Seunghyun Lee

Contact Information

mobile: 919 - 699 - 8095
e-mail: seunghy1@cs.cmu.edu

Education

Carnegie Mellon University, School of Computer Science, Pittsburgh, PA
Master of Computational Data Science, Systems Major **December 2016 (Expected)**
· Course Highlights: Distributed Systems, Cloud Computing, Machine Learning, Storage Systems

Duke University, Durham, NC
Bachelor of Science in Computer Science, Minor in Mathematics **May 2015**
· GPA: 3.68/4.0
· Awards: Dean's List for Spring 2010, Fall 2012, Fall 2013, Spring 2014
· Course Highlights: Data-Intensive Computing Systems, Computer Networks, Operating Systems, Software Design & Implementation, Data Structures, Algorithms, Computer Architecture

Research Experience

YARN Resource Management and Scheduling **March 2015 – May 2015**
Advisor: Shivnath Babu, Duke University
· Implemented the framework for profiling scheduler's metrics information using YARN's REST API.
· Set up two different scenarios (batch and interactive intensive) using MapReduce/Hive, and conducted experiments to optimize configurations for Capacity Scheduler.
· Improved the performance of the base case by 20%.

Connected Components on MapReduce **May 2014 – July 2014**
Advisor: Kamesh Munagala, Duke University
· Implemented existing MapReduce algorithms for connected components such as Hash-Min, Hash-To-All, Hash-To-Min on top of Spark.
· Conducted in-depth comparisons of algorithms by analyzing collected data such as the execution time, the number of iterations, and the size of intermediate data.

Work Experience

Frograms, Seoul, Korea **September 2014 – December 2014**
Software Engineer Intern
· Designed and implemented a quiz and poll generating tool using Ruby on Rails and MySQL.
· Designed and implemented a backend part of an analytics tool that visualizes the collected data for the effect of advertisement.

SAP Labs Korea, Seoul, Korea **July 2013 – August 2013**
Software Engineer Intern
· Conducted a rigorous performance comparisons on Hadoop/Hive, Spark/Shark, and SAP HANA using TPC-H data set and its queries.
· Deployed and managed a large cluster with two hundred cores and 2TB RAM.

Duke University, Durham, NC **January 2013 – May 2013**
Undergraduate Teaching Assistant, Department of Computer Science
· Teaching Assistant for Data Strucures and Algorithms course, Instructor : Tabitha Peck.
· Conducted weekly office hours and graded programming assignments.

Course Projects

· **Malloc**: Implemented a dynamic storage allocator in C. Investigated and implemented various data structures(explicit list, segregated list) and algorithms(best fit, first fit) for optimizing performance.
· **Proxy**: Implemented a web proxy that caches web objects. It also supports concurrent requests by multi-threading.
· **SLogo**: Designed and implemented an interpreter of Logo Language in Java. Mostly worked on the back-end part that parses and builds the execution tree of the user input commands.
· **Devil Shell**: Built a basic functional shell that supports job control, input/output redirection, error handling, and pipelines.
· **DeFiler**: Implemented a simple multi-threaded file system in Java capable of storing and managing concurrent access to the multiple files.

Skills

Java, C, Python, Ruby on Rails, Hadoop/Hive, Spark, SQL, HTML/CSS/JS, Git, LaTeX