

NSDI '19: 16th USENIX Symposium on Networked Systems Design and Implementation

February 26–28, 2019

Boston, MA, USA

Host Networking

- Datacenter RPCs can be General and Fast** 1
Anuj Kalia, *Carnegie Mellon University*; Michael Kaminsky, *Intel Labs*; David Andersen, *Carnegie Mellon University*
- Eiffel: Efficient and Flexible Software Packet Scheduling**17
Ahmed Saeed and Yimeng Zhao, *Georgia Institute of Technology*; Nandita Dukkkipati, *Google*; Ellen Zegura and Mostafa Ammar, *Georgia Institute of Technology*; Khaled Harras, *Carnegie Mellon University*; Amin Vahdat, *Google*
- Loom: Flexible and Efficient NIC Packet Scheduling**..... 33
Brent Stephens, *UIC*; Aditya Akella and Michael Swift, *UW-Madison*

Distributed Systems

- Exploiting Commutativity For Practical Fast Replication** 47
Seo Jin Park and John Ousterhout, *Stanford University*
- Flashield: a Hybrid Key-value Cache that Controls Flash Write Amplification** 65
Assaf Eisenman, *Stanford University*; Asaf Cidon, *Stanford University and Barracuda Networks*; Evgenya Pergament and Or Haimovich, *Stanford University*; Ryan Stutsman, *University of Utah*; Mohammad Alizadeh, *MIT CSAIL*; Sachin Katti, *Stanford University*
- Size-aware Sharding For Improving Tail Latencies in In-memory Key-value Stores** 79
Diego Didona, *EPFL*; Willy Zwaenepoel, *EPFL and University of Sydney*
- Monoxide: Scale out Blockchains with Asynchronous Consensus Zones** 95
Jiapeng Wang, *ICT/CAS, Sinovation AI Institute*; Hao Wang, *Ohio State University*

Modern Network Hardware

- FreeFlow: Software-based Virtual RDMA Networking for Containerized Clouds** 113
Daehyeok Kim and Tianlong Yu, *Carnegie Mellon University*; Hongqiang Harry Liu, *Alibaba*; Yibo Zhu, *Microsoft and Bytedance*; Jitu Padhye and Shachar Raindel, *Microsoft*; Chuanxiong Guo, *Bytedance*; Vyas Sekar and Srinivasan Seshan, *Carnegie Mellon University*
- Direct Universal Access: Making Data Center Resources Available to FPGA**..... 127
Ran Shu and Peng Cheng, *Microsoft Research*; Guo Chen, *Microsoft Research & Hunan University*; Zhiyuan Guo, *Microsoft Research & Beihang University*; Lei Qu and Yongqiang Xiong, *Microsoft Research*; Derek Chiou and Thomas Moscibroda, *Microsoft Azure*
- Stardust: Divide and Conquer in the Data Center Network**141
Noa Zilberman, *University of Cambridge*; Gabi Bracha and Golan Schuzkin, *Broadcom*
- Blink: Fast Connectivity Recovery Entirely in the Data Plane**161
Thomas Holterbach, Edgar Costa Molero, and Maria Apostolaki, *ETH Zurich*; Alberto Dainotti, *CAIDA/UC San Diego*; Stefano Vissicchio, *UC London*; Laurent Vanbever, *ETH Zurich*

Analytics

- Hydra: a federated resource manager for data-center scale analytics**177
Carlo Curino, Subru Krishnan, and Konstantinos Karanasos, *Microsoft*; Sriram Rao, *Facebook*; Giovanni M. Fumarola, Botong Huang, Kishore Chaliparambil, Arun Suresh, Young Chen, Solom Heddaya, Roni Burd, Sarvesh Sakalanaga, Chris Douglas, Bill Ramsey, and Raghu Ramakrishnan, *Microsoft*
- Shuffling, Fast and Slow: Scalable Analytics on Serverless Infrastructure** 193
Qifan Pu, *UC Berkeley*; Shivaram Venkataraman, *University of Wisconsin, Madison*; Ion Stoica, *UC Berkeley*

dShark: A General, Easy to Program and Scalable Framework for Analyzing In-network Packet Traces 207
Da Yu, *Brown University*; Yibo Zhu, *Microsoft and Bytedance*; Behnaz Arzani, *Microsoft*; Rodrigo Fonseca, *Brown University*; Tianrong Zhang, Karl Deng, and Lihua Yuan, *Microsoft*

