

# Shengzhe Wang

☎ 18860016206   ✉ sshanewang@163.com   📍 Xiamen  
📄 Intended Position: Software Development Engineer



## EDUCATION

<b>Xiamen University</b>	Sep 2020 - Jun 2023
Computer Science and Technology Master	GPA: 3.5/4.0
<b>Xiamen University</b>	Aug 2016 - Jun 2020
Software Engineering Bachelor	GPA: 3.5/4.0

## PROFESSIONAL EXPERIENCE

<b>Alibaba Cloud Computing</b>	Jul 2022 - Present
Research Intern	Beijing

**Research Topic:** Software and Hardware Co-design Based on DPU in Cloud Storage

**Research Purpose:** Developing DPU-based AI technology to accelerate key-value storage system (LevelDB)

**Research Progress:** (1) Exploring the opportunity of offloading the compaction process of LevelDB/RocksDB to DPU. (2) Trying to Deploy a learning-based auto-tuner for LevelDB/RocksDB in DPU.

## RESEARCH EXPERIENCE

<b>Learning-based Flash Translation Layer</b>	Jan 2022 - May 2022
-----------------------------------------------	---------------------

- **Motivation:** To reduce the double-read (one flash read for address, another flash read for data) problem caused by insufficient cache memory and high-space-overhead mapping table structure in SSD FTL.
- **Design:** (1) Replacing the mapping table structure with highly space-efficient learned indexes in the cache memory. Making sure the learned indexes can index all flash pages. (2) Retraining each learned index during garbage collections.
- **Contribution:** (1) Implemented the designs in a C-based FEMU simulator. (2) Completed the paper writing.
- **Evaluation:** Our design outperforms other competitors by 20%-50% in SSD reads.  
Research Progress: In the submission of EuroSys .

<b>Key-value storage based on SSD and Optane Memory</b>	Jul 2021 - May 2022
---------------------------------------------------------	---------------------

- **Motivation:** (1) To solve the write stall caused by the disorder of the L0 layer in LevelDB. (2) To improve LevelDB's performance under workloads with spatial locality. (3) The Optane memory (PM) can be used as a cache of DRAM.
- **Research 1:** LevelDB optimization. (1) Deploying a PM layer between DRAM and SSD. (2) Removing unsorted L1, and design a PM-based B+-Tree in PM. Key-value pairs flushed by memtables are first inserted into the PM-based B+-Tree. After PM-based B+-Tree reaches the threshold, selecting key-value pairs in cold leaves and flush them to SSD L1 layer.
- **Research 2 :** PM-based B+-Tree Optimization: (1) Removing memtables, and put key-value pairs in DRAM-based internal leaves' unused space. After each internal leaf is full, we flush the buffered key-value pairs to PM-based leaf nodes. (2) Designing a selective caching strategy to improve the flush efficiency of the internal node cache.
- **Contribution:** (1) Completed part code and paper writing of research 1. (2) Writing the code of research 2.
- **Evaluation:** For research 1, our design's overall throughput outperforms other competitors by about 70%, and the tail latency decreases by about 16 times.
- **Research Progress:** Research 1 is in the submission of EuroSys. Research 2 is in progress.

## SKILLS, CERTIFICATIONS & OTHERS

- **Skills:** C/C++; Persistent Memory; SSD; Linux
- **Certifications:** 2019 WeChat Mini-Program Contest National Third Prize; Xiamen University Academic Scholarship

# 王圣哲

☎ 18860016206    ✉ sshanewang@163.com    📍 厦门  
👤 应届生    📁 服务端研发工程师



## 教育经历

厦门大学	2020年09月 - 2023年06月
计算机科学与技术 硕士 信息学院	导师：吴素贞副教授 GPA: 3.5/4.0
厦门大学	2016年08月 - 2020年06月
软件工程 本科 软件学院	GPA: 3.5/4.0

