Peter Van Sandt

eyepatchParrot.github.io
linkedin.com/in/petervansandt

WORK EXPERIENCE

Jan 2022 - Present **Citadel Securities**, *Systems Software Engineer*, Chicago, IL

Aug 2020 - Jan 2022 *Citadel Enterprise*

- Accelerate GPU, multi-core and single-core implementations for xgboost and LightGBM forest inference with improved integration into research and production workflows and frameworks at higher accuracy
- Plan roadmap for evolution of a C++ library by collaborating with several desks to identify and prioritize development
- Survey hyperparameter dead-ends and ML training optimizations with benchmarking analysis and data synthesis to reduce wasted experiments and cost by 80%

Oct 2019 - Jun 2020 **Dell-EMC Isilon**, *Senior Software Engineer*, Seattle, WA

Jun 2018 - Oct 2019, 2017 *Software Engineer - Performance*, *Software Engineer Intern - Performance*

- Designed and developed request handling, parsing, and dispatch in development of new protocol stack: minimizing data copies and branches in create, read, delete, and list operations.
- Researched and regularly presented performance and architectural analysis of academic and competing stacks to leadership team: focusing on concurrent creates in parallel file systems and throughput in object storage.
- Developed and extended non-intrusive, trace-based and low-overhead, distribution-based workload analysis tools with dtrace and SQL, so that customers could size their clusters 200x faster and support engineers could non-disruptively assess on-going workload distributions.
- Developed and extended automated performance analysis tooling with pandas: modeling read path on a distributed file system to highlight reads that waste cache and identifying a 50% performance regression.
- Root cause analysis of performance issues in an NFS protocol stack and authentication reducing CPU utilization in hot spots and improve worst case performance.

Oct 2016 - May 2018 **University of Wisconsin-Madison, Computer Sciences**, *Undergraduate Researcher*

- Developed interpolation search algorithms for uniform and non-uniform data to use constant factor improvements and asymptotic improvements using the distribution of the data to improve throughput in numpy by 380% and LevelDB by 146%.
- Optimized LevelDB by replacing use of mutexes in the fast-path for reads with copy-on-write version updates and lock-free LRU improving read throughput by up to 45%.

Jun 2016 - Aug 2016 **Jump Trading, LLC.**, *Software Developer Intern*, Chicago, IL

- Eliminated algorithmic bottlenecks in 100KLOC internal application improving response time by 200%.

EDUCATION

May 2018 **B.S., Computer Sciences, Honors, ΦBK**, *University of Wisconsin-Madison*, GPA: 4.0/4.0

- Graduate Coursework: Advanced Algorithms, High Performance Computing, Distributed Systems
- Languages: C++, CUDA C++, Python, AVX, dtrace

TALKS & PAPERS

2021 "Dissecting the Ampere GPU Architecture through Microbenchmarking" *NVIDIA GPU Technology Conference*

2019 Van Sandt, Peter, Yannis Chronis, and Jignesh M. Patel. "Efficiently Searching In-Memory Sorted Arrays: Revenge of the Interpolation Search?." *Proceedings of the 2019 International Conference on Management of Data.* 2019.

PERSONAL PROJECTS

2017 Developed GPU addition algorithm based on Brent-Kung adder using CUDA C++ minimizing number of kernels launched and achieving $O(\log N)$ span and $O(N \log N)$ work.

2016 Vectorized insertion and rank order sorts of fixed-size arrays using C and AVX2 intrisics.

2012 Developed GPU-accelerated, Toom-Cook multiplication bignum library in C++ and CUDA C++.