Woong Shin, Ph.D.

Research Staff, Analytics & AI Methods at Scale (AAIMS) Group
Advanced Technology Section (ATS), National Center for Computational Sciences (NCCS)
Oak Ridge National Laboratory, PO Box 2008, MS6012, Oak Ridge, TN 37831-6012
Phone: +1 865-440-6284, E-Mail: shinw@ornl.gov

Research Interest

- Data driven HPC data center operations and control, optimized systems for big data analytics, systems for near-real time stream data analytics, AI & machine learning systems and applications for intelligent facilities, data engineering, machine learning engineering, NVRAM based storage system architecture, distributed systems, large-scale databases, operating systems, computer architecture, embedded systems, software engineering

Education

- **Ph.D.**, Electrical Engineering and Computer Science, Seoul National University, Republic of Korea
  Master's & Ph.D. Integrated Program (September 2010 ~ February 2017)
  Dissertation Title: *"OS I/O Stack Optimizations for Flash Solid-State Drives"*
- **BS c**, Computer Science, Korea University, Republic of Korea, 2003

Employments

- **Research Staff,** October 2020 ~ present
  Analytics & AI Methods at Scale (AAIMS) Group, Advanced Technology Section (ATS)
  National Center for Computational Sciences (NCCS)
  Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
- **Research Staff,** April 2020 ~ September 2020
  Technology Integration Group, National Center for Computational Sciences
  Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
- **Research Associate**, May 2017 ~ March 2020
  Technology Integration Group, National Center for Computational Sciences
  Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
- **Postdoctoral Research Associate**, February 2017 ~ April 2017
  Technology Integration Group, National Center for Computational Sciences
  Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
- **Research Intern**, April 2016 ~ September 2016
  Technology Integration Group, National Center for Computational Sciences
  Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
- **Research Intern**, September 2012 ~ February 2013
  IBM Austin Research Laboratory, Austin, Texas, USA
- **Research Assistant**, September 2010 ~ February 2017
  Distributed Computing Systems Laboratory, Dept. of Computer Science and Engineering,
  Seoul National University, Republic of Korea
- **Research Engineer**, June 2006 ~ February 2010
  System Management Laboratory, Research & Development Center, TmaxSoft, Republic of Korea,
  (https://www.tmaxsoft.com)
- **Software Engineer**, January 2005 ~ June 2006
  Internet Telephone Business Team, Samsung Networks, Republic of Korea
- **Software Engineer**, March 2003 ~ December 2005
  Infrastructure Technology Team, Samsung Networks, Republic of Korea

Honors and Awards

- Research accomplishment award, UT-Battelle, LLC (UT-Battelle) Awards Night, November 2022
- Best paper award at the International Conference for High Performance Computing, Networking, Storage, and Analysis (SC'21), November 2021
- Best paper award at the Second International Conference on Big Data and Smart Computing (BigComp'15), February 2015
- Technology innovation award, TmaxSoft, January 2010
- Best employee award, TmaxSoft, August 2009

Research Projects

- **Framework and Tools for Frontier Supercomputer**
  Power & Energy Monitoring, Modeling and Optimization
  Oak Ridge National Laboratory Leadership Computing Facility (OLCF)
  Analytics & AI Methods at Scale (AAIMS) Group, Advanced Technology Section (ATS)
  National Center for Computational Sciences
  Oak Ridge National Laboratory
  Research Staff, July 2021 ~ present
    - Leading a team of data scientists & researchers as a technology lead on developing framework and tools for monitoring, analyzing, modeling and optimizing power & energy operations of the Frontier supercomputer. With a focus on supporting Frontier facility operations, the project aims to research and develop near real time data center visibility and data analytics capabilities that can cover data from Frontier and its supporting central energy plant (CEP). The project aims to deliver dashboards and interactive data analytics tools that use machine learning algorithms working on top of HPC job power profiles and facility data streams powered by visual analytics techniques.
- **AI for Facilities**
  Oak Ridge National Laboratory Leadership Computing Facility (OLCF)
  Analytics & AI Methods at Scale (AAIMS) Group, Advanced Technology Section (ATS),
  National Center for Computational Sciences
  Oak Ridge National Laboratory
  Research Staff, October 2020 ~ present
    - Research and development in machine learning and artificial intelligence powered operational data analytics use cases, systems, and techniques in the context of OLCF operations. Investigates various topics in HPC system data acquisition, automatic diagnosis of failures, prediction of load, and optimization aiming for AI-powered self-driving smart facility operations. Also, with a focus on "AI & ML in operations", putting efforts in developing & implementing services for sustainable data management, collaborative data analytics & machine learning model development that can support mission critical HPC user facility operations.
- **Long term Analysis of Summit's Power, Energy and Thermal Dynamics**
  Oak Ridge National Laboratory Leadership Computing Facility (OLCF)
  Analytics & AI Methods at Scale (AAIMS) Group, Advanced Technology Section (ATS),
  National Center for Computational Sciences
  Oak Ridge National Laboratory
  Research Staff, October 2020 ~ June 2021
    - Analyzed accumulated data from the "Cooling Intelligence for Summit" project to reveal power, energy, and thermal behavior and impact of the Summit system and its cooling & electrical infrastructure. Foundational work in understanding Summit's power & energy operational data and

