Xusheng CHEN

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

Phone: (+852) 92390400
(+86) 14715487800
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 distributed 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

PROJECTS

- **Fault-tolerant large-scale distributed databases for serializable transactions in edge computing [2020-2021].**
  - *Summary:* we built DAST, the first fault-tolerant, large-scale distributed database that can ensure both serializability and low tail-latency for emerging edge computing applications. To achieve this goal, we invent a new distributed concurrency control protocol that addresses the convoy effect in existing geo-distributed database systems.
  - *My tasks:*
    1. Study existing distributed databases, summarize the key reasons why they fail to achieve low tail-latency, and derive the new concurrency control protocol.
    2. Implement the our new protocol and state-of-the-art protocols in an open-source database. Our protocol reduce the tail latency for up to 93.2%.
  - *Deliverables:* one paper accepted by EuroSys '21, one patent transferred to Huawei, code will be open-sourced soon (after patent submitted).
- **Training/inference systems for large machine learning models [2019-now].**
  - *Summary:* we built a system that can train large ML models (e.g., BERT, AmoebaNet) efficiently with pipeline-parallelism manner, and can provide high-availability for ML service graphs consisting stateful models (e.g., LSTM).
  - *My tasks:*
    1. Design and implement the key protocol for non-stop layer migration in pipeline-parallel model training.
    2. Design and implement the key protocol for non-stop replication of stateful ML models.
  - *Deliverables:* one paper accepted by DSN '20, one paper submitted to TPDS '21 (minor revision), one patent transferred to Huawei, code released on `https://github.com/hku-systems`.
- **Efficient blockchain systems for Internet and datacenter networks [2017-now].**
  - *Summary:* we built a new efficient, DoS-resistant, and scalable consensus protocol leveraging Intel SGX for large-scale blockchains running on the Internet, and a highly-efficient execution framework for blockchains systems deployed among many data centers.
  - *My tasks:*
    1. Design the new consensus protocol and the workflow of new execution framework.
    2. Implement the new consensus protocol on Ethereum codebase.
  - *Deliverables:* one paper submitted to TPDS (under review), one paper submitted to OSDI (under review), two patents transferred to Huawei, code released on `https://github.com/hku-systems`.
- **Efficient and transparent fault-tolerant systems leveraging data center networks [2017-2019].**
  - *Summary:* We leverage the network features (e.g., RDMA and DPDK) in datacenter networks to build highly-efficient fault-tolerant systems for server applications.
  - *My tasks:*
    1. Design and optimize the consensus protocols used in the fault-tolerant systems.

2. Implement the consensus protocol on QEMU and KVM.

– *Deliverables:* one paper accepted by SoCC '17, one paper accepted by NSDI '18, one paper accepted by ICPADS '18, one paper accepted by SRDS '19, one patent transferred to Huawei, code released on `https://github.com/hku-systems`.

- **Efficient and secure big data computation systems using TEEs [2017-2019].**
  – *Summary:* We built a secure JVM leveraging TEEs (e.g., Intel SGX) to provide computational integrity and data privacy for big data computation frameworks (e.g., Spark) and ML model training.
  – *My tasks:*
    1. Design and implement the protocol for loading, verification, and attestation of of trusted code into TEE enclaves.
  – *Deliverables:* one paper accepted by ASIACCS '19, one paper accepted by TDSC '20, two patents transferred to Huawei, code released on `https://github.com/hku-systems`.

PUBLICATIONS

- **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***, accepted)*
  Describes DAST, a fault-tolerant database depployed across gloablly distributed edge computing servers. DAST is the first geo-distributed database that guarantees serailizaibility, 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.
  *Proceedings of the 50th IEEE/IFIP International Conference on Dependable Systems and Networks (***DSN '20***).*
  *** Parallel First Author (annotated in the paper's first page).**
  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 nondeterminism caused by GPU threads on running AI serving models.

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

- **Eges: Efficient, DoS-resistant Consensus for Permissioned Blockchains.**
  Xusheng Chen, Shixiong Zhao, Cheng Wang, Haoze Song, Jianyu Jiang, Ji Qi, Tsz On Li, T.-H. Hubert Chan, Heming Cui.
  Submitted to IEEE TPDS, under review.
  Describes Edges, an efficient consensus protocol that can hide consensus nodes to defend against targeted DoS attacks in a large-scale permissioned blockchain system.

- **BIDL: A High-throughput, Low-latency Permissioned Blockchain Framework for Datacenter Networks.**
  Xusheng Chen, Ji Qi, Yunpeng Jiang, Jianyu Jiang, Tianxiang Shen, Shixiong Zhao, Sen Wang, Gong Zhang, Man Ho Au, Heming Cui.
  *** Parallel First Author (annotated in the paper's first page).**
  Submitted to OSDI '21, under review.

- **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***).*
  Describe 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 anonimity 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***).*
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***).*
  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***).*
  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.

- **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***), 2017.*
  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***), 2018.*
  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***), 2017.*
  Describes APUS, a fast Paxos protocol and its runtime system using fast RDMA features.

- **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.
  Submitted to IEEE TPDS, under minor revision.

- **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 SIGMETRICS '21, under review.

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.

Work Experience

- Research Assistant at Center of Cloud Computing & Big Data, Lenovo, Hong Kong. Dec 2015 - Jul 2016.
- Student Trianee 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.

Others

- 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.