# HYEONTAEK LIM

Computer Science Department
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213, USA

Phone: 1-412-626-4271
Fax: 1-412-268-5576
E-mail: hl@cs.cmu.edu
WWW: www.cs.cmu.edu/~hl

## Employment

09/2015– **Post Doctoral Fellow**. . . . . Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA

## Education

08/2009–09/2015 **Carnegie Mellon University**, Pittsburgh, PA, USA
Ph.D. in Computer Science
Thesis: "Resource-efficient data-intensive system designs for high performance and capacity"
Advisor: **David G. Andersen**
Committee: David G. Andersen, Michael Kaminsky, Andrew Pavlo, Eddie Kohler

03/2002–01/2009 **Korea Advanced Institute of Science and Technology (KAIST)**, Daejeon, Korea
B.S. in Computer Science, Summa Cum Laude, 2009
First in class of Computer Science, 2009

## Research Interests

Resource-efficient and flexible distributed systems, networking, and operating systems.

## Publications

Refereed Z. Wang, A. Pavlo, H. Lim, V. Leis, H. Zhang, M. Kaminsky, and D. G. Andersen. **Building a Bw-tree takes more than just buzz words**. In *Proceedings of ACM SIGMOD International Conference on Management of Data (SIGMOD) 2018*, June 2018.

H. Zhang, H. Lim, V. Leis, D. G. Andersen, M. Kaminsky, K. Keeton, and A. Pavlo. **SuRF: Practical range query filtering with fast succinct tries**. In *Proceedings of ACM SIGMOD International Conference on Management of Data (SIGMOD) 2018*, June 2018. Best Paper Award.

C. Canel, T. Kim, G. Zhou, C. Li, H. Lim, D. Andersen, M. Kaminsky, and S. Dulloor. **Picking interesting frames in streaming video**. In *SysML Conference 2018*, Feb. 2018.

G. Zhou, T. Kim, C. Canel, C. Li, H. Lim, D. Andersen, M. Kaminsky, and S. Dulloor. **Efficient multi-tenant inference on video using microclassifiers**. In *SysML Conference 2018*, Feb. 2018.

H. Lim, M. Kaminsky, and D. G. Andersen. **Cicada: Dependably fast multi-core in-memory transactions**. In *Proceedings of ACM SIGMOD International Conference on Management of Data (SIGMOD) 2017*, May 2017.

S. Li, H. Lim, V. W. Lee, J. H. Ahn, A. Kalia, M. Kaminsky, D. G. Andersen, S. O, S. Lee, and P. Dubey. **Achieving one billion key-value requests per second on a single server**. *IEEE Micro*, 36(3):94–104, May 2016a. Top Picks.

S. Li, H. Lim, V. W. Lee, J. H. Ahn, A. Kalia, M. Kaminsky, D. G. Andersen, S. O, S. Lee, and P. Dubey. **Full-stack architecting to achieve a billion-requests-per-second throughput on a single key-value store server platform**. *ACM Transactions on Computer Systems*, 34(2): 5:1–5:30, May 2016b.

H. Lim, D. G. Andersen, and M. Kaminsky. **Towards accurate and fast evaluation of multi-stage log-structured designs**. In *Proceedings of the 14th USENIX Conference on File and Storage Technologies (FAST)*, Feb. 2016.

D. Zhou, B. Fan, H. Lim, D. G. Andersen, M. Kaminsky, M. Mitzenmacher, R. Wang, and A. Singh. **Scaling up clustered network appliances with ScaleBricks**. In *Proceedings of ACM SIGCOMM 2015*, Aug. 2015.

S. Li, H. Lim, V. W. Lee, J. Ahn, A. Kalia, M. Kaminsky, D. G. Andersen, S. O, S. Lee, and P. Dubey. **Architecting to achieve a billion requests per second throughput on a single key-value store server platform**. In *Proceedings of the 42nd International Symposium on Computer Architecture (ISCA)*, June 2015. Selected as one of Micro's Top Picks from the Computer Architecture Conferences (Top Picks). Invited to appear as an extended version at ACM Transactions on Computer Systems (TOCS) via fast-track processing.

D. Naylor, M. K. Mukerjee, P. Agyapong, R. Grandl, R. Kang, M. Machado, S. Brown, C. Doucette, H.-C. Hsiao, D. Han, T. H.-J. Kim, H. Lim, C. Ovon, D. Zhou, S. B. Lee, Y.-H. Lin, C. Stuart, D. Barrett, A. Akella, D. Andersen, J. Byers, L. Dabbish, M. Kaminsky, S. Kiesler, J. Peha, A. Perrig, S. Seshan, M. Sirbu, and P. Steenkiste. **XIA: Architecting a more trustworthy and evolvable Internet**. *ACM SIGCOMM Computer Communications Review (CCR)*, 44(3):50–57, July 2014.

H. Lim, D. Han, D. G. Andersen, and M. Kaminsky. **MICA: A holistic approach to fast in-memory key-value storage**. In *Proceedings of the 11th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, Apr. 2014.