Data Center Network Architecture

Minimal Rewiring: Efficient Live Expansion for Clos Data Center Networks 221
Shizhen Zhao, Rui Wang, Junlan Zhou, Joon Ong, Jeffrey C. Mogul, and Amin Vahdat, *Google, Inc.*

Understanding Lifecycle Management Complexity of Datacenter Topologies 235
Mingyang Zhang, *University of Southern California*; Radhika Niranjana Mysore, *VMware Research*;
Sucha Supittayapornpong and Ramesh Govindan, *University of Southern California*

Shoal: A Network Architecture for Disaggregated Racks 255
Vishal Shrivastav, *Cornell University*; Asaf Valadarsky, *Hebrew University of Jerusalem*; Hitesh Ballani and
Paolo Costa, *Microsoft Research*; Ki Suh Lee, *Waltz Networks*; Han Wang, *Barefoot Networks*; Rachit Agarwal
and Hakim Weatherspoon, *Cornell University*

Wireless Technologies

NetScatter: Enabling Large-Scale Backscatter Networks 271
Mehrddad Hesar, Ali Najafi, and Shyamnath Gollakota, *University of Washington*

Towards Programming the Radio Environment with Large Arrays of Inexpensive Antennas 285
Zhuqi Li, Yaxiong Xie, and Longfei Shangguan, *Princeton University*; Rotman Ivan Zelaya, *Yale University*;
Jeremy Gummeson, *UMass Amherst*; Wenjun Hu, *Yale University*; Kyle Jamieson, *Princeton University*

Pushing the Range Limits of Commercial Passive RFIDs 301
Jingxian Wang, *Carnegie Mellon University*; Junbo Zhang, *Tsinghua University*; Rajarshi Saha, *IIT Kharagpur*;
Haojian Jin and Swarun Kumar, *Carnegie Mellon University*

SweepSense: Sensing 5 GHz in 5 Milliseconds with Low-cost Radios 317
Yeswanth Guddeti, *UC San Diego*; Raghav Subbaraman, *IIT Madras*; Moein Khazraee, Aaron Schulman, and
Dinesh Bharadia, *UC San Diego*

Operating Systems

Slim: OS Kernel Support for a Low-Overhead Container Overlay Network 331
Danyang Zhuo and Kaiyuan Zhang, *University of Washington*; Yibo Zhu, *Microsoft and Bytedance*; Hongqiang Harry
Liu, *Alibaba*; Matthew Rockett, Arvind Krishnamurthy, and Thomas Anderson, *University of Washington*

Shinjuku: Preemptive Scheduling for μ second-scale Tail Latency 345
Kostis Kaffes, Timothy Chong, and Jack Tigar Humphries, *Stanford University*; Adam Belay, *Massachusetts Institute
of Technology*; David Mazières and Christos Kozyrakis, *Stanford University*

Shenango: Achieving High CPU Efficiency for Latency-sensitive Datacenter Workloads 361
Amy Ousterhout, Joshua Fried, Jonathan Behrens, Adam Belay, and Hari Balakrishnan, *MIT CSAIL*

Monitoring and Diagnosis

End-to-end I/O Monitoring on a Leading Supercomputer 379
Bin Yang, *Shandong University, National Supercomputing Center in Wuxi*; Xu Ji, *Tsinghua University, National
Supercomputing Center in Wuxi*; Xiaosong Ma, *Qatar Computing Research Institute, HBKU*; Xiyang Wang, *National
Supercomputing Center in Wuxi*; Tianyu Zhang and Xiupeng Zhu, *Shandong University, National Supercomputing
Center in Wuxi*; Nosayba El-Sayed, *Emory University*; Haidong Lan and Yibo Yang, *Shandong University*; Jidong Zhai,
Tsinghua University; Weiguo Liu, *Shandong University, National Supercomputing Center in Wuxi*; Wei Xue, *Tsinghua
University, National Supercomputing Center in Wuxi*

