THE UNIVERSITY OF ALABAMA

DEPARTMENT OF COMPUTER SCIENCE

Colloquium Talk

The Millibottleneck Theory of Millisecond-Scale Performance Bugs and Its Experimental Verification

Dr. Calton Pu

Professor and John P. Imlay, Jr. Chair in Software

School of Computer Science, Georgia Institute of Technology

Friday, February 9, 2018 @ 2:00 - 3:00pm

NERC 1012

Abstract

Web-facing applications have complex deployment dependencies and stringent quality of service requirements, e.g., 99.9% of requests with response time within 0.5 seconds. However, despite continued efforts by industry and academic researchers, the latency long tail problem, where a non-trivial fraction of Very Long Response Time (VLRT) requests return after a few seconds, remains a serious research and practical challenge. Latency long tail happens even when the system utilization is still very far from saturation (e.g., 40 – 60% average CPU utilization). Using automated n-tier application benchmarks, we have reproduced several cases of these VLRT requests (e.g., due to JVM garbage collection and VM-based application consolidation), caused by millibottlenecks (resource saturations that last only tens to hundreds of milliseconds).

The Millbottleneck Theory explains these VLRT requests in a model with two parts. First, a resource millibottleneck is created, and it propagates through a chain of dependencies among system components, accumulating queuing effects that end in VLRT requests. We have released the MilliMonitor toolkit, which is capable of 50ms sampling periods and detailed event monitoring to enable the detection of millibottlenecks and the tracking of chain of dependencies that end in VLRT requests. We will describe a methodical approach to find new millibottlenecks and their chain of dependencies, so we can remove the sources of VLRT requests and improve overall system utilization in data centers while preserving high quality of service.

Biography

Calton Pu was born in Taiwan and grew up in Brazil. He received his PhD from University of Washington and served on the faculty of Columbia University and Oregon Graduate Institute. Currently, he is holding the position of Professor and John P. Imlay, Jr. Chair in Software in the College of Computing, Georgia Institute of Technology. He has worked on several projects in systems and database research. His contributions to systems research include program specialization and software feedback. His contributions to database research include extended transaction models and their implementation. His recent research has focused on automated system management in clouds (Elba project), information quality (e.g., spam processing), and big data in Internet of Things (GRAIT-DM project). He has collaborated extensively with scientists and industry researchers. He has published more than 70 journal papers and book chapters, 280 conference and refereed workshop papers. He served on more than 120 program committees, including the co-PC chairs of SRDS'95, ICDE'99, COOPIS'02, SRDS'03, DOA'07, DEBS'09, ICWS'10, CollaborateCom'11, ICAC'13, CLOUD'15, BigData Congress'16, CIC'16, and co-general chair of ICDE'97, CIKM'01, ICDE'06, DEPSA'07, CEAS'07, SCC'08, CollaborateCom'08, World Service Congress'11, CollaborateCom'12, IEEE CIC'15, and ICDCS'17. He is a Fellow of AAAS and IEEE.