Systems Modeling*, 2017.
- **Using Machine Learning for Black-Box Autoscaling**  
Muhammad Wajahat, Alexei Karve, Andrzej Kochut, and Anshul Gandhi.  
*International Green and Sustainable Computing Conference*, Hangzhou, China, November 2016.
- **Minimizing Electricity Cost for Geo-Distributed Interactive Services with Tail Latency Constraint**  
Mohammad Islam, Anshul Gandhi, and Shaolei Ren.  
*International Green and Sustainable Computing Conference*, Hangzhou, China, November 2016.
- **Dynamic Interference-Aware Load Balancing**  
Seyyed Ahmad Javadi, Himanshu Rajput, and Anshul Gandhi.  
*SOCC* (poster), Santa Clara, CA, USA, October 2016.
- **Using Predictions in Online Optimization: Looking forward with an eye on the past**  
Niangjun Chen, Joshua Comden, Zhenhua Liu, Anshul Gandhi, and Adam Wierman.  
*Sigmetrics*, Antibes, France, June 2016.
- **Analyzing the Power Consumption of the Mobile Page Load**  
Yi Cao, Javad Nejati, Pavan Maguluri, Aruna Balasubramanian, and Anshul Gandhi.  
*Sigmetrics* (poster), Antibes, France, June 2016.
- **Autoscaling for Hadoop Clusters**  
Anshul Gandhi, Parijat Dube, Andrzej Kochut, Li Zhang, and Sidhartha Thota.  
*IC2E*, Berlin, Germany, April 2016.

- **UIE: User-centric Interference Estimation for Cloud Applications**  
Seyyed Ahmad Javadi, Sagar Mehra, Bharath Vangoor, and Anshul Gandhi.  
*IC2E*, Berlin, Germany, April 2016.
- **HALO: Heterogeneity-Aware Load Balancing**  
Anshul Gandhi, Xi Zhang, and Naman Mittal.  
*MASCOTS*, Atlanta, GA, USA, October 2015.
- **The Unobservability Problem in Clouds**  
Anshul Gandhi, Parijat Dube, Alexei Karve, Andrzej Kochut, and Harsha Ellanti.  
*CAC*, Cambridge, MA, USA, September 2015.
- **Model-driven Autoscaling for Hadoop clusters**  
Anshul Gandhi, Parijat Dube, Andrzej Kochut, Li Zhang, and Sidhartha Thota.  
*SOCC* (poster), Kohala Coast, HI, USA, August 2015.
- **Model-driven Autoscaling for Hadoop clusters**  
Anshul Gandhi, Parijat Dube, Andrzej Kochut, and Li Zhang.  
*ICAC* (poster), Grenoble, France, July 2015.
- **Analyzing the Network for AWS Distributed Cloud Computing**  
Anshul Gandhi and Justin Chan.  
*DCC*, Portland, OR, USA, June 2015.
- **Optimal Load-Balancing for Heterogeneous Clusters**  
Anshul Gandhi, Naman Mittal, and Xi Zhang.  
*DCC* (poster), Portland, OR, USA, June 2015.
- **Modeling the Impact of Workload on Cloud Resource Scaling**  
Anshul Gandhi, Parijat Dube, Alexei Karve, Andrzej Kochut, and Li Zhang.  
*SBAC*, Paris, France, October 2014.
- **Adaptive, Model-driven Autoscaling for Cloud Applications**  
Anshul Gandhi, Parijat Dube, Alexei Karve, Andrzej Kochut, and Li Zhang.  
*ICAC*, Philadelphia, PA, USA, June 2014.
- **Exact Analysis of an M/M/k with Setup Times via Recursive Renewal Reward**  
Anshul Gandhi, Sherwin Doroudi, Mor Harchol-Balter, and Alan Scheller-Wolf.  
*Queueing Systems*, Volume 77, Issue 2, pp. 177-209.
- **Dynamic Management of Caching Tiers**  
Anshul Gandhi.  
*ICPE* (invited abstract), Dublin, Ireland, March 2014.
- **M/G/k with Staggered Setup**  
Anshul Gandhi and Mor Harchol-Balter.  
*Operations Research Letters*, Volume 41, Issue 4, pp. 317-320.
- **Exact Analysis of an M/M/k with Setup Times via Recursive Renewal Reward**  
Anshul Gandhi, Sherwin Doroudi, Mor Harchol-Balter, and Alan Scheller-Wolf.  
*SIGMETRICS*, Pittsburgh, PA, USA, June 2013.
- **AutoScale: Dynamic, Robust Capacity Management for Multi-Tier Data Centers**  
Anshul Gandhi, Mor Harchol-Balter, Ram Raghunathan, and Michael Kozuch.  
*Transactions on Computer Systems*, Volume 30, Issue 4, Article 14.
- **SOFTScale: Stealing Opportunistically For Transient Scaling**  
Anshul Gandhi, Timothy Zhu, Mor Harchol-Balter, and Michael Kozuch.  
*Middleware*, Montreal, Canada, December 2012.

