

# **17th USENIX Symposium on Networked Systems Design and Implementation (NSDI '20)**

Santa Clara, California, USA  
25-27 February 2020

Volume 1 of 2

ISBN: 978-1-7138-1527-3

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2020) by Usenix Association  
All rights reserved.

Printed with permission by Curran Associates, Inc. (2021)

For permission requests, please contact Usenix Association  
at the address below.

Usenix Association  
2560 Ninth Street, Suite 215  
Berkeley, California, 94710

<https://www.usenix.org/>

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2633  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

# NSDI '20: 17th USENIX Symposium on Networked Systems Design and Implementation

February 25–27, 2020

Boston, MA, USA

## Datacenter Networking 1

**Expanding across time to deliver bandwidth efficiency and low latency** ..... 1  
William M. Mellette, Rajdeep Das, Yibo Guo, Rob McGuinness, Alex C. Snoeren, and George Porter, *University of California San Diego*

**Re-architecting Congestion Management in Lossless Ethernet**..... 19  
Wenxue Cheng and Kun Qian, *Tsinghua University and Beijing National Research Center for Information Science and Technology (BNRist)*; Wanchun Jiang, *Central South University*; Tong Zhang, *Tsinghua University, Beijing National Research Center for Information Science and Technology (BNRist), and Nanjing University of Aeronautics and Astronautics*; Fengyuan Ren, *Tsinghua University and Beijing National Research Center for Information Science and Technology (BNRist)*

**Measuring Congestion in High-Performance Datacenter Interconnects** ..... 37  
Saurabh Jha and Archit Patke, *University of Illinois at Urbana-Champaign*; Jim Brandt and Ann Gentile, *Sandia National Lab*; Benjamin Lim, *University of Illinois at Urbana-Champaign*; Mike Showerman and Greg Bauer, *National Center for Supercomputing Applications*; Larry Kaplan, *Cray Inc.*; Zbigniew Kalbarczyk, *University of Illinois at Urbana-Champaign*; William Kramer, *University of Illinois at Urbana-Champaign and National Center for Supercomputing Applications*; Ravi Iyer, *University of Illinois at Urbana-Champaign*

**SP-PIFO: Approximating Push-In First-Out Behaviors using Strict-Priority Queues**..... 59  
Albert Gran Alcoz, Alexander Dietmüller, and Laurent Vanbever, *ETH Zürich*

## OS, Storage, and Hardware

**AccelTCP: Accelerating Network Applications with Stateful TCP Offloading** ..... 77  
YoungGyou Moon and SeungEon Lee, *KAIST*; Muhammad Asim Jamshed, *Intel Labs*; KyoungSoo Park, *KAIST*

**Enabling Programmable Transport Protocols in High-Speed NICs**..... 93  
Mina Tahmasbi Arashloo and Alexey Lavrov, *Princeton University*; Manya Ghobadi, *MIT*; Jennifer Rexford, David Walker, and David Wentzlaff, *Princeton University*

**FileMR: Rethinking RDMA Networking for Scalable Persistent Memory** ..... 111  
Jian Yang, *UC San Diego*; Joseph Izraelevitz, *University of Colorado, Boulder*; Steven Swanson, *UC San Diego*

**TCP  $\approx$  RDMA: CPU-efficient Remote Storage Access with i10**..... 127  
Jaehyun Hwang, Qizhe Cai, Ao Tang, and Rachit Agarwal, *Cornell University*

**NetTLP: A Development Platform for PCIe devices in Software Interacting with Hardware** .....141  
Yohei Kuga and Ryo Nakamura, *The University of Tokyo*; Takeshi Matsuya, *Keio University*; Yuji Sekiya, *The University of Tokyo*

**Near-Optimal Latency Versus Cost Tradeoffs in Geo-Distributed Storage**..... 157  
Muhammed Uluyol, Anthony Huang, Ayush Goel, Mosharaf Chowdhury, and Harsha V. Madhyastha, *University of Michigan*

## Network Verification

**NetSMC: A Custom Symbolic Model Checker for Stateful Network Verification**..... 181  
Yifei Yuan, *Intentionet*; Soo-Jin Moon, Sahil Uppal, Limin Jia, and Vyas Sekar, *Carnegie Mellon University*

**Tiramisu: Fast Multilayer Network Verification**..... 201  
Anubhavnidhi Abhashkumar, *University of Wisconsin - Madison*; Aaron Gember-Jacobson, *Colgate University*; Aditya Akella, *University of Wisconsin - Madison*

