

# Weidong Zhu

Master, Xiamen University

## Personal Information

---

**Address:** No. 422 Siming South Road, 361005, Xiamen, China  
**My Github:** [www.github.com/zwdong1994](http://www.github.com/zwdong1994)

**Phone:** +86-18959218665  
**Email:** [zwdong1994@gmail.com](mailto:zwdong1994@gmail.com)

## Education

---

2016 - now	M.S, Computer Science, School of Information and Engineering <b>Adviser:</b> Prof. Suzhen Wu and Prof. Bo Mao <b>Xiamen University, China</b>
2012 - 2016	B.S, Information Security, School of Computer Science and Technology <b>Huazhong University of Science and Technology, China</b>

## Research Interests

---

**Areas:** Storage System, Database, Parallel/Distributed Systems  
**Focus:** Deduplication, SSD, K-V Stores, SSD RAID

## Conference Publications

---

- Suzhen Wu, Weidong Zhu, Guixin Liu, Hong Jiang, and Bo Mao. **GC-aware Request Steering with Improved Performance and Reliability for SSD-based RAIDs**. To appear in Proceedings of the 32nd IEEE International Parallel & Distributed Processing Symposium (IPDPS 18). Vancouver, British Columbia, Canada, May 21-May 25, 2018.

## Journal Publications

---

- Suzhen Wu, Weidong Zhu, Bo Mao, Kuan-Ching Li. **PP: Popularity-based Proactive Data Recovery for HDFS RAID systems**. Future Generation Computer Systems (2017), <http://dx.doi.org/10.1016/j.future.2017.03.032>.

## Research & Experience

---

**Research Assistant**  
**Advanced Storage Technology Lab, Xiamen University**

**Mentor: Suzhen Wu    July 2016 - now**

- KV-based Stores**  
LSM tree is one of the most popular data structures in K-V stores implementation, because it transfers the random writes into sequential writes. However, there are still have some problem in the LSM-based K-V store, just like write/read amplification. So, I am studying the structure and characteristics of Leveldb(A LSM-based KV database.) and OpenChannel-SSD to find a way to decrease the impact that write/read amplifications brought about.
- Deduplication**  
Deduplication can significantly reduce write traffic and it has great profit for flash-based storage system, like SSD RAID. So, deduplication can help to improve the performance and reliability in the flash storage system. Moreover, with the advent of NVMe and 3D NAND flash technologies, the performance of flash-based storage systems has been improved significantly. So, this brought a new problem that the hash operation might become the bottleneck in the deduplication systems. So, I proceeded my research with this direction.

- **SSD-based Array**

SSD-based RAID's suffer from significant performance degradation whenever user I/O requests conflict with the ongoing Garbage Collection (GC) operations which introduces tail latency. And it will also brought great tail latency in the SSD-based Array. Moreover, the uncorrectable error occurred of SSD in the SSD-based RAID's will trigger the fail-recovery process. So, proposing a scheme that aware of the GC process within an SSD-based RAID, to address both the performance and reliability issues of SSD-based RAID's alluded to above is our main target in this research project. (IPDPS'18)

**Join a Competition**

**2014 - 2015**

**Information Security Lab, Huazhong University of Science and Technology**

- **A USB-based Device with Encryption**

To ensure the security of USB-based device, encryption or the other security technology used on the USB-based device. However, existing schemes were always built in the software layer and its access control function is not perfect. So, we built a security system with USB-based device by using AES algorithm to encrypt the data and provided a strong access control system on the PC. Besides, the encryption system built on the firmware of the USB-based device.

## Honors & Awards

---

Excellent Merit Student, Xiamen University, China

Nov. 2017

National Scholarship, Ministry of Education of China

Oct. 2017

Third Price of 2015 National College Students Information Security Contest, Ministry of Education of China Aug. 2015

## Computer Skills

---

**Operating System:**

Linux, Windows

**Database:**

Leveldb

**Programming:**

C, C++, Python, GO, L<sup>A</sup>T<sub>E</sub>X