

ANSHUL GANDHI

Associate 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 2020 - present)
Associate Professor, Department of Computer Science
Affiliate Faculty, Department of Applied Mathematics and Statistics
Stony Brook, NY, USA
- **Stony Brook University** (August 2014 - July 2020)
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
Pittsburgh, PA, USA
Advisor: 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)
- **Indian Institute of Technology, Kanpur** (2003-2007)
B.Tech, Computer Science and Engineering
Kanpur, UP, India
CGPA: 3.9/4.0 (Departmental Rank: 1)
Thesis: Indecomposable Graphs
Advisor: Prof. Shashank K. Mehta

WORK EXPERIENCE

- **HP Labs** (Summer 2010)
Summer Intern
Palo Alto, CA, USA
Hosts: Yuan Chen and Cullen Bash
- **Intel Research Pittsburgh** (Summer 2009)
Summer Intern
Pittsburgh, PA, USA
Hosts: Michael Kozuch and Jason Campbell
- **IBM T.J. Watson Research Center** (Summer 2008)
Summer Intern
Hawthorne, NY, USA
Hosts: Rajarshi Das and Jeff Kephart

- **Institute of Theoretical Computer Science, ETH**
Summer Intern
Hosts: Uli Wagner and Emo Welzl

(Summer 2006)
Zurich, Switzerland

TEACHING EXPERIENCE

- Instructor (Spring 2022, Stony Brook University)
CSE 544: Probability and Statistics for Data Science (Graduate Course)
- Instructor (Spring 2021, Stony Brook University)
CSE 544: Probability and Statistics for Data Science (Graduate Course)
- Instructor (Fall 2020, Stony Brook University)
CSE 357: Statistical Methods for Data Science (Undergraduate Course)
- Instructor (Spring 2020, Stony Brook University)
CSE 544: Probability and Statistics for Data Science (Graduate Course)
- Instructor (Fall 2019, Stony Brook University)
CSE 357: Statistical Methods for Data Science (Undergraduate Course)
New course developed by Gandhi
- Instructor (Spring 2019, Stony Brook University)
CSE 391: Special Topics: Probability and Statistics for Data Science (Undergraduate Course)
- Instructor (Fall 2018, Stony Brook University)
CSE 544: Probability and Statistics for Data Science (Graduate Course)
- 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)
New course co-developed by Gandhi
- 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)
New course developed by Gandhi
- Instructor (Fall 2014, Stony Brook University)
CSE 691: Energy-Efficient Computing (Graduate Course)
New course developed by Gandhi

AWARDS

- Research Excellence Award (Department of Computer Science, Stony Brook University), 2019
- ACM Sigmetrics Rising Star Award, 2019.
- Best Paper Award, IEEE International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), 2019.
- Best Student Paper Award, IEEE International Conference on Distributed and Computing Systems (ICDCS), 2018.
- NSF CAREER Award, 2018.
- Microsoft Azure Research Award, 2016.
- NSF CRII (pre-CAREER) Award, 2015.
- Google Research Award, 2015.
- IBM Faculty Award, 2015.

- SPEC Distinguished Dissertation Award, 2013.
- Best Paper Award, IEEE International Green Computing Conference (IGCC), 2011.
- Semi-finalist, MIT Clean Energy Prize 2011 competition.
- Winner of the CMU Cross Campus New Venture Competition, 2011.
- Kauffman Foundation Commercialization Fellowship, 2010.
- Intel/CMU Summer Fellowship, 2009.
- General Proficiency Medal for Best Academic Performance in the Computer Science and Engineering (CSE) Department (IIT Kanpur).
- Proficiency Medal for Best Project Work in the CSE Department (IIT Kanpur).
- Dr. V. Rajaraman Scholarship for Best Student of the CSE Department in AY 2005-06 (IIT Kanpur).
- J. N. Kapur Prize for Best Academic Performance in AY 2004-05 (IIT Kanpur).