**Automated Verification of Customizable Middlebox Properties with Gravel** ..... 221  
Kaiyuan Zhang, *University of Washington*; Danyang Zhuo, *Duke University*; Aditya Akella, *University of Wisconsin-Madison*; Arvind Krishnamurthy and Xi Wang, *University of Washington*

**APKeep: Realtime Verification for Real Networks** ..... 241  
Peng Zhang and Xu Liu, *Xi'an Jiaotong University*; Hongkun Yang, *Google*; Ning Kang, Zhengchang Gu, and Hao Li, *Xi'an Jiaotong University*

**Liveness Verification of Stateful Network Functions** ..... 257  
Farnaz Yousefi, *Johns Hopkins University*; Anubhavnidhi Abhashkumar and Kausik Subramanian, *University of Wisconsin-Madison*; Kartik Hans, *IIT Delhi*; Soudeh Ghorbani, *Johns Hopkins University*; Aditya Akella, *University of Wisconsin-Madison*

## Distributed Systems

**Sol: Fast Distributed Computation Over Slow Networks** ..... 273  
Fan Lai, Jie You, Xiangfeng Zhu, Harsha V. Madhyastha, and Mosharaf Chowdhury, *University of Michigan*

**THEMIS: Fair and Efficient GPU Cluster Scheduling** ..... 289  
Kshiteej Mahajan, Arjun Balasubramanian, Arjun Singhvi, Shivaram Venkataraman, and Aditya Akella, *University of Wisconsin-Madison*; Amar Phanishayee, *Microsoft Research*; Shuchi Chawla, *University of Wisconsin-Madison*

**Fine-Grained Replicated State Machines for a Cluster Storage System** ..... 305  
Ming Liu and Arvind Krishnamurthy, *University of Washington*; Harsha V. Madhyastha, *University of Michigan*; Rishi Bhardwaj, Karan Gupta, Chinmay Kamat, Huapeng Yuan, Aditya Jaltade, Roger Liao, Pavan Konka, and Anoop Jawahar, *Nutanix*

**Scalog: Seamless Reconfiguration and Total Order in a Scalable Shared Log** ..... 325  
Cong Ding, David Chu, and Evan Zhao, *Cornell University*; Xiang Li, *Alibaba Group*; Lorenzo Alvisi and Robbert van Renesse, *Cornell University*

## Wireless Networks 1

**Frequency Configuration for Low-Power Wide-Area Networks in a Heartbeat** ..... 339  
Akshay Gadre, *Carnegie Mellon University*; Revathy Narayanan, *Carnegie Mellon University and IIT Madras*; Anh Luong, Anthony Rowe, Bob Iannucci, and Swarun Kumar, *Carnegie Mellon University*

**ABC: A Simple Explicit Congestion Controller for Wireless Networks** ..... 353  
Prateesh Goyal, *MIT CSAIL*; Anup Agarwal, *CMU*; Ravi Netravali, *UCLA*; Mohammad Alizadeh and Hari Balakrishnan, *MIT CSAIL*

**Amphilight: Direct Air-Water Communication with Laser Light** ..... 373  
Charles J. Carver and Zhao Tian, *Department of Computer Science, Dartmouth College*; Hongyong Zhang and Kofi M. Odame, *Thayer School of Engineering, Dartmouth College*; Alberto Quattrini Li and Xia Zhou, *Department of Computer Science, Dartmouth College*

## Deployment Experience

**Gandalf: An Intelligent, End-To-End Analytics Service for Safe Deployment in Large-Scale Cloud Infrastructure** .. 389  
Ze Li, Qian Cheng, Ken Hsieh, and Yingnong Dang, *Microsoft Azure*; Peng Huang, *Johns Hopkins University*; Pankaj Singh and Xinsheng Yang, *Microsoft Azure*; Qingwei Lin, *Microsoft Research*; Youjiang Wu, Sebastien Levy, and Murali Chintalapati, *Microsoft Azure*

**Experiences with Modeling Network Topologies at Multiple Levels of Abstraction** ..... 403  
Jeffrey C. Mogul, Drago Goricanec, Martin Pool, Anees Shaikh, Douglas Turk, and Bikash Koley, *Google LLC*; Xiaoxue Zhao, *Alibaba Group Inc.*

