DESIGN AND IMPLEMENTATION: DEALING WITH LARGE-SCALE BUG TRACES

Vijay K. Gurbani

Salvatore Loreto

As our networks, hosts, and services become complex, and indeed, as the number of networks, hosts, and services grow and shrink on demand through the magic of virtualization, uncovering software defects by analyzing packet traces are becoming even more important. A by-product of Big Data appears to be Big Traces as well.

Finding errant behavior of an application by analyzing the network traffic is a reasonable technique when the latent bugs do not show themselves during unit testing or integration testing conducted in a small-scale (read laboratory) network. Running applications on a large scale uncovers temporal and causality bugs that are hard to find otherwise. An attendant problem with large-scale traces is, well, the size of the trace itself. Analyzing software defects in network applications that routinely generate gigabytes of data per minute by reasoning on the produced network traffic is a challenging task.

In the article that follows, Ying-Dar *et al.* demonstrate how to reduce the size of the captured trace while ensuring that the downsized trace is representative of the original trace. The latter characteristic is important precisely because ignoring temporal and casual events in the downsized trace will render it useless for subsequent reasoning. Ying-Dar *et al.* present two downsizing algorithms: linear downsizing and binary downsizing. They study the efficacy of these algorithms on traces generated from a beta site at National Chiao Tung University in Taiwan.

BIOGRAPHIES

VIJAY K. GURBANI [M'98] (vkg@bell-labs.com) is a Distinguished Member of Technical Staff in wireless research at Bell Laboratories, Alcatel-Lucent. He holds a B.Sc. in computer science with a minor in mathematics and an M.Sc. in computer science, both from Bradley University, and a Ph.D. in computer science from Illinois Institute of Technology. His current work focuses on peer-to-peer networks, Internet multimedia session protocols, and anomaly detection in such protocols. He is the author of over 50 journal papers and conference proceedings, five books, and 16 IETF RFCs. He is currently Co-Chair of the Application Layer Traffic Optimization (ALTO) Working Group in the IETF, which is designing a protocol to enable abstract network information to be provided to applications through a well defined API. He holds four patents and has 10 applications pending with the U.S. Patent Office. He is a Senior Member of the ACM, and a member of the IEEE Computer Society and Usenix.

SALVATORE LORETO [M'01, SM'09] (salvatore.loreto@ieee.org) has 15 years of experience in a variety of information and communication companies, and has been working in networking and telecommunications since 1999. Currently, he works as a research scientist in the MultiMedia Technology section branch, which is part of the NomadicLab, at Ericsson Research Finland. He has made contributions in Internet transport protocols (e.g., TCP, SCTP), signal protocols (e.g., SIP, XMPP), VoIP, IP-telephony convergence, conferencing over IP, 3GPP IP Multimedia Subsystem (IMS), HTTP, and web technologies. He is also an active contributor to the IETF, where he has co-authored several RFCs and Internet drafts. Currently he is serving within the IETF as Co-Chair of the SIP Overload Control (soc), Application Area (appsawg), and Bidirectional or Server-Initiated HTTP (HyBi) Working Groups. For the IEEE Communications Society, he serves as a Design and Implementation Series Co-Editor and an Associate Technical Editor for *IEEE Communications Magazine*. He received an M.S. degree in engineer computer science and a Ph.D. degree in computer networking from Napoli University in 1999 and 2006, respectively.