Abstract
From training a machine learning model with billions of operations to running a large datacenter with tens of thousands of servers, designing a large real-world system that is both scalable and fault-tolerant is never an easy task. Many design ideas have failed in the past three decades, but many more succeeded, and what we achieved so far were considerable feats of engineering. In this talk, I will share my own ideas and thoughts on the pitfalls and lessons learned on designing and implementing scalable, fast, and practical real-world systems. To support my ideas, I will share anecdotal evidence from a long history of real-world systems, ranging from a simple time synchronization protocol to distributed machine learning systems in recent academic work.

About the Speaker
Baochun Li received his B.Engr. degree from the Department of Computer Science and Technology, Tsinghua University, China, in 1995 and his M.S. and Ph.D. degrees from the Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, in 1997 and 2000. Since 2000, he has been with the Department of Electrical and Computer Engineering at the University of Toronto, where he is currently a Professor. He holds the Bell Canada Endowed Chair in Computer Engineering since August 2005. His research interests include cloud computing, distributed systems, datacenter networking, and wireless systems.

Prof. Li has co-authored more than 410 research papers, with a total of over 21000 citations, an H-index of 83 and an i10-index of 280, according to Google Scholar Citations. He was the recipient of the IEEE Communications Society Leonard G. Abraham Award in the Field of Communications Systems in 2000. In 2009, he was a recipient of the Multimedia Communications Best Paper Award from the IEEE Communications Society, and a recipient of the University of Toronto McLean Award. He is a member of ACM and a Fellow of IEEE.