Xusheng CHEN

Room 312
Chow Yei Ching Building
The University of Hong Kong
Pokfulam, Hong Kong Island
Hong Kong SAR

Phone: (+852) 92390400
(+86) 18018527062
Email: xschen@cs.hku.hk
michael.xschen@gmail.com
Homepage: http://i.cs.hku.hk/∼xschen

EDUCATION

- Ph.D. candidate in Computer Science, The University of Hong Kong. 2017-Current.
  Supervisor: Dr. Heming Cui.
  Research interests: fault-tolerant systems, distributed databases, secure and privacy-preserving systems, machine learning systems, and blockchains.
- Bachelor of Engineering in Computer Science, The University of Hong Kong. 2012-2017.
  Major: Computer Science, Minor: Mathematics. GPA: 3.7
- Exchange student at University of California, San Diego, USA. Jan-Jun, 2015.
  Major: Computer Science and Engineering, GPA: 4.0

PUBLICATIONS (IN REVERSE-CHRONOLOGICAL ORDER)

- BIDL: A High-throughput, Low-latency Permissioned Blockchain Framework for Datacenter Networks.
  Xusheng Chen*, Ji Qi*, Jianyu Jiang, Tianxiang Shen, Shixiong Zhao, Sen Wang, Gong Zhang, Man Ho Au, Heming Cui.
  *** Parallel First Author, equal contribution.**
  *Proceedings of the 28th ACM Symposium on Operating Systems Principles 2021 (***SOSP '21***, acceptance rate: 54/348 = 15.5%).*
  Describes BIDL, a permissioned blockchain framework that leveraging a new shperded sequencer abstraction to speculatively parallelize the consensus phase and execution phase of a blockchain system. BIDL features a new denylist protocol to achieve high throughput and low latecny even in the presence of malicious participants.

- Eges: Efficient, DoS-resistant Consensus for Permissioned Blockchains.
  Xusheng Chen, Shixiong Zhao, Ji Qi, Jianyu Jiang, Haoze Song, Cheng Wang, Tsz On Li, T.-H. Hubert Chan, Fengwei Zhang, Xiapu Luo, Sen Wang, Gong Zhang, Heming Cui.
  *The 39th IFIP WG 7.3 International Symposium on Computer Performance, Modeling, Measurements and Evaluation 2021 (***Performance 2021***, acceptance rate: 27/81 = 32.9%).*
  Describes Eges, an SGX-powered distriburted consensus protocol for large-scale blockchains. Eges's new stealth consensus committee abstraction can rotate and reach consensus securely, stealthly and efficiently. Among consensus protocols for permissioned blockchains, Eges is the first protocol that can ensure both safety and good performance against tough DOS attacks, including targetting attacks on consensus nodes and network partition attacks

- vPIPE: A Virtualized Acceleration System for Achieving Efficient and Scalable Pipeline Parallel DNN Training.
  Shixiong Zhao, Fanxin Li, **Xusheng Chen**, Xiuxian Guan, Jianyu Jiang, Dong Huang, Yuhao Qing, Sen Wang, Peng Wang, Gong Zhang, Cheng Li, Ping Luo, Heming Cui.
  IEEE Transactions on Parallel and Distributed Systems 2021 **(TPDS '21)**.
  Describes vPipe, a general pipepline parallel training system for big AI models (e.g., GPT and GNMT). vPipe tackles a severe load imbalance problem among the multi-GPU pipeline stages caused by inherent tensors accumulation in former stages, which can greatly downgrade training performance. vPipe's throughput and effective GPU ALU utilization are higher than existing premier pipepline parallel training systems (e.g., GPipe and Pipedream). vPipe is also the first pipeline parallel traing system that supports NAS training.

- DAST: Achieving Low Tail-latency and High Scalability for Serializable Transactions in Edge Computing
  Xusheng Chen, Haoze Song, Jianyu Jiang, Chaoyi Ruan, Cheng Li, Seng Wang, Gong Zhang, Reynold Cheng, Heming Cui.
  *Proceedings of the European Conference on Computer Systems 2021 (***EuroSys '21***, acceptance rate: 38/181 = 20.9%)*
  Describes DAST, a fault-tolerant database deployed across globally distributed edge computing servers. DAST is the first geo-distributed database that guarantees serializability, low tail-latency, and horizontal scalability.

- HAMS: High Availability for Distributed Machine Learning Service Graphs.
  Xusheng Chen*, Shixiong Zhao*, Cheng Wang, Fanxin Li, Ji Qi, Heming Cui, Cheng Li, Sen Wang.
  ***Parallel First Author, equal contribution.**
  *Proceedings of the 50th IEEE/IFIP International Conference on Dependable Systems and Networks (***DSN '20***, acceptance rate: 48/285 = 16.8%).*

Describes HAMS, an efficient fault-tolerant system for a distributed deployment graph consisting of machine learning serving models (some are stateful) with little performance overhead. HAMS is the first system that can correctly replicate the non-determinism caused by GPU threads on running AI serving models.

- DAENet: Making Strong Anonymity Scale in a Fully Decentralized Network.
  Tianxiang Shen, Jianyu Jiang, Yunpeng Jiang, **Xusheng Chen**, Ji Qi, Shixiong Zhao, Fengwei Zhang, Xiapu Luo, Heming Cui
  *IEEE Transactions on Dependable and Secure Computing 2020 (***TDSC '20***).*
  Describes DAENet, an SGX-powered secure communication system that can provide strong anonymity (e.g., hiding communication endpoints and traffic routes) and ensure reasonable communication efficiency even on packet dropping and DOS attacking scenarios. DAENet's strong anonymity and efficiency are scalable on Internet.