## 实习经历

阿里巴巴集团-阿里云智能	2022年07月 - 至今
学术合作实习生 盘古	北京
研究课题：云存储中基于DPU的软硬件协同设计	
<ul style="list-style-type: none"><li>● <b>研究目的</b>：使用基于DPU的AI技术加速键值存储系统（LevelDB）</li><li>● <b>设计思路</b>：① 利用DPU的近数据计算和独立系统特性，将LevelDB中的压缩过程卸载到DPU中进行，并设计专用算法智能对主机数据进行压缩。② 在DPU中实现利用机器学习算法对主机上的LevelDB/RocksDB进行动态参数调整。</li><li>● <b>目前进展</b>：① 调研DPU的设备特性与访问方式。② 探索如何使用DPU中硬件加速器对LevelDB的压缩过程进行加速。③ 将基于强化学习和机器学习的LevelDB/RocksDB动态调参代码部署于DPU中进行测试。</li></ul>	

## 研究经历

基于学习索引算法的固态硬盘地址翻译加速技术研究	2022年01月 - 2022年05月
<ul style="list-style-type: none"><li>● <b>动机</b>：SSD内部的缓存（DRAM）用于存放地址映射表（通常为页级别）。随着SSD容量的不断增大，将全部映射表放入缓存会带来较大的内存开销。主流的做法是只缓存热的地址映射于缓存中，冷的地址映射存放于flash中，但这种做法在随机读场景下会带来额外的flash读，增大读取延迟。</li><li>● <b>设计</b>：提出LearnedFTL，在缓存热地址映射的基础上，使用学习索引算法为flash中的冷地址映射建立超低内存占用的机器学习模型，并放入SSD的缓存中，通过缓存中的模型命中，减少为了地址翻译的额外的flash读，从而加速SSD中的随机读取效率。</li><li>● <b>主要贡献</b>：代码方面，在基于C的FEMU SSD模拟器中修改FTL层，完成了模型构建和加速。论文方面，完成论文的撰写。</li><li>● <b>指标提升</b>：相较于竞争对手，LearnedFTL在顺序读和随机读测试中提升20%-50%。</li><li>● <b>研究进展</b>：EuroSys 2023在投。</li></ul>	
基于持久性内存和高性能固态硬盘的键值存储系统研究	2021年07月 - 2022年05月
<ul style="list-style-type: none"><li>● <b>动机</b>：① LevelDB将L0数据无序化和粗粒度压缩的方式带来写停顿和读放大。② 与flash相比，容量大，速度快，可字节寻址的Optane持久性内存(PM)可以替代SSD实现L0细粒度数据管理。③ B+树比SSTable更适合实现细粒度管理。④ 与DRAM相比，PM具有低扩展性和块级别写入粒度，针对PM的B+树需要进行写入优化。⑤ 负载具有时间和空间局部性。</li><li>● <b>研究1</b>：使用B+树替换L0层的SSTable: ① 使用部署于PM上的B+树替换部署于SSD中的L0层的SSTable，实现细粒度的管理和压缩。② 设计基于叶子结点热度的节点刷新的算法，保留热节点在PM中。</li><li>● <b>研究2</b>：PM-B+树优化：① 将PM-B+树的内部节点放于DRAM中作数据写入缓存，取消memtable，数据直接写入内部节点的空闲空间，待内部节点空间满后刷新入PM上的叶子结点。② 应用选择性缓存策略，提高内部节点缓存的刷新效率。</li><li>● <b>主要贡献</b>：完成研究1部分代码与论文的写作，正在完成研究2。</li><li>● <b>指标提升</b>：研究1的整体系统吞吐量提升约70%；尾延迟降低约16倍。</li><li>● <b>研究进展</b>：研究1已完成，EuroSys 2023在投。</li></ul>	

## 技能/证书及其他

- **技能**：掌握C/C++，熟悉C++新特性，熟悉持久性内存与SSD，Linux操作系统，具有良好的英文读写能力。
- **证书/执照**：2019微信小程序大赛国家三等奖；厦门大学学业奖学金（三次）