RESEARCH FUNDING

Total Funding: \$4,237,325 *Personal Share:* \$2,364,361

- NSF – Collaborative Research: CNS Core: Large: Systems and Verifiable Metrics for Sustainable Data Centers 10/2022 – 09/2027
PI: Anshul Gandhi, Co-PI: Dongyoon Lee, Zhenhua Liu, Shuai Mu, Erez Zadok
Amount: \$928,403 personal share: \$201,481 (22%)
- SBU OVPR Seed Grant – Enabling Live Video Analytics for Mobile Cameras 08/2022 – 02/2024
PI: Shubham Jain, Co-PI: Anshul Gandhi, Samir Das
Amount: \$50,000, personal share: \$16,666 (33%)
- OpenInfra Labs – Performance Analysis of Computer Systems 05/2021 – present
Sole PI: Anshul Gandhi
Amount: \$71,968 (gift), personal share: \$71,968 (100%)
- NSF – Collaborative Research: CNS Core: Medium: Optimizing Storage Caches via Adaptive and Reconfigurable Tiering 10/2021 – 09/2025
PI: Erez Zadok, Co-PI: Anshul Gandhi
Amount: \$533,333 personal share: \$259,199 (49%)
- NSF – CNS Core: Small: Significantly Improving Networked Applications using Congestion-Aware Transport Modeling 10/2019 – 09/2022
PI: Aruna Balasubramanian, Co-PI: Anshul Gandhi
Amount: \$499,698, personal share: \$249,849 (50%)
- DoE – Grid Ready Energy Analytics Training with Data 10/2019 – 09/2024
PI: Zhenhua Liu, Co-PI: Eugene Feinberg, Anshul Gandhi, Erez Zadok, Yue Zhao
Amount: \$244,000, personal share: \$20,000 (8%)
- NSF – Student Travel Grant for the 2019 ACM Sigmetrics International Conference on Measurement and Modeling of Computer Systems 05/2019 – 04/2020
Sole PI: Anshul Gandhi
Amount: \$25,000, personal share: \$25,000 (100%)
- Amazon – AWS Cloud Credits for Research Award 05/2019 – 04/2020
Sole PI: Anshul Gandhi
Amount: \$15,000 (Azure credits), personal share: \$15,000 (100%)
- SBU OVPR Seed Grant – Exploiting Analytical Models to Realize the Full Performance and Parallelization Potential of Modern Storage Architectures for Big Data Applications 12/2018 – 06/2020
PI: Anshul Gandhi, Co-PI: Erez Zadok
Amount: \$60,000, personal share: \$30,000 (50%)
- NSF – CNS: CAREER: Enabling Predictable Performance in Cloud Computing 09/2018 – 08/2023
Sole PI: Anshul Gandhi
Amount: \$435,727, personal share: \$435,727 (100%)

- NSF – NeTS: Small: Demystifying the role of prediction models: bridging prediction algorithms and resource provisioning 09/2017 – 08/2020
PI: Anshul Gandhi, Co-PI: Zhenhua Liu
Amount: \$449,817, personal share: \$224,909 (50%)
- NSF – CRI: II-EN: Collaborative Research: Enhancing the Parasol Experimental Testbed for Sustainable Computing 07/2017 – 06/2020
PI (for SBU): Anshul Gandhi, Co-PI: Zhenhua Liu
Amount (for SBU): \$24,215, personal share: \$12,108 (50%)
- Microsoft – Azure Research Award 10/2016 – 09/2017
Sole PI: Anshul Gandhi
Amount: \$20,000 (Azure credits), personal share: \$20,000 (100%)
- NSF – CSR: Small: Scalable, heterogeneity-aware load balancing 10/2016 – 09/2019
Sole PI: Anshul Gandhi
Amount: \$394,981, personal share: \$394,981 (100%)
- NSF – EAGER: Elastic Multi-layer Memcached Tiers 06/2016 – 05/2018
PI: Anshul Gandhi, Co-PI: Erez Zadok
Amount: \$257,166, personal share: \$180,000 (70%)
- Google – Google Faculty Research Award: The effect of Web optimizations on mobile browser performance and power 02/2016 – present
PI: Aruna Balasubramanian, Co-PI: Anshul Gandhi
Amount: \$41,088 (gift), personal share: \$20,544 (50%)
- NSF – CRII: CSR: Online Performance Modeling of Opaque Cloud Applications 09/2015 – 08/2017
Sole PI: Anshul Gandhi
Amount: \$173,229, personal share: \$173,229 (100%)
- IBM – Faculty Award: AutoScaling for Cloud Applications 03/2015 – present
Sole PI: Anshul Gandhi
Amount: \$10,000 (gift), personal share: \$10,000 (100%)
- Amazon – AWS in Education Grant Award 08/2014 – 07/2015
Sole PI: Anshul Gandhi
Amount: \$3,700 (AWS credits), personal share: \$3,700 (100%)

PUBLICATIONS

Total Citations: 3,523 *h-index: 30*

(Gandhi and his [co-]advisees are underlined)

Journal Articles

- Seyyed Ahmad Javadi, Anshul Gandhi. User-Centric Interference-Aware Load Balancing for Cloud-Deployed Applications. *IEEE Transactions on Cloud Computing*, 10(1), 736–748, March 2022.
- Muhammad Wajahat, Aditya Yele, Tyler Estro, Anshul Gandhi, Erez Zadok. Analyzing the Distribution Fit for Storage Workload and Internet Traffic Traces. *Performance Evaluation*, 142, 102–121, September 2020.
- Anshul Gandhi, Parijat Dube, Alexei Karve, Andrzej Kochut, Li Zhang. Providing Performance Guarantees for Cloud-deployed Applications. *IEEE Transactions on Cloud Computing*, 8(1), 269–281, March 2020.
- Muhammad Wajahat, Alexei Karve, Andrzej Kochut, Anshul Gandhi. MLscale: A Machine Learning based Application-Agnostic Autoscaler. *Sustainable Computing, Informatics and Systems*, 22, 287–299, June 2019.