- **Saving Cash by Using Less Cache**  
Timothy Zhu, Anshul Gandhi, Mor Harchol-Balter, and Michael Kozuch.  
*HotCloud*, Boston, MA, USA, June 2012.
- **Are sleep states effective in data centers?**  
Anshul Gandhi, Mor Harchol-Balter, and Michael Kozuch.  
*International Green Computing Conference*, San Jose, CA, USA, June 2012.
- **Hybrid Resource Provisioning for Minimizing Data Center SLA Violations and Power Consumption**  
Anshul Gandhi, Yuan Chen, Daniel Gmach, Martin Arlitt, and Manish Marwah.  
*Sustainable Computing: Informatics and Systems*, Volume 2, Issue 2, pp. 91-104.
- **The Case for Sleep States in Servers**  
Anshul Gandhi, Mor Harchol-Balter, and Michael Kozuch.  
*HotPower*, Cascais, Portugal, October 2011.
- **Distributed, Robust Auto-Scaling Policies for Power Management in Compute Intensive Server Farms**  
Anshul Gandhi, Mor Harchol-Balter, Ram Raghunathan, and Michael Kozuch.  
*Open Cirrus Summit*, Atlanta, GA, USA, October 2011.
- **How Data Center Size Impacts the Effectiveness of Dynamic Power Management**  
Anshul Gandhi and Mor Harchol-Balter.  
*49th Annual Allerton Conference on Communication, Control, and Computing*, Allerton, IL, USA, September 2011.
- **Minimizing Data Center SLA Violations and Power Consumption via Hybrid Resource Provisioning (Best Paper Award)**  
Anshul Gandhi, Yuan Chen, Daniel Gmach, Martin Arlitt, and Manish Marwah.  
*International Green Computing Conference*, Orlando, FL, USA, July 2011.
- **Optimality Analysis of Energy-Performance Trade-off for Server Farm Management**  
Anshul Gandhi, Varun Gupta, Mor Harchol-Balter, and Michael Kozuch.  
*Performance Evaluation*, Vol. 67, Issue 11, pp. 1155-1171.
- **Server Farms with Setup Costs**  
Anshul Gandhi, Mor Harchol-Balter, and Ivo Adan.  
*Performance Evaluation*, Vol. 67, Issue 11, pp. 1123-1138.
- **M/G/k with Exponential Setup**  
Anshul Gandhi, Mor Harchol-Balter, and Ivo Adan.  
*Madrid Conference on Queueing Theory*, Madrid, Spain, June 2010.
- **Decomposition Results for an M/M/k with Staggered Setup**  
Anshul Gandhi, Mor Harchol-Balter, and Ivo Adan.  
*Performance Evaluation Review*, Vol. 38, Issue 2, pp. 48-50.
- **Optimal Power Allocation in Server Farms**  
Anshul Gandhi, Mor Harchol-Balter, Rajarshi Das, and Charles Lefurgy.  
*SIGMETRICS/Performance*, Seattle, WA, USA, June 2009.
- **Power Capping Via Forced Idleness**  
Anshul Gandhi, Mor Harchol-Balter, Rajarshi Das, Jeffrey Kephart, and Charles Lefurgy.  
*Workshop on Energy-Efficient Design*, Austin, TX, USA, June 2009.
- **Database Summarization and Publishing in Wireless Environments**  
Anshul Gandhi and R.K.Ghosh.  
*International Conference on Distributed Computer Networking*, Guwahati, India, December 2006.