its use for ML based methods. As the first author and lead investigator, delivered a research paper that provides a first-order examination and analysis of power consumption at the component-level, node-level, and system-level, from all 4,626 Summit compute nodes, each with over 100 metrics at 1Hz frequency over the entire year of 2020. The paper also investigates the power characteristics and energy efficiency of over 840k Summit jobs and 250k GPU failure logs for further operational insight.

- Paper accepted and won the ***best paper award*** at the International Conference for High Performance Computing, Networking, Storage, and Analysis (SC 2021) [1]. The team was recognized with the ***research accomplishment award*** at the UT-Battelle awards night in 2022.

- **Operational Data Analytics for Energy-Efficient HPC**
  Operational Data Analytics Team
  Energy Efficient HPC Working Group (EEHPCWG)
  Team Member, *November 2019 ~ present*
  - As a member of a team of experts from HPC sites and vendors such as LRZ, LLNL, ORNL, LBNL and HPE, investigated the trends, architectures, guidelines, roadmaps for operational data analytics (ODA) representing ORNL's ODA efforts. Participated in the analysis of a global survey on ODA co-authoring a global survey report as the second author [5]. Also participated in a team effort in proposing a conceptual framework for ODA that can help map the landscape of data analytics use-cases using operational data in the midst of plethora of data analytics techniques from domains such as data science, control theory and optimization. Co-authored a vision paper describing this conceptual framework as the second author [3].

- **Cooling Intelligence for the Summit Supercomputer**
  Oak Ridge National Laboratory Leadership Computing Facility (OLCF)
  Technology Integration Group, National Center for Computational Sciences
  Oak Ridge National Laboratory
  Research Associate, May 2018 ~ April 2019 (development), May 2019 ~ Current (operation & maintenance)
  - Delivered and maintained a near real-time monitoring and data analytics system that aggregated data from the facility and Summit which was required to push maximum energy efficiency. The system provides near real-time visualization of the current state of the Summit system and the cooling plant, adding confidence in adjusting high-impact cooling parameters. OLCF was able to address over-cooling issues imposed by conservative vendor cooling specifications, introducing significant cooling energy savings. After delivery of the system, undertook efforts in maintaining the quality of the monitoring system and the data streams / data sets throughout the lifetime of the target system Summit and its central energy plant (CEP) for facility engineers and researchers.

- **Multi-tiered Storage Systems for HPC centers**
  Technology Integration Group, National Center for Computational Sciences
  Oak Ridge National Laboratory
  Research Associate, April 2016 ~ March 2020
  - Evaluating NVM-e SSDs from multiple vendors (Intel, Samsung & HGST) to understand the impact on our systems when employed as a component (i.e., burst buffer) within an HPC multi-tiered storage environment helping OLCF to determine its SSDs for node local NVRAM.
  - Perform research in designing a system software called "Data Jockey" that aims to solve the difficulty of manually managing the namespace, placement and movement of various data across multiple storage tiers such as burst buffers, parallel file systems, object stores and long-term data backup systems. Results presented in IPDPS'19 [6] ***(Acceptance rate 27.7%)***

