

Wei Bai

14820 NE 36th Street, Building 99,
Redmond, Washington, 98052 USA
✉ baiwei0427@gmail.com
📄 [baiwei0427.github.io](https://github.com/baiwei0427.github.io)

Research Interests

I am broadly interested in computer networking, with special focuses on data center networking, congestion control, and high performance networked systems.

Education

- 2013-2017 **The Hong Kong University of Science and Technology (HKUST)**, Hong Kong SAR, China
Ph.D. Computer Science and Engineering
Advisor: Prof. Kai Chen
Thesis: "Congestion Control Mechanisms for Data Center Networks"
- 2009-2013 **Shanghai Jiao Tong University (SJTU)**, Shanghai, China
B.E. Information Security

Work Experience

- 2019-Present **Microsoft Research Redmond**, Redmond, USA
Research Software Development Engineer, Mobility and Networking Research Group
- 2017-2019 **Microsoft Research Asia**, Beijing, China
Researcher (formerly Associate Researcher 2), Networking Research Group

Technology Transfer at Microsoft

- 2019-Present **RDMA congestion control**: I am working with Azure networking to design and implement high performance congestion control mechanisms to support large scale remote direct memory access (RDMA) deployments in Microsoft data centers.
- 2018-2019 **SONiC chassis**: I worked with Azure networking to design and implement a disaggregated chassis switch for SONiC. Different from traditional proprietary chassis switches, SONiC chassis disaggregates the chassis switch into a Clos network with commodity packet switching chips and uses SONiC as the switch operating system.

Awards and Honors

- 2016, 2017 HKUST Research Travel Grant
2015, 2016 USENIX NSDI Student Grant
2015 **Microsoft Research Asia Fellowship**
2013-2017 HKUST Postgraduate Scholarship
2013 SJTU Outstanding Bachelor Thesis Award (Top 1%)

2012 First Prize, Software Group, Information Security Contest, National Undergraduate Electronic Design Contest

Publications

Conference Publications

- [C21] Shuihai Hu, **Wei Bai**, Gaoxiong Zeng, Zilong Wang, Baochen Qiao, Kai Chen, Kun Tan, Yi Wang, "Aeolus: A Building Block for Proactive Transport in Datacenters", in *Proceedings of the ACM SIGCOMM 2020 Conference (SIGCOMM)*, Virtual Conference, August 10-14, 2020.
- [C20] Qun Huang, Haifeng Sun, Patrick P. C. Lee, **Wei Bai**, Feng Zhu, Yungang Bao, "OmniMon: Re-architecting Network Telemetry with Resource Efficiency and Full Accuracy", in *Proceedings of the ACM SIGCOMM 2020 Conference (SIGCOMM)*, Virtual Conference, August 10-14, 2020.
- [C19] **Wei Bai**, Shuihai Hu, Kai Chen, Kun Tan, Yongqiang Xiong, "One More Config is Enough: Saving (DC)TCP for High-speed Extremely Shallow-buffered Datacenters", in *Proceedings of the 39th IEEE International Conference on Computer Communications (INFOCOM)*, Virtual Conference, July 6-9, 2020.
- [C18] Junxue Zhang, **Wei Bai**, Kai Chen, "Enabling ECN for Datacenter Networks with RTT Variations", in *Proceedings of the 15th International Conference on emerging Networking EXperiments and Technologies (CoNEXT)*, Orlando, Florida, December 9-12, 2019.
- [C17] Gaoxiong Zeng, **Wei Bai**, Ge Chen, Kai Chen, Dongsu Han, Yibo Zhu, Lei Cui, "Congestion Control for Cross-Datacenter Networks", in *Proceedings of the 27th IEEE International Conference on Network Protocols (ICNP)*, Chicago, Illinois, October 7-10, 2019.
- [C16] Xiaodong Yi, Junjie Wang, Jingpu Duan, **Wei Bai**, Chuan Wu, Yongqiang Xiong, Dongsu Han, "FlowShader: a Generalized Framework for GPU-accelerated VNF Flow Processing", in *Proceedings of the 27th IEEE International Conference on Network Protocols (ICNP)*, Chicago, Illinois, October 7-10, 2019.
- [C15] Bojie Li, Tianyi Cui, Zibo Wang, **Wei Bai**, Lintao Zhang, "SocksDirect: Datacenter Sockets can be Fast and Compatible" in *Proceedings of the ACM SIGCOMM 2019 Conference (SIGCOMM)*, Beijing, China, August 19-24, 2019.
- [C14] Yang Cheng, Dan Li, Zhiyuan Guo, Binyao Jiang, Jiaxin Lin, Xi Fan, Jinkun Geng, Xinyi Yu, **Wei Bai**, Lei Qu, Ran Shu, Peng Cheng, Yongqiang Xiong, Jianping Wu, "DLBooster: Boosting End-to-End Deep Learning Workflows with Offloading Data Preprocessing Pipelines" in *Proceedings of the 48th International Conference on Parallel Processing (ICPP)*, Kyoto, Japan, August 5-8, 2019.
- [C13] Zhao Lucis Li, Mike Chieh-Jan Liang, **Wei Bai**, Qiming Zheng, Yongqiang Xiong, Guangzhong Sun, "Accelerating Rule-matching Systems with Learned Rankers" in *Proceedings of the 2019 USENIX Annual Technical Conference (ATC)*, Renton, Washington, July 10-12, 2019 (Short Paper).