- [Anshul Gandhi](#), Parijat Dube, Alexei Karve, Andrzej Kochut, Li Zhang. Model-driven optimal resource scaling in cloud. *Software and Systems Modeling*, 17(2), 509–526, May 2018.
- [Anshul Gandhi](#), Sherwin Doroudi, Mor Harchol-Balter, Alan Scheller-Wolf. Exact Analysis of an M/M/k with Setup Times via Recursive Renewal Reward. *Queueing Systems*, 77(2), 177–209, June 2014.
- [Anshul Gandhi](#), Mor Harchol-Balter. M/G/k with Staggered Setup. *Operations Research Letters*, 41(4), 317–320, July 2013.
- [Anshul Gandhi](#), Mor Harchol-Balter, Ram Raghunathan, Michael Kozuch. AutoScale: Dynamic, Robust Capacity Management for Multi-Tier Data Centers. *ACM Transactions on Computer Systems*, 30(4), Article 14, November 2012.
- [Anshul Gandhi](#), Yuan Chen, Daniel Gmach, Martin Arlitt, Manish Marwah. Hybrid Resource Provisioning for Minimizing Data Center SLA Violations and Power Consumption. *Sustainable Computing: Informatics and Systems*, 2(2), 91-104, June 2012.
- [Anshul Gandhi](#), Varun Gupta, Mor Harchol-Balter, Michael Kozuch. Optimality Analysis of Energy-Performance Trade-off for Server Farm Management. *Performance Evaluation*, 67(11), 1155-1171, November 2010.
- [Anshul Gandhi](#), Mor Harchol-Balter, Ivo Adan. Server Farms with Setup Costs. *Performance Evaluation*, 67(11), 1123-1138, November 2010.

Refereed Conference Papers

- Santiago Vargas, Gautham Gunapati, Aruna Balasubramanian, [Anshul Gandhi](#). Are Mobiles ready for BBR? In *Proceedings of the 22nd ACM Internet Measurement Conference (IMC)*, Nice, France, October 2022.
- Gagan Somashekar, Mohammad Delasay, [Anshul Gandhi](#). Truncating Multi-Dimensional Markov Chains with Accuracy Guarantee. In *Proceedings of the 30th International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*, Nice, France, October 2022.
- Gagan Somashekar, Amoghavarsha Suresh, Saurabh Tyagi, Vikas Dhyani, [Krishna Donkada](#), [Anurag Pradhan](#), [Anshul Gandhi](#). Reducing the Tail Latency of Microservices Applications via Optimal Configuration Tuning. In *Proceedings of the 3rd IEEE International Conference on Autonomic Computing and Self-Organizing Systems (ACSOS)*, Virtual Event, September 2022.
- [Sri Pramodh Rachuri](#), [Arun Gantasala](#), [Prajeeth Emanuel](#), [Anshul Gandhi](#), Robert Foley, Peter Puhov, Theodoros Gkountouvas, Hui Lei. Optimizing Near-Data Processing for Spark. In *Proceedings of the 42nd IEEE International Conference on Distributed Computing Systems (ICDCS)*, Bologna, Italy, July 2022.
- Gagan Somashekar, [Anurag Dutt](#), [Rohith Vaddavalli](#), Sai Bhargav Varanasi, [Anshul Gandhi](#). B-MEG: Bottlenecked-Microservices Extraction Using Graph Neural Networks. In *Proceedings of the 13th ACM/SPEC International Conference on Performance Engineering (ICPE)*, Virtual Event, April 2022.
- [Ubaid Ullah Hafeez](#), Xiao Sun, [Anshul Gandhi](#), Zhenhua Liu. Towards Optimal Placement and Scheduling of DNN Operations with Pesto. In *Proceedings of the 2021 ACM/IFIP International Middleware Conference (Middleware)*, Virtual Event, December 2021.
- Amoghavarsha Suresh, [Anshul Gandhi](#). ServerMore: Opportunistic Execution of Serverless Functions in the Cloud. In *Proceedings of the 2021 ACM Symposium on Cloud Computing (SOCC)*, Virtual Event, November 2021.