**Firecracker: Lightweight Virtualization for Serverless Applications** ..... 419  
Alexandru Agache, Marc Brooker, Andreea Florescu, Alexandra Iordache, Anthony Liguori, Rolf Neugebauer, Phil Pivonka, and Diana-Maria Popa, *Amazon Web Services*

**Rex: Preventing Bugs and Misconfiguration in Large Services Using Correlated Change Analysis** ..... 435  
Sonu Mehta, Ranjita Bhagwan, Rahul Kumar, Chetan Bansal, Chandra Maddila, B. Ashok, and Sumit Asthana, *Microsoft Research India*; Christian Bird, *Microsoft Research Redmond*; Aditya Kumar, *Microsoft Research India*

**Building An Elastic Query Engine on Disaggregated Storage** ..... 449  
Midhul Vuppapalapati, Justin Miron, and Rachit Agarwal, *Cornell University*; Dan Truong, Ashish Motivala, and Thierry Cruanes, *Snowflake Computing*

<b>Millions of Tiny Databases</b> .....	<b>463</b>
Marc Brooker, Tao Chen, and Fan Ping, <i>Amazon Web Services</i>	
<b>Measurement and Adaptation</b>	
<b>Diamond-Miner: Comprehensive Discovery of the Internet's Topology Diamonds</b> .....	<b>479</b>
Kevin Vermeulen, <i>Sorbonne Université</i> ; Justin P. Rohrer and Robert Beverly, <i>Naval Postgraduate School</i> ; Olivier Fourmaux and Timur Friedman, <i>Sorbonne Université</i>	
<b>Learning <i>in situ</i>: a randomized experiment in video streaming</b> .....	<b>495</b>
Francis Y. Yan and Hudson Ayers, <i>Stanford University</i> ; Chenzhi Zhu, <i>Tsinghua University</i> ; Sadjad Fouladi, James Hong, Keyi Zhang, Philip Levis, and Keith Winstein, <i>Stanford University</i>	
<b>Is Big Data Performance Reproducible in Modern Cloud Networks?</b> .....	<b>513</b>
Alexandru Uta and Alexandru Custura, <i>Vrije Universiteit Amsterdam</i> ; Dmitry Duplyakin, <i>University of Utah</i> ; Ivo Jimenez, <i>UC Santa Cruz</i> ; Jan Rellermeyer, <i>TU Delft</i> ; Carlos Maltzahn, <i>UC Santa Cruz</i> ; Robert Ricci, <i>University of Utah</i> ; Alexandru Iosup, <i>Vrije Universiteit Amsterdam</i>	
<b>Learning Relaxed Belady for Content Distribution Network Caching</b> .....	<b>529</b>
Zhenyu Song, <i>Princeton University</i> ; Daniel S. Berger, <i>Microsoft Research &amp; Carnegie Mellon University</i> ; Kai Li and Wyatt Lloyd, <i>Princeton University</i>	
<b>Fault Tolerance and Availability</b>	
<b>Meaningful Availability</b> .....	<b>545</b>
Tamás Hauer, Philipp Hoffmann, John Lunney, Dan Ardelean, and Amer Diwan, <i>Google</i>	
<b>Understanding, Detecting and Localizing Partial Failures in Large System Software</b> .....	<b>559</b>
Chang Lou, Peng Huang, and Scott Smith, <i>Johns Hopkins University</i>	
<b>Check before You Change: Preventing Correlated Failures in Service Updates</b> .....	<b>575</b>
Ennan Zhai, <i>Alibaba Group</i> ; Ang Chen, <i>Rice University</i> ; Ruzica Piskac, <i>Yale University</i> ; Mahesh Balakrishnan, <i>Facebook</i> ; Bingchuan Tian, <i>Nanjing University</i> ; Bo Song and Haoliang Zhang, <i>Google</i>	
<b>Gryff: Unifying Consensus and Shared Registers</b> .....	<b>591</b>
Matthew Burke, <i>Cornell University</i> ; Audrey Cheng and Wyatt Lloyd, <i>Princeton University</i>	
<b>CableMon: Improving the Reliability of Cable Broadband Networks via Proactive Network Maintenance</b> .....	<b>619</b>
Jiyao Hu, Zhenyu Zhou, and Xiaowei Yang, <i>Duke University</i> ; Jacob Malone, <i>CableLabs</i> ; Jonathan W Williams, <i>The University of North Carolina at Chapel Hill</i>	
<b>Datacenter Networking 2</b>	
<b>Batchy: Batch-scheduling Data Flow Graphs with Service-level Objectives</b> .....	<b>633</b>
Tamás Lévai, <i>Budapest University of Technology and Economics &amp; University of Southern California</i> ; Felicián Németh, <i>Budapest University of Technology and Economics</i> ; Barath Raghavan, <i>University of Southern California</i> ; Gábor Rétvári, <i>MTA-BME Information Systems Research Group &amp; Ericsson Research, Hungary</i>	
<b>Adapting TCP for Reconfigurable Datacenter Networks</b> .....	<b>651</b>
Matthew K. Mukerjee, <i>Carnegie Mellon University / Nefeli Networks</i> ; Christopher Canel, <i>Carnegie Mellon University</i> ; Weiyang Wang, <i>UC San Diego</i> ; Daehyeok Kim, <i>Carnegie Mellon University / Microsoft Research</i> ; Srinivasan Seshan, <i>Carnegie Mellon University</i> ; Alex C. Snoeren, <i>UC San Diego</i>	
<b>A High-Speed Load-Balancer Design with Guaranteed Per-Connection-Consistency</b> .....	<b>667</b>
Tom Barbette, Chen Tang, Haoran Yao, Dejan Kostić, Gerald Q. Maguire Jr., Panagiotis Papadimitratos, and Marco Chiesa, <i>KTH Royal Institute of Technology</i>	
<b>Programmable Calendar Queues for High-speed Packet Scheduling</b> .....	<b>685</b>
Naveen Kr. Sharma, Chenxingyu Zhao, and Ming Liu, <i>University of Washington</i> ; Pravein G Kannan, <i>School of Computing, National University of Singapore</i> ; Changhoon Kim, <i>Barefoot Networks</i> ; Arvind Krishnamurthy, <i>University of Washington</i> ; Anirudh Sivaraman, <i>NYU</i>	

