Department of Computer Science
香港城市大學
City University of Hong Kong

COMPUTER SCIENCE COLLOQUIUM

Making Serverless Compu3ng Efficient and Scalable

SPEAKER **Dr. K. K. Ramakrishnan**

Distinguished Professor
Computer Science and Engineering
University of California, Riverside

DATE 22 Sep, 2025 (Mon)
TIME 9:30 AM - 10:30 AM
VENUE CS Seminar Room, Y6405, 6th Floor, Yellow Zone, Yeung Kin Man Academic Building, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong

ABSTRACT

We have been working on Serverless Computing, a fast-growing area of cloud computing, for a few years now. Serverless computing simplifies the development, deployment, and automated management of modular software functions, exploiting the microservices paradigm while promising efficient, low-cost computing for users. Serverless computing has evolved by integrating multiple existing cloud computing software components, thus quickly offering its capabilities in the cloud, while possibly sacrificing efficiency. We developed SPRIGHT, to exploit event-driven shared memory processing to improve the serverless data plane's efficiency and scalability by eliminating data copies and avoiding unnecessary protocol processing and serialization-deserialization overheads. The extended Berkeley Packet Filter (eBPF) enabled us to create true event-driven processing, especially to replace the typical heavyweight sidecar proxy used in serverless computing. After a brief recap of SPRIGHT, we will look at another aspect of our work, SURE, a unikernel-based serverless framework, to address the need for a secure data plane and fast-function startup. SURE's data plane supports distributed zero-copy communication via the seamless interaction between zero-copy protocol stack (Z-stack) and local shared memory processing. We leverage Intel's Memory Protection Keys (MPK) as a lightweight capability to ensure safe access to the shared memory data plane and the use of a library-based sidecar. We have recently extended the shared memory of our serverless computing framework to multiple nodes using RDMA to achieve even higher performance and lower latency.

BIOGRAPHY

Dr. K. K. Ramakrishnan is a Distinguished Professor of Computer Science and Engineering at the University of California, Riverside. Previously, he was a Distinguished Member of Technical Staff at AT&T Labs-Research. He joined AT&T Bell Labs in 1994 and was with AT&T Labs-Research since its inception in 1996. Prior to 1994, he was a Technical Director and Consulting Engineer in Networking at Digital Equipment Corporation. Between 2000 and 2002, he was at TeraOptic Networks, Inc., as Founder and Vice President. Dr. Ramakrishnan is an ACM Fellow, IEEE Fellow, and an AT&T Fellow, recognized for his fundamental contributions to communication networks, congestion control, traffic management, VPN services, and a lasting impact on AT&T and the industry. His work on the "DECbit" congestion avoidance protocol received the ACM Sigcomm Test of Time Paper Award in 2006. K. K. received the ACM Sigcomm Lifetime Achievement Award in 2024. He has published over 300 papers and has over 180 patents issued in his name. K. K. received his MTech from the Indian Institute of Science (1978, recently recognized as one of IISc's Distinguished Alumni), MS (1981), and Ph.D. (1983) in Computer Science from the University of Maryland, College Park, USA.

All are welcome!

In case of questions, please contact Prof JIA Xiaohua at csjia@cityu.edu.hk, or visit the CS Departmental Seminar Web at https://www.cs.cityu.edu.hk/events/cs-seminars/recent-cs-colloquiums.