## Invited Talks

- Modeling and Analysis of Performance under Interference in the Cloud. *Computer Science Department, Carnegie Mellon University*, November 2017.
- DIAL: Dynamic Interference-Aware Load Balancing. *Department of Computer Science, University of Illinois, Urbana-Champaign*, June 2017.
- Dynamic Server Provisioning for Data Center Power Management. *The 11th International Conference & Expo on Emerging Technologies for a Smarter World (CEWIT2014)*, October 2014.
- Providing Performance Guarantees for Cloud Applications. *School of Computer Science and Informatics, University College Dublin*, March 2014.
- Providing Performance Guarantees for Cloud Applications. *DIMACS Working Group on Algorithms for Green Data Storage, Rutgers University*, December 2013.
- Performance Modeling for Data Center Power Management. *Department of Technology Management, University of California, Santa Cruz*, April 2013.
- Dynamic Capacity Management for Multi-Tier Data Centers. *Computer Science and Engineering Department, University at Buffalo*, March, 2013.
- Dynamic Capacity Management for Multi-Tier Data Centers. *Microsoft Research*, March, 2013.
- Performance Modeling for Data Center Power Management. *Rotman School of Management, University of Toronto*, March 2013.
- Dynamic Capacity Management for Multi-Tier Data Centers. *Department of Computer Science, Stony Brook University*, March, 2013.
- Dynamic Capacity Management for Multi-Tier Data Centers. *Computer Science Department, Binghamton University*, March, 2013.
- Dynamic Capacity Management for Multi-Tier Data Centers. *School of Computing & Information Sciences, Florida International University*, March, 2013.
- Dynamic Capacity Management for Multi-Tier Data Centers. *IBM T.J. Watson Research Center*, December, 2012.
- Exact Analysis of the M/M/2 with Setup Times and other Hard Variants. *INFORMS*, October 2012.
- Dynamic Capacity Management for Multi-Tier Data Centers. *Computer Science Department, New York University*, November 2012.
- Dynamic Capacity Management for Multi-Tier Data Centers. *Computer Science Department Colloquium, Rutgers University*, November 2012.
- Dynamic Capacity Management for Multi-Tier Data Centers. *Computer Science Department, University of California, Santa Cruz*, June 2012.
- AutoScale: Dynamic Power Management for Multi-Tier Data Centers. *CERCS Seminar, Georgia Institute of Technology*, November 2011.
- POW Solutions. *MIT Clean Energy Prize Semi-Finals*, April 2011.
- POW Solutions. *CMU Cross-Campus New Venture Competition*, February 2011.
- Power-efficient Server Provisioning in Server Farms. *Invited Talk, IBM Student Workshop for Frontiers of Cloud Computing*, September 2010.
- POW Solutions. *USF International Business Plan Competition*, March 2010.
- Optimal Power Allocation in Server Farms. *Guest Lecture, Algorithmic Power Management, University of Pittsburgh*, February 2010.
- Optimizing Server Farm Performance in Power-Constrained Environments. *TTC Technology Commercialization Advisory Board Meeting*, October 2009.
- Optimal Power Allocation in Server Farms. *Workshop on Quantitative Models for Production and Communication Networks, Eindhoven University of Technology*, January 2009.
- Power Management in Server Farms. *Parallel Data Laboratory Retreat*, November 2008.

## Conference Talks

- Autoscaling for Hadoop Clusters. *IC2E 2016*, April 2016.
- UIE: User-centric Interference Estimation for Cloud Applications. *IC2E 2016*, April 2016.

- HALO: Heterogeneity-Aware Load Balancing. *MASCOTS 2015*, October 2015.
- The Unobservability Problem in Clouds. *CAC 2015*, September 2015.
- Analyzing the Network for AWS Distributed Cloud Computing. *DCC 2015*, June 2015.
- Adaptive, Model-driven Autoscaling for Cloud Applications. *ICAC 2014*, June 2014.
- DynamicManagementofCachingTiers. *ICPE 2014*, March 2014.
- Exact Analysis of an M/M/k with Setup Times via Recursive Renewal Reward. *Sigmetrics 2013*, June 2013.
- SOFTScale: Stealing Opportunistically For Transient Scaling. *Middleware 2012*, December 2012.
- Are sleep states effective in data centers? *International Green Computing Conference*, June 2012.
- Distributed, Robust Auto-Scaling Policies for Power Management in Compute Intensive Server Farms. *Open Cirrus Summit*, October 2011.
- The Case for Sleep States in Servers. *HotPower 2011*, October 2011.
- Minimizing Data Center SLA Violations and Power Consumption via Hybrid Resource Provisioning. *International Green Computing Conference*, July 2011.
- Optimality Analysis of Energy-Performance Trade-off for Server Farm Management. *Performance 2010*, November 2010.
- Decomposition Results for an M/M/k with Staggered Setup. *MAMA 2010*, June 2010.
- M/G/k with Exponential Setup. *Madrid Conference on Queueing Theory*, June 2010.
- Power Capping Via Forced Idleness. *Workshop on Energy-Efficient Design*, June 2009.
- Optimal Power Allocation in Server Farms. *Sigmetrics 2009*, June 2009.