- Santiago Vargas, Rebecca Drucker, Aiswarya Renganathan, Aruna Balasubramanian, Anshul Gandhi. BBR Bufferbloat in DASH Video. In *Proceedings of The Web Conference (WWW)*, Virtual Event, April 2021.
- Ubaid Ullah Hafeez, Anshul Gandhi. Empirical Analysis and Modeling of ComputeTimes of CNN Operations on AWS Cloud. In *Proceedings of the 2020 IEEE International Symposium on Workload Characterization (IISWC)*, Beijing, China, October 2020.
- Amoghavarsha Suresh, Gagan Somashekar, Anandh Varadarajan, Veerendra Kakarla, Hima Upadhyay, Anshul Gandhi. ENSURE: Efficient Scheduling and Autonomous Resource Management in Serverless Environments. In *Proceedings of the 1st IEEE International Conference on Autonomic Computing and Self-Organizing Systems (ACSOS)*, Washington, DC, USA, August 2020.
- Muhammad Wajahat, Bharath Balasubramanian, Anshul Gandhi, Gueyoung Jung, Shankar Narayanan. MERIT: Model-driven Rehoming for VNF Chains. In *Proceedings of the 1st IEEE International Conference on Autonomic Computing and Self-Organizing Systems (ACSOS)*, Washington, DC, USA, August 2020.
- Seyyed Ahmad Javadi, Amoghavarsha Suresh, Muhammad Wajahat, Anshul Gandhi. Scavenger: A black-box batch workload resource manager for improving utilization in cloud environments. In *Proceedings of the 10th ACM Symposium on Cloud Computing (SOCC)*, Santa Cruz, CA, USA, November 2019.
- Yi Cao, Javad Nejati, Aruna Balasubramanian, Anshul Gandhi. ECON: Modeling the network to improve application performance. In *Proceedings of the 19th ACM Internet Measurement Conference (IMC)*, Amsterdam, Netherlands, October 2019.
- Yi Cao, Arpit Jain, Kriti Sharma, Aruna Balasubramanian, Anshul Gandhi. When to use and when not to use BBR: An empirical analysis and evaluation study. In *Proceedings of the 19th ACM Internet Measurement Conference (IMC)*, Amsterdam, Netherlands, October 2019.
- Ubaid Ullah Hafeez, Alexei Karve, Braullio Dumba, Anshul Gandhi, Sai Zeng. Automated Patch Management in a Hybrid Cloud with HCPM. In *Proceedings of the 17th International Conference on Service-Oriented Computing (ICSOC)*, Toulouse, France, October 2019.
- Muhammad Wajahat, Aditya Yele, Tyler Estro, Anshul Gandhi, Erez Zadok. Distribution Fitting and Performance Modeling for Storage Traces (**Best Paper Award**). In *Proceedings of the 27th IEEE International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*, Rennes, France, October 2019.
- Scott Votke, Jazeem Abdul Jaleel, Amoghavarsha Suresh, Mohammad Delasay, Shewin Doroudi, Anshul Gandhi. Optimal Markovian Dynamic Control of Interference-Prone Server Farms. In *Proceedings of the 27th IEEE International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*, Rennes, France, October 2019.
- Amoghavarsha Suresh, Anshul Gandhi. Using Variability as a Guiding Principle to Reduce Latency in Web Applications via OS Profiling (**Oral Presentation**). In *Proceedings of the World Wide Web Conference (WWW)*, San Francisco, CA, USA, May 2019.
- K. R. Jayaram, Anshul Gandhi, Hongyi Xin, Shu Tao. Adaptively Accelerating Map-Reduce/Spark with GPUs: A Case Study. In *Proceedings of the 16th IEEE International Conference on Autonomic Computing (ICAC)*, Umeå, Sweden, June 2019.
- Vasudevan Nagendra, Arani Bhattacharya, Anshul Gandhi, Samir Das. Scalable and Resource Efficient Control Plane for Next Generation Cellular Packet Core. In *Proceedings of the 5th ACM Symposium on SDN Research (SOSR)*, San Jose, CA, USA, April 2019.

