Yan Li

TuneUp.ai
San Francisco Bay Area

Phone: (831) 713-6616
Email: yanli@tuneup.ai
Website: http://elliot.li

TL;DR

Ph.D. Systems and Deep Learning Researcher. Inventor. Founder. Yan focuses on applying the latest machine learning techniques to computer system problems. Yan's Ph.D. advisor was Professor Darrell Long. Yan has an Erdős number 3.

JOBS

March '17

Founder, TuneUp.ai
Invented a method that uses Deep Reinforcement Learning to tune the parameters of storage systems for better performance in a fully automated, unsupervised manner. Achieved 45% increase of throughput in evaluations on Lustre. Paper (http://elliot.li/res/li-capes-sc17.pdf) was nominated for best student paper in Supercomputing '17. Developed the method into a scalable online service that offers fully automated parameter tuning for infrastructure, storage, or any computer/industrial systems. See https://tuneup.ai for more information.

Summer '15

Research Intern, TurboStor
Developed a fully functional, lock free, high performance key-value store backend for Ceph in C++ using Hierarchical B+Tree based Trie. Also formally proved the correctness of the lock free multithreaded design using +Cal/TLA+. A high level introduction can be found at https://elliot.li/res/lockless-ds-and-proof-distilled.pdf.

2013 - 2016

Administrator, ACM Transactions on Storage
ACM Transactions on Storage (http://tos.acm.org/) is the preeminent journal for computer storage research.

Summer '14

Research Intern, Symantec Research Labs
System management and optimization automation, reducing complexity and costs.

Summer '13

Research Intern, IBM Almaden Research Center
Implemented a smart Quality-of-Service system for IBM's GPFS file system. The system automatically selects rate limits to prevent congestion collapse and to reduce response time according to dynamic workload.

Summer ’12 **Engineering Intern, Google**

Worked in the Linux kernel performance team and developed a new Linux kernel scheduler with reduced context switching overhead, aiming at fixing the long tail latency issue of certain queries. The kernel mode scheduler coordinates with a user space scheduler, which is aware of workload specific information, to achieve optimal scheduling.

2008 - 2011 **Software Engineer, MeeGo core system developer, Intel**

Worked on the MeeGo Linux distro for Intel’s mobile platforms. Responsibilities included kernel, performance, stability, hardware compatibility, file systems, enterprise desktop, email client, Chromium browser, GPL-compliance, and others.

MeeGo was the next-generation research operating system sponsored by Intel, Linux Foundation, and Samsung for emerging mobile platforms, such as handset, tablet and in-vehicle infotainment systems. Yan’s worked covered:

- Linux kernel maintainer: maintained Intel’s patchsets on the latest Linux kernel, oversaw the kernel release process, was responsible for the overall kernel quality.
- Ported Linux kernels and Xorg to pre-release Intel hardware.
- Developed a secure file-system for MeeGo (eCryptfs based).
- Developed persistent on-disk caching of KBD compilation of Xorg, increased the Xorg start-up speed by 1 to 2 seconds on all platforms.
- Designed the process for ensuring legal compliance of using GPLv3 licensed software in MeeGo.
- Developed infrastructure and certain components (VPN, Microsoft Exchange email client, WPA2 EAP, etc.) for using MeeGo in enterprise environment.

2010 **Co-admin and mentor of Google Summer of Code 2010 for Intel’s MeeGo Platform**

Co-administered the MeeGo/Maemo project for Google Summer of Code 2010. Mentored one student who worked on cloud storage support for MeeGo.

2009 - 2012 **GNOME Foundation Member**

GNOME Foundation (http://www.gnome.org/about/) is a non-profit desktop software development organization. Yan worked as a contributor to Evolution, the personal information management application.

2006 - 2008 **Staff Software Engineer, IBM**

QA Project Manager of the Asianux on POWER team. Led a 4-person QA team and worked on bug-fixing, validation, and performance tuning of several major Linux distributions (includes Red Hat Enterprise Linux, Novell SUSE Linux Enterprise, IBM MCP, and Asianux) on IBM POWER systems.

PROJECTS

2017 - now **TuneUp.ai LightningCloud**

LightningCloud is a cloud-based performance tuning solution that uses Deep Reinforcement Learning to tune the parameters of the target system, such as process and I/O schedulers for the OS, readahead settings for block devices, or keepalive_requests and limit_rate of NGINX. The client agent runs on the target system, collecting runtime information and changing parameters according to instructions from the engine running in the cloud.

The client agent is open sourced and can be found at https://github.com/mlogic/tuclient. The cloud engine is proprietary.

Patents:

- Yan Li. "Parameter Tuning with Continuous Deep Q Reinforcement Learning". US Patent Application. Pending. Not published. Full text available upon request. 2018.
- Yan Li. "Methods and Apparatus for Parameter Tuning Using a Cloud Service". US Patent Application. Pending. Published as US20180351816A1 (https://patents.google.com/patent/US20180351816A1). 2017.