## Routing

- Contra: A Programmable System for Performance-aware Routing** . . . . . 701  
Kuo-Feng Hsu, *Rice University*; Ryan Beckett, *Microsoft Research*; Ang Chen, *Rice University*; Jennifer Rexford, Praveen Tamma, and David Walker, *Princeton University*
- FLAIR: Accelerating Reads with Consistency-Aware Network Routing** . . . . . 723  
Hatem Takturi, Ibrahim Kettaneh, Ahmed Alquraan, and Samer Al-Kiswany, *University of Waterloo*
- Towards Logically Centralized Interdomain Routing** . . . . . 739  
Shahrooz Pouryousef, Lixin Gao, and Arun Venkataramani, *University of Massachusetts at Amherst*

## Security

- XRD: Scalable Messaging System with Cryptographic Privacy** . . . . . 759  
Albert Kwon, *MIT*; David Lu, *MIT PRIMES*; Srinivas Devadas, *MIT*
- High Throughput Cryptocurrency Routing in Payment Channel Networks** . . . . . 777  
Vibhaalakshmi Sivaraman, *Massachusetts Institute of Technology*; Shaileshh Bojja Venkatakrishnan, *Ohio State University*; Kathleen Ruan, *Carnegie Mellon University*; Parimarjan Negi and Lei Yang, *Massachusetts Institute of Technology*; Radhika Mittal, *University of Illinois at Urbana-Champaign*; Giulia Fanti, *Carnegie Mellon University*; Mohammad Alizadeh, *Massachusetts Institute of Technology*
- PrivateEye: Scalable and Privacy-Preserving Compromise Detection in the Cloud** . . . . . 797  
Behnaz Arzani, *Microsoft Research*; Selim Ciraci, *Microsoft*; Stefan Saroiu, Alec Wolman, and Jack Stokes, *Microsoft Research*; Geoff Outhred and Lechao Diwu, *Microsoft*
- Telekine: Secure Computing with Cloud GPUs** . . . . . 817  
Tyler Hunt, Zhipeng Jia, Vance Miller, Ariel Szekely, and Yige Hu, *The University of Texas at Austin*; Christopher J. Rossbach, *The University of Texas at Austin and VMware Research*; Emmett Witchel, *The University of Texas at Austin*
- TimeCrypt: Encrypted Data Stream Processing at Scale with Cryptographic Access Control** . . . . . 835  
Lukas Burkhalter, *ETH Zurich*; Anwar Hithnawi, *UC Berkeley, ETH Zurich*; Alexander Viand and Hossein Shafagh, *ETH Zurich*; Sylvia Ratnasamy, *UC Berkeley*
- Ghstor: Toward a Secure Data-Sharing System from Decentralized Trust** . . . . . 851  
Yuncong Hu, Sam Kumar, and Raluca Ada Popa, *University of California, Berkeley*