- Muhammad Wajahat, Bharath Balasubramanian, Anshul Gandhi, Gueyoung Jung, Shankaranarayanan Puzhavakath Narayanan. A Graybox Approach to Rehoming Service Chains. In *Proceedings of the 26th IEEE International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*, Milwaukee, WI, USA, September 2018.
- Ubaid Ullah Hafeez, Muhammad Wajahat, Anshul Gandhi. EMMem: Towards an Elastic Memcached System (**Best Student Paper Award**). In *Proceedings of the 38th IEEE International Conference on Distributed Computing Systems (ICDCS)*, Vienna, Austria, July 2018.
- Seyyed Ahmad Javadi, Harsh Gupta, Robin Manhas, Shweta Sahu, Anshul Gandhi. EASY: Efficient segment assignment strategy for reducing tail latencies in Pinot. In *Proceedings of the 38th IEEE International Conference on Distributed Computing Systems (ICDCS)*, Vienna, Austria, July 2018.
- Muhammad Wajahat, Salman Masood, Abhinav Sau, Anshul Gandhi. Lessons Learnt from Software Tuning of a Memcached-Backed, Multi-Tier, Web Cloud Application. In *Proceedings of the 8th IEEE International Green and Sustainable Computing Conference (IGSC)*, Orlando, FL, USA, October 2017.
- Scott Votke, Seyyed Ahmad Javadi, Anshul Gandhi. Modeling and Analysis of Performance under Interference in the Cloud. In *Proceedings of the 25th IEEE International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*, Banff, Canada, September 2017.
- Seyyed Ahmad Javadi Anshul Gandhi. DIAL: Reducing Tail Latencies for Cloud Applications via Dynamic Interference-aware Load Balancing. In *Proceedings of the 14th IEEE International Conference on Autonomic Computing (ICAC)*, Columbus, OH, USA, July 2017.
- Yi Cao, Javad Nejati, Muhammad Wajahat, Aruna Balasubramanian, Anshul Gandhi. Deconstructing the Energy Consumption of the Mobile Page Load. In *Proceedings of the 43rd ACM International Conference on Measurement and Modeling of Computer Science (SIGMETRICS)*, Urbana-Champaign, IL, USA, June 2017.
- Muhammad Wajahat, Alexei Karve, Andrzej Kochut, Anshul Gandhi. Using Machine Learning for Black-Box Autoscaling. In *Proceedings of the 7th IEEE International Green and Sustainable Computing Conference (IGSC)*, Hangzhou, China, November 2016.
- Mohammad Islam, Anshul Gandhi, Shaolei Ren. Minimizing Electricity Cost for Geo-Distributed Interactive Services with Tail Latency Constraint. In *Proceedings of the 7th IEEE International Green and Sustainable Computing Conference (IGSC)*, Hangzhou, China, November 2016.
- Niangjun Chen, Joshua Comden, Zhenhua Liu, Anshul Gandhi, Adam Wierman. Using Predictions in Online Optimization: Looking forward with an eye on the past. In *Proceedings of the 42nd ACM International Conference on Measurement and Modeling of Computer Science (SIGMETRICS)*, Antibes, France, June 2016.
- Anshul Gandhi, Parijat Dube, Andrzej Kochut, Li Zhang, Sidhartha Thota. Autoscaling for Hadoop Clusters. In *Proceedings of the 4th IEEE International Conference on Cloud Engineering (IC2E)*, Berlin, Germany, April 2016.
- Seyyed Ahmad Javadi, Sagar Mehra, Bharath Vangoor, Anshul Gandhi. UIE: User-centric Interference Estimation for Cloud Applications. In *Proceedings of the 4th IEEE International Conference on Cloud Engineering (IC2E)*, Berlin, Germany, April 2016.
- Anshul Gandhi, Xi Zhang, Naman Mittal. HALO: Heterogeneity-Aware Load Balancing. In *Proceedings of the 23rd IEEE International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*, Atlanta, GA, USA, October 2015.
- Anshul Gandhi, Parijat Dube, Alexei Karve, Andrzej Kochut, Harsha Ellanti. The Unobservability Problem in Clouds. In *Proceedings of the 3rd IEEE International Conference on Cloud and Autonomic Computing (CAC)*, Cambridge, MA, USA, September 2015.

- [Anshul Gandhi](#), Parijat Dube, Alexei Karve, Andrzej Kochut, Li Zhang. Modeling the Impact of Workload on Cloud Resource Scaling. In *Proceedings of the 26th IEEE International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)*, Paris, France, October 2014.
- [Anshul Gandhi](#), Parijat Dube, Alexei Karve, Andrzej Kochut, Li Zhang. Adaptive, Model-driven Autoscaling for Cloud Applications. In *Proceedings of the 11th IEEE International Conference on Autonomic Computing (ICAC)*, Philadelphia, PA, USA, June 2014.
- [Anshul Gandhi](#), Sherwin Doroudi, Mor Harchol-Balter, Alan Scheller-Wolf. Exact Analysis of an M/M/k with Setup Times via Recursive Renewal Reward. In *Proceedings of the 39th ACM International Conference on Measurement and Modeling of Computer Science (SIGMETRICS)*, Pittsburgh, PA, USA, June 2013.
- [Anshul Gandhi](#), Timothy Zhu, Mor Harchol-Balter, Michael Kozuch. SOFTScale: Stealing Opportunistically For Transient Scaling. In *Proceedings of the 13th ACM/IFIP/USENIX International Middleware Conference (Middleware)*, Montreal, Canada, December 2012.
- [Anshul Gandhi](#), Mor Harchol-Balter, Michael Kozuch. Are sleep states effective in data centers? In *Proceedings of the 3rd IEEE International Green Computing Conference (IGCC)*, San Jose, CA, USA, June 2012.
- [Anshul Gandhi](#), Mor Harchol-Balter, Ram Raghunathan, Michael Kozuch. Distributed, Robust Auto-Scaling Policies for Power Management in Compute Intensive Server Farms. In *Proceedings of the 6th Open Cirrus Summit (OCS)*, Atlanta, GA, USA, October 2011.
- [Anshul Gandhi](#) Mor Harchol-Balter. How Data Center Size Impacts the Effectiveness of Dynamic Power Management. In *Proceedings of the 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Allerton, IL, USA, September 2011.
- [Anshul Gandhi](#), Yuan Chen, Daniel Gmach, Martin Arlitt, Manish Marwah. Minimizing Data Center SLA Violations and Power Consumption via Hybrid Resource Provisioning (**Best Paper Award**). In *Proceedings of the 2nd IEEE International Green Computing Conference (IGCC)*, Orlando, FL, USA, July 2011.
- [Anshul Gandhi](#), Varun Gupta, Mor Harchol-Balter, Michael Kozuch. Optimality Analysis of Energy-Performance Trade-off for Server Farm Management. In *Proceedings of the 28th IFIP International Symposium on Computer Performance, Modeling, Measurements and Evaluation (PERFORMANCE)*, Namur, Belgium, November 2010.
- [Anshul Gandhi](#), Mor Harchol-Balter, Ivo Adan. Server Farms with Setup Costs. In *Proceedings of the 28th IFIP International Symposium on Computer Performance, Modeling, Measurements and Evaluation (PERFORMANCE)*, Namur, Belgium, November 2010.
- [Anshul Gandhi](#), Mor Harchol-Balter, Ivo Adan. M/G/k with Exponential Setup. In *Proceedings of the 3rd Madrid Conference on Queueing Theory (MCQT)*, Madrid, Spain, June 2010.
- [Anshul Gandhi](#), Mor Harchol-Balter, Rajarshi Das, Charles Lefurgy. Optimal Power Allocation in Server Farms. In *Proceedings of the 35th ACM International Conference on Measurement and Modeling of Computer Science (SIGMETRICS)*, Seattle, WA, USA, June 2009.
- [Anshul Gandhi](#), R. K. Ghosh. Database Summarization and Publishing in Wireless Environments. In *Proceedings of the 8th International Conference on Distributed Computer Networking (ICDCN)*, Guwahati, India, December 2006.