- [C12] Hong Zhang, Junxue Zhang, **Wei Bai**, Kai Chen, Mosharaf Chowdhury, "Resilient Datacenter Load Balancing in the Wild" in *Proceedings of the ACM SIGCOMM 2017 Conference (SIGCOMM)*, Los Angeles, California, August 21-25, 2017.
- [C11] Ziyang Li, **Wei Bai**, Kai Chen, Dongsu Han, Yiming Zhang, Dongsheng Li, Hongfang Yu, "Rate-Aware Flow Scheduling for Commodity Data Center Networks" in *Proceedings of the 36th Annual IEEE International Conference on Computer Communications (INFOCOM)*, Atlanta, Georgia, May 1-4, 2017.
- [C10] **Wei Bai**, Kai Chen, Li Chen, Changhoon Kim, Haitao Wu, "Enabling ECN over Generic Packet Scheduling" in *Proceedings of the 12th International Conference on emerging Networking EXperiments and Technologies (CoNEXT)*, Irvine, California, December 12-15, 2016.
- [C9] Li Chen, Kai Chen, **Wei Bai**, Mohammad Alizadeh "Scheduling Mix-flows in Commodity Datacenters with Karuna" in *Proceedings of the ACM SIGCOMM 2016 Conference (SIGCOMM)*, Florianopolis, Brazil, August 22-26, 2016.
- [C8] Shuihai Hu, **Wei Bai**, Kai Chen, Chen Tian, Ying Zhang, Haitao Wu, "Providing Bandwidth Guarantees, Work Conservation and Low Latency Simultaneously in the Cloud", in *Proceedings of the 35th Annual IEEE International Conference on Computer Communications (INFOCOM)*, San Francisco, California, April 10-14, 2016.
- [C7] **Wei Bai**, Li Chen, Kai Chen, Haitao Wu "Enabling ECN in Multi-Service Multi-Queue Data Centers" in *Proceedings of the 13th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, Santa Clara, California, March 16-18, 2016.
- [C6] **Wei Bai**, Li Chen, Kai Chen, Dongsu Han, Chen Tian, Hao Wang, "Information-Agnostic Flow Scheduling for Commodity Data Centers" in *Proceedings of the 12th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, Oakland, California, May 4-6, 2015.
- [C5] Shuihai Hu, Kai Chen, Haitao Wu, **Wei Bai**, Chang Lan, Hao Wang, Hongze Zhao, Chuanxiong Guo, "Explicit Path Control in Commodity Data Centers: Design and Applications" in *Proceedings of the 12th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, Oakland, California, May 4-6, 2015.
- [C4] Yangming Zhao, Kai Chen, **Wei Bai**, Minlan Yu, Chen Tian, Yanhui Geng, Yiming Zhang, Dan Li, Sheng Wang, "RAPIER: Integrating Routing and Scheduling for Coflow-aware Data Center Networks" in *Proceedings of the 34th Annual IEEE International Conference on Computer Communications (INFOCOM)*, Hong Kong, April 26-May 1, 2015.
- [C3] Hong Zhang, Kai Chen, **Wei Bai**, Dongsu Han, Chen Tian, Hao Wang, Haibing Guan, Ming Zhang, "Guaranteeing Deadlines for Inter-Datacenter Transfers" in *Proceedings of the 10th European Conference on Computer Systems (EuroSys)*, Bordeaux, France, April 21-24, 2015.
- [C2] **Wei Bai**, Kai Chen, Haitao Wu, Wuwei Lan, Yangming Zhao, "PAC: Taming TCP Incast Congestion Using Proactive ACK Control" in *Proceedings of the IEEE 22nd International Conference on Network Protocols (ICNP)*, Research Triangle, North Carolina, October 21-24, 2014.

- [C1] Yang Peng, Kai Chen, Guohui Wang, **Wei Bai**, Zhiqiang Ma, Lin Gu, "HadoopWatch: A First Step Towards Comprehensive Traffic Forecasting in Cloud Computing" in *Proceedings of the 33rd Annual IEEE International Conference on Computer Communications (INFOCOM)*, Toronto, Canada, April 27-May 2, 2014.