## Wireless Networks 2

- Fawkes: Faster Mobile Page Loads via App-Inspired Static Templating** . . . . . 879  
Shaghayegh Mardani, *UCLA*; Mayank Singh, *IIT Delhi*; Ravi Netravali, *UCLA*
- VMscatter: A Versatile MIMO Backscatter** . . . . . 895  
Xin Liu, Zicheng Chi, Wei Wang, Yao Yao, and Ting Zhu, *University of Maryland, Baltimore County*
- Performant TCP for Low-Power Wireless Networks** . . . . . 911  
Sam Kumar, Michael P Andersen, Hyung-Sin Kim, and David E. Culler, *University of California, Berkeley*
- Comb Decoding towards Collision-Free WiFi** . . . . . 933  
Shangqing Zhao, Zhe Qu, Zhengping Luo, Zhuo Lu, and Yao Liu, *University of South Florida*

## Debugging

- Plankton: Scalable network configuration verification through model checking** . . . . . 953  
Santhosh Prabhu, Kuan-Yen Chou, Ali Kheradmand, P. Brighten Godfrey, and Matthew Caesar, *University of Illinois at Urbana-Champaign*
- Config2Spec: Mining Network Specifications from Network Configurations** . . . . . 969  
Rüdiger Birkner, *ETH Zürich*; Dana Drachler-Cohen, *Technion*; Laurent Vanbever and Martin Vechev, *ETH Zürich*
- Network Error Logging: Client-side measurement of end-to-end web service reliability** . . . . . 985  
Sam Burnett and Lily Chen, *Google*; Douglas A. Creager, *GitHub*; Misha Efimov, Ilya Grigorik, and Ben Jones, *Google*; Harsha V. Madhyastha, *Google and University of Michigan*; Pavlos Papageorge, Brian Rogan, Charles Stahl, and Julia Tuttle, *Google*

**Finding Network Misconfigurations by Automatic Template Inference** ..... 999  
Siva Kesava Reddy Kakarla and Alan Tang, *UCLA*; Ryan Beckett, *Microsoft Research*; Karthick Jayaraman, *Microsoft Azure*; Todd Millstein, *UCLA / Intentionet*; Yuval Tamir and George Varghese, *UCLA*

**tpprof: A Network Traffic Pattern Profiler** .....1015  
Nofel Yaseen, John Sonchack, and Vincent Liu, *University of Pennsylvania*

## Sensor Networks

**TinySDR: Low-Power SDR Platform for Over-the-Air Programmable IoT Testbeds** .....1031  
Mehrdad Hesar, Ali Najafi, Vikram Iyer, and Shyamnath Gollakota, *University of Washington*

**RFocus: Beamforming Using Thousands of Passive Antennas** ..... 1047  
Venkat Arun and Hari Balakrishnan, *Massachusetts Institute of Technology*

**CarMap: Fast 3D Feature Map Updates for Automobiles** ..... 1063  
Fawad Ahmad and Hang Qiu, *University of Southern California*; Ray Eells, *California State Polytechnic University, Pomona*; Fan Bai, *General Motors*; Ramesh Govindan, *University of Southern California*

**Food and Liquid Sensing in Practical Environments using RFIDs**..... 1083  
Unsoo Ha, Junshan Leng, and Alaa Khaddaj, and Fadel Adib, *Massachusetts Institute of Technology*

**Eingerprint: Robust Energy-related Fingerprinting for Passive RFID Tags**.....1101  
Xingyu Chen, Jia Liu, Xia Wang, Haisong Liu, Dong Jiang, and Lijun Chen, *Nanjing University*

**LocAP: Autonomous Millimeter Accurate Mapping of WiFi Infrastructure** .....1115  
Roshan Ayyalasomayajula, Aditya Arun, Chenfeng Wu, Shrivatsan Rajagopalan, Shreya Ganesaraman, Aravind Seetharaman, and Ish Kumar Jain, and Dinesh Bharadia, *University of California, San Diego*