Refereed Workshop Papers

- Umit Akgun, Santiago Vargas, Michael Arkhangelskiy, Andrew Burford, Michael McNeill, Aruna Balasubramanian, [Anshul Gandhi](#), Erez Zadok. Predicting Network Buffer Capacity for BBR Fairness. In *Proceedings of the 1st Workshop on ML for Systems*, New Orleans, LA, USA, December 2022.

- [Anshul Gandhi](#), Kanad Ghose, Kartik Gopalan, Syed Rafiul Hussain, Dongyoon Lee, Yu David Liu, Zhenhua Liu, Patrick McDaniel, Shuai Mu, Erez Zadok. Metrics for Sustainability in Data Centers. In *Proceedings of the 1st Workshop on Sustainable Computer Systems Design and Implementation (HotCarbon)*, San Diego, CA, USA, July 2022.
- Gagan Somashekar, [Anshul Gandhi](#). Towards Optimal Configuration of Microservices. In *Proceedings of the 1st Workshop on Machine Learning and Systems (EuroMLSys)*, Virtual Event, April 2021.
- Amoghavarsha Suresh, [Anshul Gandhi](#). FnSched: An Efficient Scheduler for Serverless Functions. In *Proceedings of the 5th International Workshop on Serverless Computing (WoSC)*, Davis, CA, USA, December 2019.
- Gagan Somashekar, Mohammad Delasay, [Anshul Gandhi](#). Tighter Lyapunov Truncation for Multi-Dimensional Continuous Time Markov Chains with Known Moments. In *Proceedings of the 21st Workshop on Mathematical Performance Modeling and Analysis (MAMA)*, Phoenix, AZ, USA, June 2019.
- [Anshul Gandhi](#) and [Justin Chan](#). Analyzing the Network for AWS Distributed Cloud Computing. In *Proceedings of the 3rd Annual Workshop on Distributed Cloud Computing (DCC)*, Portland, OR, USA, June 2015.
- Timothy Zhu, [Anshul Gandhi](#), Mor Harchol-Balter, and Michael Kozuch. Saving Cash by Using Less Cache. In *Proceedings of the 4th USENIX Workshop on Hot Topics in Cloud Computing (HotCloud)*, Boston, MA, USA, June 2012.
- [Anshul Gandhi](#), Mor Harchol-Balter, and Michael Kozuch. The Case for Sleep States in Servers. In *Proceedings of the 4th Workshop on Power-Aware Computing and Systems (HotPower)*, Cascais, Portugal, October 2011.
- [Anshul Gandhi](#), Mor Harchol-Balter, and Ivo Adan. Decomposition Results for an M/M/k with Staggered Setup. In *Proceedings of the 12th Workshop on Mathematical Performance Modeling and Analysis (MAMA)*, New York, NY, USA, June 2010.
- [Anshul Gandhi](#), Mor Harchol-Balter, Rajarshi Das, Jeffrey Kephart, and Charles Lefurgy. Power Capping Via Forced Idleness. In *Proceedings of the 1st Workshop on Energy-Efficient Design (WEED)*, Austin, TX, USA, June 2009.