Workshop Publications

- [W5] Jiacheng Xia, Gaoxiong Zeng, Junxue Zhang, Weiyan Wang, **Wei Bai**, Junchen Jiang, Kai Chen, "Rethinking transport layer design for distributed machine learning" in *Proceedings of the 3rd Asia-Pacific Workshop on Networking (APNet)*, Beijing, China, August 17-18, 2019.
- [W4] Shuihai Hu, **Wei Bai**, Baochen Qiao, Kai Chen, Kun Tan, "Augmenting Proactive Congestion Control with Aeolus" in *Proceedings of the 2nd Asia-Pacific Workshop on Networking (APNet)*, Beijing, China, August 2-3, 2018.
- [W3] Gaoxiong Zeng, **Wei Bai**, Ge Chen, Kai Chen, Dongsu Han, Yibo Zhu, "Combining ECN and RTT for Datacenter Transport" in *Proceedings of the 1st Asia-Pacific Workshop on Networking (APNet)*, Hong Kong, China, August 3-4, 2017.
- [W2] **Wei Bai**, Kai Chen, Shuihai Hu, Kun Tan, Yongqiang Xiong, "Congestion Control for High-speed Extremely Shallow-buffered Datacenter Networks" in *Proceedings of the 1st Asia-Pacific Workshop on Networking (APNet)*, Hong Kong, China, August 3-4, 2017.
- [W1] **Wei Bai**, Li Chen, Kai Chen, Dongsu Han, Chen Tian, Weicheng Sun, "PIAS: Practical Information-Agnostic Flow Scheduling for Data Center Networks" in *Proceedings of the 13th ACM Workshop on Hot Topics in Networks (HotNets)*, Los Angeles, California, October 27-28, 2014.

Journal Publications

- [J6] Shuihai Hu, **Wei Bai**, Kai Chen, Chen Tian, Ying Zhang, Haitao Wu, "Providing Bandwidth Guarantees, Work Conservation and Low Latency Simultaneously in the Cloud", in *IEEE Transactions on Cloud Computing (TCC)*, 2018.
- [J5] Shuhao Liu, Hong Xu, Libin Liu, **Wei Bai**, Kai Chen, Zhiping Cai, "RepNet: Cutting Latency with Flow Replication in Data Center Networks", in *IEEE Transactions on Services Computing (TSC)*, 2018.
- [J4] **Wei Bai**, Li Chen, Kai Chen, Dongsu Han, Chen Tian, Hao Wang, "PIAS: Practical Information-Agnostic Flow Scheduling for Commodity Data Centers" in *IEEE/ACM Transactions on Networking (ToN)*, 2017.
- [J3] Hong Zhang, Kai Chen, **Wei Bai**, Dongsu Han, Chen Tian, Hao Wang, Haibing Guan, Ming Zhang, "Guaranteeing Deadlines for Inter-Datacenter Transfers" in *IEEE/ACM Transactions on Networking (ToN)*, 2017.
- [J2] Shuihai Hu, Kai Chen, Haitao Wu, **Wei Bai**, Chang Lan, Hao Wang, Hongze Zhao, Chuanxiong Guo, "Explicit Path Control in Commodity Data Centers: Design and Applications" in *IEEE/ACM Transactions on Networking (ToN)*, 2016.
- [J1] Yang Peng, Kai Chen, Guohui Wang, **Wei Bai**, Yangming Zhao, Hao Wang, Yanhui Geng, Zhiqiang Ma, Lin Gu, "Towards Comprehensive Traffic Forecasting in Cloud Computing: Design and Application" in *IEEE/ACM Transactions on Networking (ToN)*, 2016.

Selected Talks

- Jan. 2019 Build Reliable Cloud Networks with SONiC and ONE, OCP China Technology Day, Shenzhen, China
- Oct. 2018 Congestion Control Mechanisms for Data Center Networks, HotDC 2018, Beijing, China
- Aug. 2017 Experiments with Data Center Congestion Control Research, APNet 2017, Hong Kong, China
- Dec. 2016 Enabling ECN over Generic Packet Scheduling, CoNEXT 2016, Irvine, California
- Mar. 2016 Enabling ECN in Multi-Service Multi-Queue Data Centers, NSDI 2016, Santa Clara, California
- May 2015 Information-Agnostic Flow Scheduling for Commodity Data Centers, NSDI 2015, Oakland, California

Professional Activities

Technical Program Committee

- The 4th Asia-Pacific Workshop on Networking (**APNet 2020**)
- The 11th ACM SIGOPS Asia-Pacific Workshop on Systems (**APSys 2020**)
- The 3rd Asia-Pacific Workshop on Networking (**APNet 2019**)
- The 38th IEEE International Conference on Distributed Computing Systems (**ICDCS 2018**), "Cloud Computing and Data Centers" Track

Organizing Committee

- Publicity Chair, The 4th Asia-Pacific Workshop on Networking (**APNet 2020**)
- Registration Chair, ACM SIGCOMM 2019 Conference (**SIGCOMM 2019**)
- Web Chair, The 3rd Asia-Pacific Workshop on Networking (**APNet 2019**)
- Web Chair, The 2nd Asia-Pacific Workshop on Networking (**APNet 2018**)

Reviewer

- IEEE/ACM Transactions on Networking
- Elsevier Computer Networks
- IEEE Transactions on Network and Service Management
- IEEE Communications Letters
- ACM Transactions on Storage
- IEEE Transactions on Cloud Computing
- ACM Transactions on Architecture and Code Optimization
- Elsevier Future Generation Computer Systems

References

Available upon request