Zeno: Diagnosing Performance Problems with Temporal Provenance 395
Yang Wu, *Facebook*; Ang Chen, *Rice University*; Linh Thi Xuan Phan, *University of Pennsylvania*

Confluo: Distributed Monitoring and Diagnosis Stack for High-speed Networks 421
Anurag Khandelwal, *UC Berkeley*; Rachit Agarwal, *Cornell University*; Ion Stoica, *UC Berkeley*

(continued on next page)

DETER: Deterministic TCP Replay for Performance Diagnosis 437
Yuliang Li, *Harvard University*; Rui Miao, *Alibaba Group*; Mohammad Alizadeh, *Massachusetts Institute of Technology*;
Minlan Yu, *Harvard University*

Improving Machine Learning

JANUS: Fast and Flexible Deep Learning via Symbolic Graph Execution of Imperative Programs 453
Eunji Jeong, Sungwoo Cho, Gyeong-In Yu, Joo Seong Jeong, Dong-Jin Shin, and Byung-Gon Chun, *Seoul National University*

BLAS-on-flash: An Efficient Alternative for Large Scale ML Training and Inference? 469
Suhas Jayaram Subramanya and Harsha Vardhan Simhadri, *Microsoft Research India*; Srajan Garg, *IIT Bombay*;
Anil Kag and Venkatesh Balasubramanian, *Microsoft Research India*

Tiresias: A GPU Cluster Manager for Distributed Deep Learning. 485
Juncheng Gu, Mosharaf Chowdhury, and Kang G. Shin, *University of Michigan, Ann Arbor*; Yibo Zhu, *Microsoft and Bytedance*; Myeongjae Jeon, *Microsoft and UNIST*; Junjie Qian, *Microsoft*; Hongqiang Liu, *Alibaba*;
Chuanxiong Guo, *Bytedance*

Network Functions

Correctness and Performance for Stateful Chained Network Functions 501
Junaid Khalid and Aditya Akella, *University of Wisconsin - Madison*

Performance Contracts for Software Network Functions 517
Rishabh Iyer, Luis Pedrosa, Arseniy Zaostrovnykh, Solal Pirelli, Katerina Argyraki, and George Candea, *EPFL*

FlowBlaze: Stateful Packet Processing in Hardware 531
Salvatore Pontarelli, *Axbryd/CNIT*; Roberto Bifulco, *NEC Laboratories Europe*; Marco Bonola, *Axbryd/CNIT*;
Carmelo Cascone, *Open Networking Foundation*; Marco Spaziani and Valerio Bruschi, *CNIT/University of Rome Tor Vergata*;
Davide Sanvito, *Politecnico di Milano*; Giuseppe Siracusano, *NEC Laboratories Europe*; Antonio Capone, *Politecnico di Milano*;
Michio Honda and Felipe Huici, *NEC Laboratories Europe*; Giuseppe Bianchi, *CNIT/University of Rome Tor Vergata*

Network Characterization

SIMON: A Simple and Scalable Method for Sensing, Inference and Measurement in Data Center Networks 549
Yilong Geng, Shiyu Liu, and Zi Yin, *Stanford University*; Ashish Naik, *Google Inc.*; Balaji Prabhakar and Mendel Rosenblum, *Stanford University*; Amin Vahdat, *Google Inc.*

Is advance knowledge of flow sizes a plausible assumption? 565
Vojislav Đukić, *ETH Zurich*; Sangeetha Abdu Jyothi, *University of Illinois at Urbana–Champaign*; Bojan Karlaš, Muhsen Owaida, Ce Zhang, and Ankit Singla, *ETH Zurich*