Other Publications

- [Ubaid Ullah Hafeez](#), [Anshul Gandhi](#). SLO-Aware Space-Time GPU Sharing for DL Workloads. Poster abstract in *Proceedings of the 13th ACM Symposium on Cloud Computing (SOCC)*, San Francisco, CA, USA, October 2022.
- [Anshul Gandhi](#), [Amoghavarsha Suresh](#). Leveraging Queueing Theory and OS Profiling to Reduce Application Latency. Invited tutorial at the *20th ACM/IFIP International Middleware Conference (Middleware)*, Davis, CA, USA, December 2019.
- Vasudevan Nagendra, Arani Bhattacharya, [Anshul Gandhi](#), Samir Das. 5GCoreLite: Scalable and Resource Efficient Next Generation Cellular Packet Core. Extended abstract (poster) at the *16th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, Boston, MA, USA, February 2019.
- Seyyed Ahmad Javadi, [Shalini Bhaskara](#), [Rahul Doshi](#), Prashanth Soundarapandian, [Muhammad Wajahat](#), [Anshul Gandhi](#). Application-Agnostic Batch Workload Management in Cloud Environments. Poster abstract in *Proceedings of the 9th ACM Symposium on Cloud Computing (SOCC)*, Carlsbad, CA, USA, October 2018.
- Anjul Tyagi, Ayush Kumar, [Anshul Gandhi](#), Klaus Mueller. Analysis and Visualization of Road Accidents in the UK. Poster at the *IEEE Visualization Conference (VIS)*, Berlin, Germany, October 2018.

- [Ubaid Ullah Hafeez](#), [Deepthi Male](#), [Sharath Kumar Naeni](#), [Muhammad Wajahat](#), and [Anshul Gandhi](#). Realizing an Elastic Memcached via Cached Data Migration. In *Proceedings of the 18th ACM/IFIP/USENIX International Middleware Conference (Middleware): Posters and Demos*, Las Vegas, NV, USA, December 2017.
- [Seyyed Ahmad Javadi](#), [Piyush Shyam Banginwar](#), [Vaishali Chanana](#), [Rashmi Narvekar](#), [Mitesh Kumar Savita](#), and [Anshul Gandhi](#). Improving server utilization via resource-adaptive batch VMs. In *Proceedings of the 18th ACM/IFIP/USENIX International Middleware Conference (Middleware): Posters and Demos*, Las Vegas, NV, USA, December 2017.
- [Yi Cao](#), [Aruna Balasubramanian](#), and [Anshul Gandhi](#). Rethinking TCP Throughput and Latency Modeling. In *Proceedings of the ACM SIGCOMM Posters and Demos*, Los Angeles, CA, USA, August 2017.
- [Seyyed Ahmad Javadi](#), [Himanshu Rajput](#), and [Anshul Gandhi](#). Dynamic Interference-Aware Load Balancing. Poster abstract in *Proceedings of the 7th ACM Symposium on Cloud Computing (SOCC)*, Santa Clara, CA, USA, October 2016.
- [Yi Cao](#), [Javad Nejati](#), [Pavan Maguluri](#), [Aruna Balasubramanian](#), and [Anshul Gandhi](#). Analyzing the Power Consumption of the Mobile Page Load. Extended abstract in *Proceedings of the 42nd ACM International Conference on Measurement and Modeling of Computer Science (SIGMETRICS)*, Antibes, France, June 2016.
- [Anshul Gandhi](#), [Parijat Dube](#), [Andrzej Kochut](#), [Li Zhang](#), and [Sidhartha Thota](#). Model-driven Autoscaling for Hadoop clusters. Poster abstract in *Proceedings of the 6th ACM Symposium on Cloud Computing (SOCC)*, Kohala Coast, HI, USA, August 2015.
- [Anshul Gandhi](#), [Parijat Dube](#), [Andrzej Kochut](#), and [Li Zhang](#). Model-driven Autoscaling for Hadoop clusters. Extended abstract in *Proceedings of the 12th IEEE International Conference on Autonomic Computing (ICAC)*, Grenoble, France, July 2015.
- [Anshul Gandhi](#), [Naman Mittal](#), and [Xi Zhang](#). Optimal Load-Balancing for Heterogeneous Clusters. Extended abstract in *Proceedings of the 3rd Annual Workshop on Distributed Cloud Computing (DCC)*, Portland, OR, USA, June 2015.
- [Anshul Gandhi](#). Dynamic Management of Caching Tiers. Invited abstract in *Proceedings of the 5th ACM/SPEC International Conference on Performance Engineering (ICPE)*, Dublin, Ireland, March 2014.
- [Paul Enders](#), [Anshul Gandhi](#), [Varun Gupta](#), [Laurens Debo](#), [Mor Harchol-Balter](#), and [Alan Scheller-Wolf](#). Inducing Optimal Scheduling with Selfish Users. Technical Report, CMU GSIA Working Paper #2008-E23, Pittsburgh, PA, USA, March 2008.

