

HEERAK LIM / 임희락

1992.12.29 / 1 Gwanak-ro, Gwanak-gu, Seoul 08826

+82 10-4814-3800 / rockylim@snu.ac.kr

Summary

Expected to master's degree graduate in August 2019 with two years of experience working on database systems and storage systems including Key-Value store, RocksDB, Open-Channel SSDs and parallel computing. Looking for research and engineering position in system, storage and database.

Academic Experience

External Merge Sort based on flash base storage system

Distributed Computing System Laboratory (DCSLAB), Seoul National University (May 2019 –)

- Design and Implement External Merge Sort considering internal parallelism of flash-based storage.
- Managed codebases using GitHub - <https://github.com/RockyLim92/ocm-sort>

데이터베이스 workload 기반 NVMe SSD 의 신규 command set 연구 (SAMSUNG 산학 연구 과제)

Distributed Computing System Laboratory (DCSLAB), Seoul National University (Jun 2018 – Dec 2018)

- Profiled characteristics of database workload in Linux kernel module level
- Analyzed MySQL(InnoDB) I/O process and function call stack by SQL query type (inset, select, delete, etc.)
- Designed and implemented I/O benchmark tool for Open-Channel SSDs using user-space I/O library
- Managed codebases using GitHub - https://github.com/RockyLim92/ocssd_bench

LSM-tree based Database System Optimization using Application-Driven Flash Management

Distributed Computing System Laboratory (DCSLAB), Seoul National University (Sep 2018 – May 2019)

- Optimized an LSM-tree based Key-Value store system by processing I/O taking into account of application context
- Modified RocksDB storage backend to manage Flash based storage device directly using user-space I/O library
- Deployed RocksDB on High Performance Computing (HPC) System
- Open-source contribution (bug fix) and managed codebases using GitHub - <https://github.com/OpenChannelSSD/rocksdb/pull/3>
- Learned how to quantitatively evaluate and analyze system performance
- Documented the results of the project through technical documentation such as thesis and report
- Contributed to international conferences and presented academic conferences
- Managed codebases using GitHub - <https://rockylim92.github.io/research/dapdb/>

Performance Modeling and Measurement of Selective Page-Mapping Table on the OpenSSD Platform

Ajou University (Apr 2016 – June 2016)

- Developed Demand-based Flash Translation Layer (DFTL) on OpenSSD platform
- Evaluated DFTL on OpenSSD with macro benchmark tool
- Established a mathematical model to predict the performance of the DFTL
- Established a performance model through linear regression and multiple linear regression methods using R programming environment

Work Experience

Programming Instructor

CIT Code Academy, Korea (Jul 2017 – Jan 2018)

- Computer Architecture Class
- C, C++, Python Class

Research Intern

Commonwealth Scientific and Industrial Research Organisation (CSIRO) Australia. Data61, Distributed Sensing System Team. (Jan 2017 – Feb 2017),

- Development of Radio Tomographic Imaging (RTI) system on low-power sensor networks
- Designed and implemented multiple sensor network node synchronization algorithm
- Designed and Implemented RTI application on Contiki OS
- Modified network device driver as low-overhead manner to achieve superlative frequency
- Developed Image reconstruction utility using python
- Managed codebases using GitHub - <https://github.com/RockyLim92/radio-tomography>
- <https://rockylim92.github.io/research/csiro06/>

Research Intern

차세대융합기술연구원 [Advanced Institute of Convergence Technology; AICT], 의료-IT 연구 센터 [Medical-IT research center] (Jul 2016 – Aug 2016)

- Developed web page crawler using Python to build database system for digital pathology.
- Developed C# scripts for 3D model that simulates a skeletal movement

Education

Master of Science in Computer Science

Seoul National University, Distributed Computing System Laboratory (DCSLAB) (2019),

Graduated with a bachelor's degree in Department of Software Engineering

Ajou University (2017)

Publications

International conference

- **Lim, Heerak.** "Application-Driven Flash Management: LSM-tree based Database Optimization through Read/Write Isolation." Proceedings of the Doctoral Symposium of the 19th International Middleware Conference. ACM, 2018.

Domestic conference

- 임희락, 염현영, and 손용석. "플래시 내 I/O 분리 처리를 통한 LSM-tree 기반 데이터베이스 성능 최적화." *한국정보과학회 학술발표논문집*(2018): 19-21.
- 임희락, et al. "Deep Neural Network 기반의 비언어 인식을 통한 청각장애인 위험 인지 보조 시스템." *한국지능정보시스템학회 학술대회논문집*(2017): 121-122.
- 임희락. "Open Channel SSD 플랫폼에서 쓰기 버퍼 및 스레드 구성에 따른 성능 분석." *한국정보과학회 학술발표논문집*(2017): 73-75.
- 강동현, 하영남, 임희락, & 김영재. (2016). OpenSSD 플랫폼에서 선택적 캐싱 기반 페이지 매핑 테이블의 캐시 비율에 따른 성능 모델 연구. *한국정보과학회 학술발표논문집*, 1245-1247.

Technical Document

- The Jasmine OpenSSD Platform: Technical Reference Manual (v1.4, in English) (By Preethika Kasu, Donghyun Kang, Heerak Lim), http://www.openssd-project.org/mediawiki/images/Jasmine_Tech_Ref_Manual_v.1.4e.pdf
- The Jasmine OpenSSD Platform: FTL Developer's Guide (v1.2, in English) (By Preethika Kasu, Donghyun Kang, Heerak Lim). http://www.openssd-project.org/mediawiki/images/Jasmine_FTL_Dev_Guide_v.1.2e.pdf

Awards and Honours

- Awarded Excellence Paper Award at the 45th Korea Software Congress (Dec 2018)
- Scholarship of 21 items including Honor scholarship S (the greatest) for all semesters at Ajou
- Ajou Greative (Great + Creative) software concert, Excellence Award (Jun 2017)
- Ajou Greative (Great + Creative) software concert, Excellence Award (Aug 2016)

Technical Skills

- Extensive experience in C, C++ Programing language
- Deep understanding in Linux operating system and storage system
- Deep understanding in LSM-tree data structure and it's application such as RocksDB, LevelDB
- Extensive experience in Open-Channel SSDs, Application-Driven Flash Management, FTL
- Extensive experience in development under Linux/Unix environment
- Worked on system performance profiling and monitoring
- Practical experience in large-scale source code analysis
- Practical experience in system performance tuning
- Experienced in data analysis using Python (numpy, pandas, matplotlib, etc.)

Areas of interest

- Database, Storage system, Parallel Computing
- Key-Value store, NoSQL
- Linux/Unix environment

- System performance improvement
- Application-Driven Storage Management
- Troubleshooting considering multi-thread concurrency
- Automation in the development process

Language Skills

Abounding experience in technical writing. Intermediate English speaker. Achieved Test Of English for International Communication (TOEIC) score 850.

Others

- Research Blog - <https://rockylim92.github.io/>