Stable and Practical AS Relationship Inference with ProbLink 581
Yuchen Jin, *University of Washington*; Colin Scott, *UC Berkeley*; Amogh Dhamdhare, *CAIDA*; Vasileios Giotsas, *Lancaster University*; Arvind Krishnamurthy, *University of Washington*; Scott Shenker, *UC Berkeley, ICSI*

NetBouncer: Active Device and Link Failure Localization in Data Center Networks 599
Cheng Tan, *NYU*; Ze Jin, *Cornell University*; Chuanxiong Guo, *Bytedance*; Tianrong Zhang, *Microsoft*; Haitao Wu, *Google*; Karl Deng, Dongming Bi, and Dong Xiang, *Microsoft*

Privacy and Security

Riverbed: Enforcing User-defined Privacy Constraints in Distributed Web Services 615
Frank Wang, *MIT CSAIL*; Ronny Ko and James Mickens, *Harvard University*

Hyperscan: A Fast Multi-pattern Regex Matcher for Modern CPUs 631
Xiang Wang, Yang Hong, and Harry Chang, *Intel*; Kyoungsoo Park, *KAIST*; Geoff Langdale, *branchfree.org*; Jiayu Hu and Heqing Zhu, *Intel*

Deniable Upload and Download via Passive Participation 649
David Sommer, Aritra Dhar, Luka Malisa, and Esfandiar Mohammadi, *ETH Zurich*; Daniel Ronzani, *Ronzani Schlauri Attorneys*; Srdjan Capkun, *ETH Zurich*

CAUDIT: Continuous Auditing of SSH Servers To Mitigate Brute-Force Attacks 667
Phuong M. Cao, Yuming Wu, and Subho S. Banerjee, *UIUC*; Justin Azoff and Alex Withers, *NCSA*; Zbigniew T. Kalbarczyk and Ravishankar K. Iyer, *UIUC*

Network Modeling

Dataplane equivalence and its applications 683
Dragos Dumitrescu, Radu Stoenescu, Matei Popovici, Lorina Negreanu, and Costin Raiciu, *University Politehnica of Bucharest*

Alembic: Automated Model Inference for Stateful Network Functions 699
Soo-Jin Moon, *Carnegie Mellon University*; Jeffrey Helt, *Princeton University*; Yifei Yuan, *Intentionet*; Yves Bieri, *ETH Zurich*; Sujata Banerjee, *VMware Research*; Vyas Sekar, *Carnegie Mellon University*; Wenfei Wu, *Tsinghua University*; Mihalis Yannakakis, *Columbia University*; Ying Zhang, *Facebook, Inc.*

Model-Agnostic and Efficient Exploration of Numerical State Space of Real-World TCP Congestion Control Implementations 719
Wei Sun and Lisong Xu, *University of Nebraska-Lincoln*; Sebastian Elbaum, *University of Virginia*; Di Zhao, *University of Nebraska-Lincoln*

Wireless Applications

Scaling Community Cellular Networks with CommunityCellularManager 735
Shaddi Hasan, *UC Berkeley*; Mary Claire Barela, *University of the Philippines, Diliman*; Matthew Johnson, *University of Washington*; Eric Brewer, *UC Berkeley*; Kurtis Heimerl, *University of Washington*

TrackIO: Tracking First Responders Inside-Out 751
Ashutosh Dhekne, *University of Illinois at Urbana-Champaign*; Ayon Chakraborty, Karthikeyan Sundaresan, and Sampath Rangarajan, *NEC Labs America, Inc.*

3D Backscatter Localization for Fine-Grained Robotics 765
Zhihong Luo, Qiping Zhang, Yunfei Ma, Manish Singh, and Fadel Adib, *MIT Media Lab*

Many-to-Many Beam Alignment in Millimeter Wave Networks 783
Suraj Jog, Jiaming Wang, Junfeng Guan, Thomas Moon, Haitham Hassanieh, and Romit Roy Choudhury, *UIUC*