D. Zhou, B. Fan, H. Lim, D. G. Andersen, and M. Kaminsky. **Scalable, high performance Ethernet forwarding with CuckooSwitch**. In *Proc. 9th International Conference on emerging Networking EXperiments and Technologies (CoNEXT)*, Dec. 2013.

B. Fan, D. Zhou, H. Lim, M. Kaminsky, and D. G. Andersen. **When cycles are cheap, some tables can be huge**. In *Proceedings of the 14th USENIX conference on Hot Topics in Operating Systems (HotOS)*, May 2013.

H. Lim, D. G. Andersen, and M. Kaminsky. **Practical batch-updatable external hashing with sorting**. In *Proceedings of the Algorithm Engineering and Experiments (ALENEX)*, Jan. 2013.

D. Han, A. Anand, F. Dogar, B. Li, H. Lim, M. Machado, A. Mukundan, W. Wu, A. Akella, D. Andersen, J. Byers, S. Seshan, and P. Steenkiste. **XIA: Efficient support for evolvable internetworking**. In *Proceedings of the 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, Apr. 2012.

A. Anand, F. Dogar, D. Han, B. Li, H. Lim, M. Machado, W. Wu, A. Akella, D. G. Andersen, J. W. Byers, S. Seshan, and P. Steenkiste. **XIA: An architecture for an evolvable and trustworthy Internet**. In *Proceedings of the Tenth ACM Workshop on Hot Topics in Networks (HotNets)*, Nov. 2011.

H. Lim, B. Fan, D. G. Andersen, and M. Kaminsky. **SILT: A memory-efficient, high-performance key-value store**. In *Proceedings of the 23rd ACM Symposium on Operating Systems Principles (SOSP)*, Oct. 2011.

B. Fan, H. Lim, D. G. Andersen, and M. Kaminsky. **Small cache, big effect: Provable load balancing for randomly partitioned cluster services**. In *Proceedings of the 2nd ACM Symposium on Cloud Computing (SOCC)*, Oct. 2011.

H. Kim, H. Lim, J. Jeong, H. Jo, J. Lee, and S. Maeng. **Transparently bridging semantic gap in CPU management for virtualized environments**. *Journal of Parallel and Distributed Computing (JPDC)*, 71(6):758–773, 2011. Special Issue on Cloud Computing.

H. Kim, H. Lim, J. Jeong, H. Jo, and J. Lee. **Task-aware virtual machine scheduling for I/O performance**. In *Proceedings of ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments (VEE)*, Mar. 2009.

Other H. Lim, D. G. Andersen, and M. Kaminsky. **3LC: Lightweight and effective traffic compression for distributed machine learning**. Technical Report arXiv:1802.07389, arXiv.org, Feb. 2018.

T. Kim, S. Boucher, H. Lim, D. G. Andersen, and M. Kaminsky. **Simple cache partitioning for networked workloads**. Technical Report CMU-CS-17-125, CMU, Oct. 2017.

H. Lim, V. Sekar, Y. Abe, and D. G. Andersen. **NetMemex: Providing full-fidelity traffic archival**. Technical Report arXiv:1603.04387, arXiv.org, Mar. 2016. Originally submitted to ACM SIGCOMM 2013.

H. Lim. ***Resource-efficient data-intensive system designs for high performance and capacity.*** PhD thesis, Carnegie Mellon University, Sept. 2015.

A. Anand, F. Dogar, D. Han, B. Li, H. Lim, M. Machado, W. Wu, A. Akella, D. Andersen, J. Byers, S. Seshan, and P. Steenkiste. **XIA: An architecture for an evolvable and trustworthy Internet**. Technical Report CMU-CS-11-100, Carnegie Mellon University, Feb. 2011.

H. Kim, H. Lim, J. Jeong, H. Jo, J. Lee, and S. Maeng. **Transparently bridging semantic gap in CPU management for virtualized environments**. Technical Report CS/TR-2009-311, KAIST, June 2009.

Patent H. Lim. **The method and system for rendering surface slope using a slope vector**. Republic of Korea Patent 10-0546865, issued January 2006. Held by NEXON Corporation, Seoul, Korea.

## Research Experience

04/2013– **MICA** . . . . . CMU, Pittsburgh, PA, USA

Designed and implemented MICA, a fast in-memory key-value store, which takes a holistic approach to combine a software architecture for efficient parallel processing, compact and fast data structures, and the extensive use of modern hardware support to provide a high-throughput key-value cache and store service for multi-10 gigabit networks using a general-purpose server.

11/2010– **The eXpressive Internet Architecture** . . . . . CMU, Pittsburgh, PA, USA

Devised a DAG (directed acyclic graph) addressing scheme and packet forwarding mechanism. Implemented a high-speed (25 Gbps) XIA software router prototype on the Click modular router. Designed and performed live migration experiments on XIA. Co-worked on ScaleBricks and CuckooSwitch, which improve the scalability and performance of networked appliances by using space-efficient and fast forwarding tables.

04/2014–02/2016 **Multi-stage log-structured designs** . . . . . CMU, Pittsburgh, PA, USA

Devised a new evaluation method for multi-stage log-structured designs to accurately and quickly estimate and optimize the performance of write-optimized stores such as LevelDB and RocksDB.