2015 - 2017 **CAPES: Deep Reinforcement Learning Based Performance Tuning**

CAPES was Yan's PhD thesis project and was an early research prototype of using Deep Reinforcement Learning to tune the I/O queue depth and rate limits of a Lustre storage cluster to increase its performance. Evaluation of a prototype on a Lustre file system demonstrates an increase in I/O throughput by up to 45% at saturation point.

Source code: https://github.com/mlogic/capes-oss

Publication:

- "CAPES: Unsupervised Storage Performance Tuning Using Neural Network-Based Deep Reinforcement Learning." *The 2017 International Conference for High Performance Computing, Networking, Storage and Analysis* (SC17), Denver, CO, USA: November 13–16, 2017. http://elliot.li/res/li-capes-sc17.pdf.

Oct '15 - now **Pilot: Automatic performance measurement**

Carrying out even the simplest form of performance benchmark requires considerable knowledge of statistics and computer systems, and painstakingly following many error-prone steps in order to meet all the requirements in accuracy, precision, comparability, repeatability, and control of overhead. Yan designed Pilot, which implements a collection of algorithms and heuristics to automate the steps that are necessary to produce statistically sound performance measurement results. It covers the collection, storing, analyses, and comparing of performance measurements and can help to reduce human effort and error.

Source code and documentation: https://tuneup.ai/en/pilot

2013 - 2016 **ASCAR: Automatic Rule-based Storage Congestion Control**

Heavy storage workloads could cause congestion, which greatly reduces throughput and response time for high performance storage systems like Lustre. ASCAR is designed to control congestion by using traffic control rules to limit the I/O rate in a smart manner. It automatically designs and evaluates traffic control rules for the target system in an unsupervised manner. In one evaluation, ASCAR improved the throughput of a NASA NPB BTIO checkpoint workload by 33.5% and reduced its speed variance by 55.4% at the same time.

Publication:

- "ASCAR: Automating Contention Management for High-Performance Storage Systems." *31st International Conference on Massive Storage Systems and Technologies* (MSST 2015), Santa Clara, CA, USA: May 30–June 5, 2015.
  http://elliot.li/res/ascar-msst15-final.pdf

Source code:

- https://github.com/mlogic/ascar-lustre-2.9-client
- https://github.com/mlogic/ascar-lustre-2.4-client
- https://github.com/mlogic/ascar-lustre-sharp

2011-2012 **Horus, Fine-Grained Encryption-Based Security for Large-Scale Storage**

Management of encryption keys on a fine-grained level is essential for keeping shared datasets secure. However, the sheer size of these keys could easily reach GB-level for a large dataset, imposing challenges for managing and storing them. Horus solved this problem by using keyed hash trees to securely and deterministically generate unique encryption keys on-demand, for regions of any desired size, providing fine-grained security control with much lower overhead than other solutions.

Source code: https://github.com/mlogic/horus

Publication:

- "Horus: Fine-Grained Encryption-Based Security for Large-Scale Storage," *Proceedings of the 11th USENIX Conference on File and Storage Technologies* (FAST '13), San Jose, CA, USA: Usenix Association, February 12–15, 2013. http://elliot.li/res/li-fast13.pdf

HONORS AND AWARDS

- 2017, Honorable Mention for CS Dissertation Award, UC Santa Cruz.
- 2014, Symantec Research Labs Graduate Fellowship.
- 2012, UCSC Graduate Research Symposium Alumni Association Award.
- 2009, Special Recognition Award, Intel.
- 2008, First Patent Plateau Award of IBM for filing 4 patent applications.
- 2007, IBM First Patent Award.
- 2004, China National Third Prize for the Electric Grid Line Loss Online Management System, by Chinese Society for Electrical Engineering. Yan Li was the software architect and one of the main developers.
- 2003, National Fellowship for Graduate Study during Master program.
- 2002, Shandong Province Undergraduate Electronic Design Contest, China, First Prize.
- 2001, China National Undergraduate Electronic Design Contest, First Prize.
- 1997, China National Olympiad in Informatics, Shandong District, Gold Medal.
- 1995, Shandong Olympiad in Informatics, China, First Prize.

