Yuanjiang Ni

CONTACT INFORMATION	556 Redwood Rd. Felton, CA 95018	831-295-3975 yni6@ucsc.edu

RESEARCH INTERESTS

Non-Volatile Memory, File/Storage Systems, Operating Systems, Distributed Systemss

EDUCATION

UC Santa Cruz, Santa Cruz, CA   Sep. 2016 - now

Ph.D., Computer Science
- Advisor: Ethan L. Miller

Sichuan University, Chengdu, Sichuan   Sep. 2011 - Jun. 2015

B.S., Software Engineering
- GPA: Overall 85.2/100, Major 89.5/100

RESEARCH EXPERIENCE

Research Assistant   Mar. 2017 - now
Storage System Research Center   **Advisor:** Ethan L. Miller

Optimized Shadow Page for Storage Consistency on BNVM

- When logging is used, old or new data needs to be copied to somewhere else. Additional cacheline flushing caused by persisting the copy is proven to be a performance killer. Our work is driven by the observation that shadow page is able to eliminate the data copying. We use sub-page validity to reduce the overhead brought by the mismatch between the page granularity and the actual byte ranges that are modified.
- Created a performance model; Investigated an application-level implementation that requires no hardware/OS change; Prototyped a hardware-assist implementation in the McSimA+ simulator.

PM-aware Process Checkpointing

- In this work, we investigated using process checkpointing techniques that have been widely used for process migration and high performance computing. Checkpointing would be promising for PM-based system since it can address the issue of crash consistency in an application-transparent way. We proposed PM-aware checkpointing, which leverages the Kernel's CoW semantic to facilitate checkpointing in a PM-attached system.
- Studied the Linux process subsystem; Implemented the PM-aware checkpointing prototype in FreeBSD kernel; Identified several problems of the preliminary design (e.g kernel crossing overhead, CoW overhead, etc); Proposed possible architecture changes that can make process checkpointing more applicable in a PM-attached system.

Research Internship & Software Engineer   Aug. 2014 - May 2016
Storage System Group, ICT, CAS   **Advisors:** Jin Xiong and Dejun Jiang

S-RAC: SSD Friendly Caching for Data Center Workloads (SYSTOR' 16)

- SSD caching faces two challenges: i) SSD has limited write endurance, which requires to reduce write amount to SSD, and ii) data-center workloads exhibit a diverse I/O access patterns, which requires to figure out SSD caching friendly patterns. We propose an SSD cache manager S-RAC with re-adding blocks and ghost cache adaptation to retain SSD friendly blocks in SSD.

- Proposed a novel SSD caching policy to deal with diverse data center workloads; Implemented S-RAC as a block-layer cache in Linux; Carried out extensive evaluation to show the efficiency of my caching design; Participated in writing the conference paper.

COSMO: Cache Oriented SSD Management and Optimization

- This work is driven by the observation that a unified view of garbage collection and cache eviction in SSD caches is able to reduce overall Write Amplification, thus boosting applications performance and elongating SSD lifetime. COSMO integrated cache and flash management and explicitly controls the page placement on SSDs, clustering data in a way that facilitates future garbage collection.
- Studied the Linux block layer and device mapper; Investigated several existing SSD cache systems (dmcache, flashcache, bcache, etc); Implemented a prototype of our COSMO cache manager (based on dmcache) into Linux kernel.

PROJECTS	**Understanding the Latency of File systems on NVM** 2017

- Identified several potential overhead for different FS operations (open, read, write, etc.); Studied the code of two Linux file systems, Ext3 and PMFS; Instrumented these file systems to break down the access latency.

Synthetic Memory Access Workloads Generation & Simulation 2017

- This is a side project of NVM checkpointing in which we want to model the performance of our checkpointing.
- Implemented a memory accesses generator that can generate workloads with customizable spatial/temporal locality, read/write ratio, etc. Developed a framework which allow users to easily simulate the behavior of their system under different access patterns.

URL Source Detection, Tracking and Statistics Module 2014

- Developed a prototype to crawl, locate and track web information; Trained a Naive Bayesian Classifier to filter out the target.

CSDN Open Source Summer Camp 2014

- Seafile is an open source cloud storage system. Administrators use a fuse-based file system to manage a git-like visioning file repository.
- Implemented several missing functionalities for the fuse-based file system.

PUBLICATIONS

1. **Y. Ni**, J. Jiang, D. jiang, X. Ma, J. Xiong, "S-RAC: SSD Friendly Caching for Data Center Workloads", **SYSTOR**, 2016

TEACHING
EXPERIENCE

- **CMPS111 Introduction To Operating Systems**, UCSC, TA   Winter 2017

AWARDS

- Outstanding Undergraduate Thesis (**Top 1%**)   2015
- Second Prize Scholarship (**Top 5%**)   2014
- Second Prize, China International Software Design Competition (**Top 5%**)   2014
- First Prize, Waterloo Cup   2014
- Outstanding Team, ICS&S Internship Project (**2 out of 26**)   2014
- Individual Scholarship   2012,2013

SKILLS

Kernel: Coding experience with Linux block I/O layer, device mapper, process subsystem
Programming: C/C++, Shell Scripting, Python, Java
Tools: Git, Dtrace, LaTeX, Pin, PMDK, Fio, Filebench