08/2010–03/2013 **FAWN: A Fast Array of Wimpy Nodes** . . . . . CMU, Pittsburgh, PA, USA

Worked on SILT (Small Index Large Table) that greatly improves the memory efficiency of a flash-based key-value store while preserving its high performance, enabling it to scale to billions of keys in a single key-value store node. Co-worked on SCBE (Small Cache Big Effect), which strives to achieve provably good load balance in a randomly partitioned cluster service by adding a small popularity-based query cache to the front-end query director.

10/2009–03/2014 **A traffic archiving system** . . . . . CMU, Pittsburgh, PA, USA

Designed and implemented a fast and cost-effective database system optimized for archiving the whole Internet traffic and processing forensic and other retrospective queries.

05/2012–05/2013 **Data-parallel computing** . . . . . Microsoft Research Redmond, WA, USA
CMU, Pittsburgh, PA, USA

02/2008–05/2009 **Task-aware virtual machine scheduling for I/O performance** . . . . . KAIST, Daejeon, Korea

Co-worked on a new virtual machine scheduling method that improves the I/O performance of virtual machines whose workloads are mixtures of CPU- and I/O-intensive tasks. Formulated a machine learning scheme for inferring I/O-bound tasks without requiring explicit task scheduling information from guest OSes. Designed exact-time credit accounting. Devised an evaluation framework and conducted preliminary evaluation. Inspected anomalies in the Xen credit scheduler.

02/2008–09/2008 **Mimicking failures for fairness in cooperative backup** . . . . . KAIST, Daejeon, Korea

Devised and evaluated a strategy enforcing fairness among cooperative backup nodes in a simulated open network by mimicking data loss symptoms and validating peers' responses.

07/2007–08/2007 **A page-granularity garbage collector** . . . . . KAIST, Daejeon, Korea
Studied on a new garbage collector that traces object references in granularity of pages for fast garbage collection. Developed an object reference analyzer in Jikes RVM for a preliminary investigation.

03/2007–06/2007 **KAIST Educational Network System (KENS)** . . . . . KAIST, Daejeon, Korea
Improved KENS for advanced education in computer networks. Added emulation of congestion and link failures. Implemented a reference version of Routing Information Protocol (RIP).

## Work Experience

04/2015–08/2015 **Software Engineer** . . . . . Marianas Labs, Inc., CA, USA

05/2012–08/2012 **Research Intern** . . . . . Microsoft Research Redmond, WA, USA
Mentored by Srikanth Kandula and Peter Bodik.

01/2004–06/2006 **Software Engineer** . . . . . NEXON Corporation, Seoul, Korea
Developed and deployed client software for massively multiplayer online games. Designed and implemented networking engines, graphics engines, and a peer-to-peer game image sharing module. Devised user interfaces and information visualization methods. Fulfilled military duty in Korea.

## TA

Spring 2014 Distributed Systems (Instructors: David G. Andersen, Srinivasan Seshan)

Spring 2012 Advanced Operating Systems and Distributed Systems (Instructor: Hui Zhang)

## Coursework (CMU Only)

Spring 2012 Special Topics in Software Systems: Memory and Resource-Efficient Big Data Computation (Instructor: David G. Andersen)

Fall 2011 Graphics and Imaging Architectures (Instructor: Kayvon Fatahalian)

Spring 2011 Machine Learning (Instructor: Tom M. Mitchell)

Fall 2010 Computer Networks (Instructor: Srinivasan Seshan)

Fall 2010 Algorithms in the Real World (Instructors: Guy E. Blelloch, Jeremy T. Fineman)

Spring 2010 Advanced Operating Systems and Distributed Systems (Instructor: David G. Andersen)

Spring 2010 Graduate Algorithms (Instructor: Gary L. Miller)

Fall 2009 Computer Architecture (Instructor: Todd C. Mowry)

Fall 2009 Type Systems for Programming Languages (Instructor: Robert Harper)

## Select Honors and Awards

09/2012–05/2013 Facebook Fellowship

02/2009 President's Award, KAIST Alumni Association

09/2008–08/2012 Doctoral Study Abroad Scholarship, The Korea Foundation for Advanced Studies (KFAS)

09/2003–12/2007 Department Scholarship, Division of Computer Science, KAIST

09/2002–06/2003 Merit-based Scholarship, KAIST

03/2002–06/2002 Admission Scholarship, KAIST

03/2002–06/2008 Full Tuition Scholarship, KAIST

## Service and Other Activities

**Reviewer**, IEEE Transactions on Computing 2014–2018.

**External Reviewer**, ACM SIGMETRICS 2017, ACM EuroSys 2017, ACM ASPLOS 2017, IEEE INFOCOM 2016.

01/2013– **Member**, SIAM.

07/2011– **Member**, ACM, IEEE, USENIX.

11/2002 **Staff**. . . The 2002 ACM International Collegiate Programming Contest (ICPC) Asia Regional Contest, Daejeon, Korea

08/2002 **Technical Engineer**. . . The 14th International Olympiad in Informatics (IOI 2002), Yongin, Korea