- **Enhanced OS & SSD Cross Layer I/O Optimization (Greybox SSDs)**
  Distributed Computing Laboratory, Seoul National University

Research Assistant, January 2015 ~ February 2017
- ○ Research on advanced cross-layer optimization techniques by introducing new SSD functionalities and extending the interface between the OS and the SSD. Employed multiple SSD evaluation platforms for research. Developed SSD FTL code enhancements as well as corresponding block device driver codes for the OpenSSD Cosmos SSD evaluation platform. Also built a DRAM backed FPGA based SSD latency emulation platform with the Xilinx VC709 evaluation board. Developed Linux block device driver code (v3.5.0), FTL enhancements running on SSD microprocessors (ARM Cortex A9, Xilinx Microblaze) as well as the FPGA logic (Xilinx Zynq-7000 & Xilinx Virtex7) for latency emulation, interrupt control and PCI-e communication. OpenSSD platform results presented in USENIX HotStorage'16 [7].

- **Hardware & Software Co-optimization
for High Performance / High Functionality SSDs in Data Centers**
Research subcontracted from Samsung Electronics, Seoul National University
Research Assistant, January 2014 ~ December 2014
  - ○ Research on cross-layer optimizations for key value storages targeting to reduce latency variations caused by mixed read and write operations on a flash SSD. Developed a multi-SSD storage engine which employs the capability of performing read-write separation as well as exploiting garbage collection API calls provided by the SSDs (Customized Samsung SSDs). Developed a custom storage engine integrated with Memcached 1.4.17 to demonstrate the effectiveness of tight cross-layer optimizations. Was able to reduce the 99.9999 percentile tail latency of foreground read operations to be under 1ms.
  Reported the results in ICPADS'14 [9] & BigComp'15 [8] ***(Best paper awards)***.

- **OS Optimizations for Multiple Flash SSDs**
Distributed Computing Laboratory, Seoul National University
Research Assistant, June 2013 ~ June 2015
  - ○ Research on optimizing the Linux 3.2.40 file I/O path to enhance the IOPS scalability when using multiple SSDs. Extended the OS block storage interface, enhanced the interactions between the OS and the device. Developed a hardware abstraction layer that provides a generic interface to multiple types of HBA controllers (AHCI, LSI Megaraid, NVM-e, DRAM SSD) in order to demonstrate the capability of interface extensions involving exposing H/W abstractions such as tags, queues & interrupts. Total 10k LoC Linux Kernel code developed to demonstrate the idea. Achieved 100% IOPS gain (from 300k IOPS to 600k IOPS) with the code. Reported the result in USENIX ATC'14 [10] ***(Acceptance rate 14.9%)***.

- **Optimized Systems for Big Data: Workload analysis for applying NVRAM components**
Internship, IBM Austin Research Laboratory, USA
Research Intern, September 2012 ~ June 2013
  - ○ Research on evaluating the impact of using NVRAM components on systems for big data analytics. Built a research prototype workload characterization module which extracts condensed application I/O models on the fly to solve the difficulties of extracting large trace data from distributed data intensive applications such as Hadoop. Hooked system calls such as open / close, read/write and lseek system calls per I/O thread to extract multiple Markov models from runtime traces with less than 5% performance degradation of the host program. The condensed Markov models were used to reproduce the application behavior. Total 10k LoC C code and 2.1k LoC Python code developed to demonstrate the idea.

- **Development of a File System based on Phase Change Memory
to Support more than 200,000 IOPS for Peta-scale Computing**
Research Grant, Ministry of Knowledge Economy (MKE) of Republic of Korea, Seoul National University
Research Assistant, May 2011 ~ August 2012