PATENTS

- [Shankar Narayanan](#), [Bharath Balasubramanian](#), [Gueyoung Jung](#), [Muhammad Wajahat](#), [Anshul Gandhi](#). Model-Driven Technique for Virtual Network Function Rehoming for Service Chains. US Patent 11,212,173, Issued on December 2021.
- [Anshul Gandhi](#), [Hui Lei](#), [Jayaram K. R.](#), [Charles O. Schulz](#), [Shu Tao](#). Graphics Processing Unit Resource Sharing. US Patent 9,830,677, Issued on November 2017.
- [Anshul Gandhi](#), [Parijat Dube](#), [Alexei Karve](#), [Andrzej Kochut](#), [Li Zhang](#). Systems and Methods for Scaling a Cloud Infrastructure. US Patent 9,300,553, Issued on March 2016.
- [Mor Harchol-Balter](#), [Anshul Gandhi](#), [Rajarshi Das](#), [Jeff Kephart](#). Power Budget Allocation in Multi-Processor Systems. US Patent 9,052,895, Issued on June 2015.
- [Mor Harchol-Balter](#), [Anshul Gandhi](#), [Varun Gupta](#), [Mike Kozuch](#). Dynamic Capacity Management of Multiple Parallel Connected Computing Resources. US Patent 8,806,018, Issued on August 2014.

- Mor Harchol-Balter, [Anshul Gandhi](#). Systems and methods for managing power consumption and performance of a processor. US Patent 8,589,709, Issued on November 2013.
- Yuan Chen, [Anshul Gandhi](#), Daniel Gmach, Chris Hyser, Martin Arlitt, Manish Marwah, Cullen Bash. Provisioning Data Center Resources. Patent pending, Filed on April 2011.

INVITED TALKS

- Resource Management and Scheduling in Today's Cloud. *Department of Computer Science and Engineering, The Pennsylvania State University*, November 2022.
- Resource Management and Scheduling in Today's Cloud. *INRIA, Grenoble*, September 2022.
- Modeling TCP congestion control to improve network performance. *Internet Network Architectures Lab, TU Berlin*, September 2022.
- Leveraging Queuing Theory and OS Profiling to Reduce Application Latency. *Invited Tutorial, 20th ACM/IFIP International Middleware Conference (Middleware)*, December 2019.
- Using Variability as a Guiding Principle to Reduce Latency in Web Applications. *Department of Computer Science, Columbia University*, November 2019.
- Using Variability as a Guiding Principle to Reduce Latency in Web Applications. *College of Information and Computer Sciences, University of Massachusetts, Amherst*, October 2019.
- Leveraging theory to guide the design of energy-efficient systems. *ACM Sigmetrics (Rising Star Award Talk)*, June 2019.
- Elastic Memory Caches. *Department of Computer Science and Engineering, The Pennsylvania State University*, October 2018.
- Elastic Memory Caches. *Department of Computer Science and Engineering, The Ohio State University*, October 2018.
- Interference-aware Resource Management in Cloud Environments. *13th Cloud Control Workshop*, June 2018.
- 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,inghamton 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

- EASY: Efficient segment assignment strategy for reducing tail latencies in Pinot. *ICDCS 2018*, July 2018.
- 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.
- Dynamic Management of Caching Tiers. *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.

PROFESSIONAL SERVICE

Editorial Board and Steering Committee

- Editorial Board, Journal of Performance Evaluation (2018 – present)

- Steering Committee Member, ACM/IFIP/USENIX Middleware, 2017 – present

Organizing Committee

- Program Committee Co-Chair, MASCOTS 2023
- Program Committee Co-Chair, Sigmetrics 2021
- Workshops Chair, Sigmetrics 2020
- Track Co-Chair (Intelligent Systems and Infrastructure track), WWW 2020
- Student Activities Chair, Sigmetrics 2019
- Industry and Sponsorship Chair, ICAC 2019
- 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

- 2023: Sigmetrics, Middleware
- 2020: Sigmetrics, ISPASS, ICAC, MASCOTS
- 2019: Sigmetrics, WWW, HPDC
- 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

Panels

- NSF Computer Systems Research (CSR) review panel, 2022
- Moderator and Organizer, ACM Sigmetrics 2019 Student Workshop
- NSF Computer Systems Research (CSR) review panel, 2018
- Panelist, ACM Sigmetrics 2017 Student Workshop
- NSF Computer Systems Research (CSR) review panel, 2015

Refereeing

- ACM Transactions on Computer Systems, IEEE Transactions on Computers, IEEE Transactions on Cloud Computing, ACM/IEEE Transactions on Networking, Performance Evaluation, Annals of Operations Research, Operations Research Letters, 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 Memberships

- IEEE Senior Member
- ACM Senior Member