PUBLICATIONS

Refereed Conference Papers

1. **Yan Li**, Kenneth Chang, Oceane Bel, Ethan L. Miller, Darrell D. E. Long. "CAPES: Unsupervised Storage Performance Tuning Using Neural Network-Based Deep Reinforcement Learning." *The 2017 International Conference for High Performance Computing, Networking, Storage and Analysis* (SC17), Denver, CO, USA: November 13–16, 2017.
2. **Yan Li**, Yash Gupta, Ethan L. Miller, Darrell D. E. Long. "Pilot: A Framework that Understands How to Do Performance Benchmarks The Right Way." *IEEE 24th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems* (MASCOTS 2016), London, UK: September 19–21, 2016.
3. Yulai Xie, Dan Feng, **Yan Li**, Darrell D. E. Long, "Oasis: An active storage framework for object storage platform." *Future Generation Computer Systems 56 (2016) 746-758*.
4. **Yan Li**, Xiaoyuan Lu, Ethan L. Miller, Darrell D. E. Long. "ASCAR: Automating Contention Management for High-Performance Storage Systems." *31st International Conference on Massive Storage Systems and Technologies* (MSST 2015), Santa Clara, CA, USA: May 30–June 5, 2015.
5. **Yan Li**, Darrell D. E. Long. "Which Storage Device is the Greenest? Modeling the Energy Cost of I/O Workloads." *IEEE 22nd International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS 2014)*, Paris, France: September 9–11, 2014.
6. **Yan Li**, Nakul Sanjay Dhotre, Yasuhiro Ohara, Thomas M. Kroeger, Ethan L. Miller, Darrell D. E. Long. "Horus: Fine-Grained Encryption-Based Security for Large-Scale Storage," *Proceedings of the 11th USENIX Conference on File and Storage Technologies* (FAST '13), San Jose, CA, USA: Usenix Association, February 12–15, 2013.
7. Yulai Xie, Kiran-Kumar Muniswamy-Reddy, Dan Feng, **Yan Li**, Darrell D. E. Long, Zhipeng Tan, Lei Chen. "A Hybrid Approach for Efficient Provenance Storage," *The 21st ACM Conference on Information and Knowledge Management* (CIKM '12), Maui, HI, USA: ACM, October 29–November 2, 2012.
8. **Yan Li**, Ethan L. Miller, Darrell D. E. Long. "Understanding Data Survivability in Archival Storage Systems," *The 5th Annual International Systems and Storage Conference (SYSTOR 2012)*, Haifa, Israel: ACM, June 4–6, 2012.

Invited Lectures

1. "Horus: Fine-Grained Encryption-Based Security for Large-Scale Storage," Sandia National Laboratories, 2012.

Industrial Conference Talks

1. **Yan Li**. "ASCAR: Increasing Performance Through Automated Contention Management," (video: https://goo.gl/lRKcpn, slides: http://goo.gl/8BbjCa), *Lustre User Group 2016*, Portland, Oregon, U.S.A., April, 2016.
2. **Yan Li**. "MeeGo Enterprise Goes Beyond Desktop," (http://sf2011.meego.com/program/sessions/meego-enterprise-goes-beyond-desktop), *MeeGo Conference 2011*, San Francisco, U.S.A., May, 2011.
3. **Yan Li**. "MeeGo and GPLv3," (http://sf2011.meego.com/program/sessions/meego-and-gplv3), *MeeGo Conference 2011*, San Francisco, U.S.A., May, 2011.
4. **Yan Li**. "Using MeeGo as Enterprise Desktop," (http://conference2010.meego.com/session/using-meego-enterprise-desktop). *MeeGo Conference 2010*, Dublin, Ireland, November, 2010.

Patents and Related Publications

1. **Yan Li**. 2018. "Parameter Tuning with Continuous Deep Q Reinforcement Learning". Application not published yet. Full text available upon request.
2. **Yan Li**. 2017. "Methods and Apparatus for Parameter Tuning Using a Cloud Service". Application. US20180351816A1.
3. Yao Qi, **Yan Li**, Wei Ying Yu, Yong Zheng. 2007. "Providing Customizable, Process-specific Just-in-time Debugging in an Operating System". Granted. US9128837B2.

4. Yao Qi, **Yan Li**, Wei Ying Yu, Yong Zheng. 2007. “Method and System for Debugging a Program in a Multi-thread Environment”. Granted. US8201152B2.

Journals

1. Yulai Xie, Kiran-Kumar Muniswamy-Reddy, Dan Feng, **Yan Li**, Darrell D. E. Long. Evaluation of A Hybrid Approach for Efficient Provenance Storage. ACM Transactions on Storage. 9(4). Nov. 2013. http://elliot.li/res/a14-xie.pdf
2. **Yan Li**. 2005. Use WANT to Build Delphi Project. Programmer Magazine, Feb 2005.
3. Jin Li, **Yan Li**. 2002. Information Collect System in Power Department. Journal of Qingdao Institute of Chemical Technology Vol. 23: 65-67.

Book

1. **Yan Li**, Jingjing Liu, and Yiming Yu. 2006. J2ME Development and Application. China Machine Press, Beijing, China. ISBN: 7111188349

Technical Report

1. **Yan Li**. 2006. Research on Refactorable Software Platform: A Framework Reuse Solution. Qingdao: Ocean University of China.

EDUCATION

2017	Ph.D.	University of California, Santa Cruz, Computer Science
2006	M.E.	Ocean University of China, Computer Applied Technology
2003	B.E.	Ocean University of China, Electronic Engineering

Last updated: October 22, 2019

`http://elliot.li/yanli-cv.pdf`