- ○ Research on optimizing the Linux 2.6.32 file I/O path to better support high performance storage devices, projected to be using new memory technologies. Performed profiling and optimization activities with the Linux 2.6.32 page cache & block layer code using a high-performance PCI-e 2.0 SSD backed with a large DRAM array. Fixed performance degradation when read-ahead was enabled (20%~30%) by patching the read-ahead code in the page cache which deals with multiple sequential streams. Results reported in HotStorage'12 [13] and TOCS [11]. Found unnecessary memory copies with DMA pages residing above the 32 bit limit (non DMA32 pages). In response, developed a new type of page cache code which places the page cache under the 32 bit limit to eliminate the additional memory copies. Total 3.6k LoC patched for additional functionality.
- **A Research on a Hierarchical Storage Architecture for Cloud Computing**
  Research Grant, National Research Foundation of Korea (NRF), Seoul National University
  Research Assistant, August 2010 ~ August 2012
  - ○ Research on an RDMA based remote memory disk cache for virtual machines. Developed a Linux kernel (v2.6.32) block device driver (1.4k LoC) which employs kernel level RDMA transactions to exploit remote DRAM exposed by remote memory donors, with the capability of direct DMA memory mapping support for virtualized environments. Employed and enhanced the dm-cache (Linux) code to better use our RDMA block device. Results reported in [14].

Industry Projects
- **Zero Failure Project**
  Software Engineer, TmaxSoft as subcontractor to LG Display ltd., June, 2009 ~ October, 2010
  - ○ Delivered an integrated library function intercept-based middleware application transaction monitoring system to an MES (Manufacture Execution System) at LG Display (South Korea). Integration involved developing customized function hooking modules for in-house MES applications running on TIBCO middleware & Oracle RDBMS.
  - ○ Among the 112k LoC of the delivered code, 31k LoC of on-site development has been done for this project on customization, stability enhancement, runtime management enhancements, testing and deployment. Product deployed on systems with the HP-UX IA64 architecture.
- **Non-intrusive C/C++ middleware transaction instrumentation for SysMaster**
  Research Engineer, TmaxSoft, June, 2006 ~ June, 2009
  - ○ Research and development on extracting end to end traces and performance metrics (latency) of multi-tiered enterprise applications by capturing & monitoring native application activities via LD_PRELOAD based system call & library call instrumentation which otherwise requires source code instrumentation.
  - ○ Designed and developed enterprise application end-to-end performance monitoring modules based on run-time function intervention (LD_PRELOAD) of middleware framework calls. Built the foundation to support multiple middleware target products with function hooking. Developed function hooking based transaction latency monitoring modules for products such as BEA Tuxedo & TmaxSoft Tmax and Oracle.
  - ○ Total 91k LoC in C developed from scratch for function hooking base functionality, lock-free shared memory-based IPC data exports and runtime management capabilities. Code was targeted for various UNIX systems such as HP-UX (PA-RISC & IA64), AIX (PPC64), Solaris (SPARC, x86_64) & Linux (x86_64).
- **Network Monitoring System Package Development (NMSPlus 3.0~3.1)**
  Software Engineer, Samsung Networks, August, 2003 ~ March, 2005
  - ○ Maintained and enhanced a collection of daemons which collects SNMP & ping statistics from various network devices (Cisco, Alcatel, Juniper).
  - ○ Contributed by adding additional SNMP based data collection support for ATM switches, L4

switches while involving multiple on-site customer support tasks.

Publications

[1] *<u>Woong Shin</u>*, Vladyslav Oles, Ahmad M. Karimi, J. Austin Ellis, Feiyi Wang, **"Revealing Power, Energy and Thermal Dynamics of a 200PF Pre-Exascale Supercomputer"**, in The International Conference for High Performance Computing, Networking, Storage, and Analysis (SC 2021), November 2021
<u>Acceptance rate 23.5%</u> **(Best paper award)**

[2] Chengcheng Li, Ahmad M. Karimi, *<u>Woong Shin</u>*, Hairong Qi, Feiyi Wang, **"The Challenge of Disproportionate Importance of Temporal Features in Predicting HPC Power Consumption"**, in the Workshop on Monitoring and Analysis for HPC Systems Plus Applications (HPCMASPA) held in conjunction with IEEE Cluster 2021 (CLUSTER2021), September 2021

[3] Alessio Netti, *<u>Woong Shin,</u>* Michael Ott, Torsten Wilde, Natalie Bates, **"A Conceptual Framework for HPC Operational Data Analytics"**, in the 2021 Energy Efficient High-Performance Computing Working Group State of Practice Workshop held in conjunction with IEEE Cluster 2021 (CLUSTER2021), September 2021