## Funding

Total Funding: ~\$1.77 Million      Personal Share: ~\$1.4 Million

- NSF - CNS: CAREER: Enabling Predictable Performance in Cloud Computing  
\$400,336, 9/18 - 8/23, sole PI.
- NSF - NeTS: Small: Demystifying the role of prediction models: bridging prediction algorithms and resource provisioning  
\$449,817, 9/17 - 8/20, PI.
- NSF - CRI: II-EN: Collaborative Research: Enhancing the Parasol Experimental Testbed for Sustainable Computing  
\$24,215, 7/17 - 6/20, PI (for SBU).
- Microsoft - Azure Research Award  
\$20,000 (Azure credits), 10/16 - 9/17, sole PI.
- NSF - CSR: Small: Scalable, heterogeneity-aware load balancing  
\$394,981, 10/16 - 9/19, sole PI.
- NSF - EAGER: Elastic Multi-layer Memcached Tiers  
\$257,166, 6/16 - 5/18, PI.
- Google - Research Award: The effect of Web optimizations on mobile browser performance and power  
\$41,088 (gift), co-PI.
- NSF - CRII: CSR: Online Performance Modeling of Opaque Cloud Applications  
\$173,229, 9/15 - 8/17, sole PI.
- IBM - Faculty Award: AutoScaling for Cloud Applications  
\$10,000 (gift), sole PI.
- Amazon - AWS in Education Grant Award  
\$3,700 (AWS credits), 8/14 - 7/15, sole PI.

## Patents

- Systems and Methods for Scaling a Cloud Infrastructure. Anshul Gandhi, Parijat Dube, Alexei Karve, Andrzej Kochut and Li Zhang (Issued: US 9,300,553).

- Power Budget Allocation in Multi-Processor Systems. Mor Harchol-Balter, Anshul Gandhi, Rajarshi Das and Jeff Kephart (Issued: US 9,052,895).
- Dynamic Capacity Management of Multiple Parallel Connected Computing Resources. Mor Harchol-Balter, Anshul Gandhi, Varun Gupta and Mike Kozuch (Issued: US 8,806,018).
- Systems and methods for managing power consumption and performance of a processor. Mor Harchol-Balter and Anshul Gandhi (Issued: US 8,589,709).
- Graphics Processing Unit Resource Sharing. Anshul Gandhi, Hui Lei, Jayaram K. R., Charles O. Schulz, Shu Tao (filed).
- Provisioning Data Center Resources. Yuan Chen, Anshul Gandhi, Daniel Gmach, Chris Hyser, Martin Arlitt, Manish Marwah and Cullen Bash (filed).

## Organizing Committee Member

- General Co-Chair, MIDDLEWARE 2017
- Program Vice Chair (Co-Chair of Cloud Computing and Data Centers track), ICDCS 2017
- Tutorials Co-Chair, Sigmetrics 2017
- Publicity Co-Chair, ICAC 2017
- Web Chair, Sigmetrics 2013

## Program Committee Member

- 2018: Sigmetrics, ICDCS, Middleware, IC2E, ICAC, HotCloud
- 2017: ICAC, IC2E, ICCAC, Cloud
- 2016: Sigmetrics, ICDCS, IGSC, MASCOTS, ICCAC, ICCCN, QEST, DCC, GPCDP, DIDC
- 2015: Middleware, IGSC, MASCOTS, GPCDP
- 2014: Sigmetrics, HotCloud, ICPP, CCWS
- 2010: Sigmetrics (Shadow PC)

## Editorial Board

- Performance Evaluation (2018 onwards)

## Refereeing

- IEEE Transactions on Computers, IEEE Transactions on Cloud Computing, ACM/IEEE Transactions on Networking, Performance Evaluation, Annals of Operations Research, IEEE Transactions on Parallel and Distributed Systems, ACM Transactions on Modeling and Performance Evaluation of Computing Systems, ACM Transactions on Autonomous and Adaptive Systems, IEEE Internet Computing, IEEE Transactions on Services Computing, IEEE Communications Letters, IEEE Transactions on Network and Service Management, IEEE Transactions on Dependable and Secure Computing, Journal of Parallel and Distributed Computing, Canadian Operations Research Journal, Computing, Parallel Computing, Sustainable Computing, International Journal of Parallel Programming, Simulation Modeling Practice and Theory

## Professional Membership

- ACM, IEEE