- UPA: An Automated, Accurate and Efficient Differentially Private Big-data Mining System.
  Tzs On Li, Jianyu Jiang, Ji Qi, Chi Chiu So, Jiacheng Ma, **Xusheng Chen**, Tianxiang Shen, Heming Cui, Yuexuan Wang, Peng Wang.
  *Proceedings of the 50th IEEE/IFIP International Conference on Dependable Systems and Networks (***DSN '20***, acceptance rate: 48/285 = 16.8%).*
  Describes UPA, the first automated, efficient, and precise differentially private big-data mining system that can preserve individual privacy while supporting general Spark big-data queries.

- Uranus: Simple, Efficient SGX Programming and Its Applications.
  Jianyu Jiang, **Xusheng Chen**, Tzs On Li, Cheng Wang, Tianxiang Shen, Shixiong Zhao, Heming Cui, Cho-Li Wang, Fengwei Zhang.
  *Proceedings of the 15th ACM ASIA Conference on Computer and Communications Security (***ASIACCS '20***, acceptance rate: 67/308 = 21.8%).*
  Describes Uranus, the first SGX-compatible and Spark-compatible secure big-data computing system that can practically support typical big-data datasets with our new big-data aware Java memory reuse protocols.

- Fulva: Efficient Live Migration for In-memory Key-Value Stores with Zero Downtime.
  Jiewen Hai, Cheng Wang, **Xusheng Chen**, Tsz On Li, Heming Cui.
  *Proceedings of the 38th International Symposium on Reliable Distributed Systems (***SRDS '19***, acceptance rate: 31/117 = 26.5%).*
  Describes Fulva, an in-memory key-value store migration system that can almost pertain the store's normal-case performance with no downtime during migration.

- PLOVER: Fast, Multi-core Scalable Virtual Machine Fault-tolerance.
  Xusheng Chen*, Cheng Wang*, Weiwei Jia, Haoran Qiu, Boxuan Li, Shixiong Zhao, Heming Cui.
  ***Parallel First Author, equal contribution.**
  *Proceedings of The 15th USENIX Symposium on Networked Systems Design and Implementation 2018 (***NSDI '18***, acceptance rate: 40/159 = 15.4%).*
  Describes PLOVER, a multi-core scalable VM fault-tolerance system built on the Virtualized State Machine Replication (VSMR) concept.

- A Fast, General Storage Replication Protocol for Active-Active Virtual Machine Fault Tolerance.
  Cheng Wang, **Xusheng Chen**, Zixu Wang, Youwei Zhu, Heming Cui.
  *Proceedings of the IEEE 23rd International Conference on Parallel and Distributed Systems (***ICPADS '18***).*
  Describes Gannet, a replication protocol for efficiently replicating virtual machines.

- Effectively Mitigating I/O Inactivity in vCPU Scheduling.
  Weiwei Jia, Cheng Wang, **Xusheng Chen**, Jianchen Shan, Xiaowei Shang, Heming Cui, Xiaoning Ding, Luwei Cheng, F.C.M. Lau, Yuexuan Wang, Yuangang Wang.
  *Proceedings of the 2018 USENIX Annual Technical Conference (***ATC '18***, acceptance rate: 76/378 = 20.1%*)
  Describes vMigrator, a lightweight, easy to use tool that can effectively mitigate I/O inactivity in vCPU scheduling, greatly improving I/O performance in VMs.

- APUS: Fast and Scalable PAXOS on RDMA.
  Cheng Wang, Jianyu Jiang, **Xusheng Chen**, Ning Yi, Heming Cui.
  *Proceedings of the ACM Symposium on Cloud Computing (***SoCC '17***, acceptance rate: 48/203 = 23.6%.*)
  Describes APUS, a fast Paxos protocol and its runtime system using fast RDMA features.

- SLARM: SLA-aware, Reliable and Efficient Transaction Dissemination for Permissioned Blockchains.
  Ji Qi, Tianxiang Shen, Yunpeng Jiang, **Xusheng Chen**, Jianyu Jiang, Xiapu Luo, Fengwei Zhang, Sen Wang, Heming Cui.
  Submitted to CoNext '21, under review.

Work Experience

- Research intern at Microsoft Research, Asia. Jun 2021 - Now.
- Research assistant at Center of Cloud Computing & Big Data, Lenovo, Hong Kong. Dec 2015 - Jul 2016.
- Student Trainee at Global Banking and Markets, HSBC, Hong Kong. Aug 2015 - Nov 2015.
- Summer Intern at Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Science, Beijing. June 2014 - Aug 2014.

Technology Transfer (Patents)

- A high-performance DNN training system with efficient and scalable pipelined parallelism on GPUs.
- A system in achieving low tail-latency and high scalability for serializable transactions in edge computing.
- A High-throughput, Low-latency Permissioned Blockchain Framework for Datacenter Networks.
- An Automated, Accurate and Efficient Differentially Private Big-data Mining System.
- A decentralized, secure and reliable network communication system via SGX.
- An Efficient, Secure Big-data Processing and Programming System based on Trusted Execution Environment.
- An Efficient, DoS Resistant Consensus Protocol for Permissioned Blockchains.
- A Distributed Fault-Tolerance Storage System via Virtualized State Machine Replication.

Others

- AE Committee member for **SOSP '21**.
- Reviewer for IEEE Transactions on Dependable and Secure Computing (**TDSC' 20**).
- Lee Shaw Kee Scholarships, 2014.
- Three times Deans Honours List, 2014-2015, 2015-2016, 2016-2017.
- Two times project co-leader of Huawei Innovation Research Programs (HIRP), 2017-2018, 2018-2020.
- Two times champion in the university basketball league.