[4] Justin Thaler, *<u>Woong Shin</u>*, Steven Roberts, Jim Rogers, Todd Rosedahl, **"Hybrid Approach to HPC Cluster Telemetry and Hardware Log Analytics"**, in The 2020 IEEE High Performance Extreme Computing Virtual Conference, September 2020

[5] Michael Ott, *<u>Woong Shin</u>*, Norman Bourassa, Torsten Wilde, Stefan Ceballos, Melissa Romanus, Natalie Bates, **"Global Experiences with HPC Operational Data Measurement, Collection and Analysis"**, in the 2020 Energy Efficient High-Performance Computing Working Group State of Practice Workshop held in conjunction with IEEE Cluster 2020 (CLUSTER2020), September 2020

[6] *<u>Woong Shin</u>*, Christopher D Brumgard, Bing Xie, Sudharshan S Vazhkudai, Devarshi Ghoshal, Sarp Oral, Lavanya Ramakrishnan, **"Data Jockey: Automatic data management for HPC multi-tiered storage systems"**, in The 33rd IEEE International Parallel and Distributed Processing Symposium (IPDPS'19), Rio de Janeiro, Brazil, May 2019, *<u>Acceptance rate 27.7%</u>*

[7] *<u>Woong Shin</u>*, Jaehyun Park, Heon, Y. Yeom, **"Unblinding the OS to Optimize User-Perceived Flash SSD latency"**, in The 8th USENIX Workshop on Hot Topics in Storage and File Systems (USENIX HotStorage'16), Denver, CO, June 2016

[8] *<u>Woong Shin</u>*, Myeongcheol Kim, Kyudong. Kim, and Heon. Y. Yeom, **"Providing QoS through Host Controlled Flash SSD Garbage Collection and Multiple SSDs"**, in The Second International Conference on Big Data and Smart Computing (BigComp'15), Jeju Island, South Korea, 9-12 February, 2015,
(Best paper award)

[9] *<u>Woong Shin</u>*, Myeongcheol Kim, Jinyoung. Choi, Hyeonsang. Eom, and Heon. Y. Yeom, **"HIOPS-KV : Exploiting Multiple Flash Solid-state Drives for Key Value Stores",** in The fourth International Workshop on Extreme Scale Computing Application Enablement - Modeling and Tools (ESCAPE'14) in conjunction with ICPADS'14, Hsinchu, Taiwan, 16-19 December, 2014.

[10] *<u>Woong Shin</u>*, Quichen Chen, Myoung Won Oh, Hyeonsang Eom and Heon Y. Yeom, **"OS I/O Path Optimizations for Flash Solid-state Drives"**, USENIX ATC'14, June 2014, *<u>Acceptance rate 14.9%</u>*

[11] Young Jin Yu, Dong In Shin, *<u>Woong Shin</u>*, Nae Young Song, Jae Woo Choi, Hyeong Seog Kim, Hyeonsang Eom and Heon Y. Yeom, **"Optimizing Block I/O Subsystem for Fast Storage Devices"**, ACM Transactions on Computer Systems (TOCS), June 2014. (SCI)

[12] NaeYoung Song, Young Jin Yu, ***Woong Shin***, Hyeonsang Eom, Heon Y. Yeom, **"Low-latency Memory-Mapped I/O for Data-Intensive Applications on Fast Storage Devices"**, High Performance Computing, Networking, Storage and Analysis (SCC), 2012

[13] Young Jin Yu, Dong In Shin, ***Woong Shin***, Nae Young Song, Hyeonsang Eom, and Heon Y. Yeom, **"Exploiting Peak Device Throughput from Random Access Workload"**, USENIX HotStorage'12, June 2012

[14] Hyuck Han, Young Choon Lee, ***Woong Shin***, Hyungsoo Jung, Heon Y. Yeom, Albert Y. Zomaya, **"Cashing in on the Cache in the Cloud"**, IEEE Transactions on Parallel and Distributed Systems (TPDS), August 2012. (SCI)

Technical Reports

[1] Abston et. al., **"US Department of Energy, Office of Science, High Performance Computing Facility Operational Assessment 2019 Oak Ridge Leadership Computing Facility"**, Oak Ridge National Laboratory (ORNL), ORNL/SPR-2020/1499, April 2020

Presentations

- **2021 Energy Efficient High-Performance Computing State of Practice Workshop**
  Keynote presentation "Energy-Efficiency HPC in the Exascale era, What's Next?", September 2021

Professional Services

- **Smoky Mountain Computational Sciences & Engineering Conference, 2021**
  Program Committee
- **Energy Efficient High-Performance Computing Working Group State of Practice Workshop, 2021**
  Program Committee
- **Energy Efficient High-Performance Computing Working Group Workshop, 2020**
  Organizing Committee / Session Lead & Moderator - ARM Operational Experiences
- **Energy Efficient High-Performance Computing Working Group State of Practice Workshop, 2020**
  Program Committee
- **Cluster Computing - The Journal of Networks, Software Tools and Applications, 2019**
  Reviewer – Volume 22, Issue 2 – Special Issue - Storage System Challenges for Big Data

Professional Organizations

- **Energy-efficient HPC Working Group (https://eehpcwg.llnl.gov), 2019 ~ current**
  Operational Data Analytics Team, General member

Certificates

- Six Sigma Green Belt Certification, Samsung Networks, November 2005
- Korean Managing Projects Course, The George Washington University School of Business and Public Management, Project Management Institute, August 2005
- Sun Certified Programmer for the Java2 Platform 1.4, Sun Microsystems, December 2004

Technical Skills & Experiences

- **Software Development**
  - C language based cross platform system programming on Unix systems such as AIX, HPUX, Solaris, and Linux developing various daemons, CLI tools and libraries
  - Python application development for data stream processing (Kafka), task management, REST API

based web applications, CLI tools, and system automation applications
- Test-driven Python application testing using robotframework & PyTest
- Kubernetes based application development, deployment & maintenance
- GitOps flavored CI/CD pipeline design and implementation for Kubernetes & HPC targeted applications using Gitlab CICD runners and ArgoCD.
- **Distributed Systems**
  - High-availability setup, operation of stateful components such as Redis, MongoDB, Zookeeper, etcd, PostgreSQL, MySQL, RabbitMQ, MinIO, Druid and Kafka on top of Kubernetes
  - HPC job management on HPC resource managers such as Slurm, LSF and Torque Moab
  - AMQP based message broker programming on RabbitMQ
  - Prometheus + Grafana based distributed infrastructure monitoring setup and operations
  - Infrastructure configuration & state management using Ansible
  - Workflow engine (Apache Airflow) setup, deployment and use with data engineering and machine learning engineering workflows orchestrated on top of various compute and data resources.
- **Data Analytics**
  - Fault tolerant data pipeline design, implementation & maintenance of Kafka based data wrangling & processing pipelines using Python & Apache Spark
  - Large scale distributed data processing (TB range), wrangling & analytics using PyArrow, Pandas, Dask, and Apache Spark (Pyspark) processing highly compressed Parquet files on HPC systems such as Summit and Andes at OLCF
- **Data Center Instrumentation for Monitoring**
  - Network data collection & monitoring using SNMP & Netflow, interacting with network switches and routers from vendors such as Cisco & Juniper
  - IBM Power9 out-of-band instrumentation using OpenBMC telemetry streams
  - NVidia DCGM based GPU usage instrumentation of HPC jobs
  - Data collection from HPC resource managers such as Slurm & LSF
  - Monitoring building automation data from CEP PLC tag objects exposed by vendor proprietary gateway modules
  - Instrument industrial systems using BACNET/IP & Modbus/TCP
  - LD_PRELOAD based API call instrumentation of Oracle OCI, Tibco messaging, and OS file I/O system calls
- **Operating Systems, Embedded systems, and Hardware Digital Design**
  - Linux kernel modification & device driver development (Linux 2.6.32 ~ Linux 3.5.0) for storage system research: scheduler, block I/O subsystem, SCSI subsystem, NVM-e device driver modification & enhancement
  - Xilinx FPGA (Zynq-7000, Virtex7) programming using Verilog & HLS (Vivado) and embedded C programming (Xilinx Microblaze & ARM Cortex A9) to develop prototype PCI-e SSDs for research
  - Performance evaluation, tuning & optimization of high-performance SATA3 SSDs, 1st & 2nd generation NVM-e SSDs under Linux as a host-system