UT Dallas Professor Shows a Promising Approach to a Critical Problem in Cloud Computing

August 27, 2012

Virtualization has enabled cloud computing, and has opened many opportunities to devise new virtual machine (VM) services for system administration, security, and reliability. But the biggest hurdle when implementing the VM services is the "semantic gap" problem. This problem can be described as follows: In regular software development, programmers have rich semantics such as APIs to use. However, the view at the VM Monitor (VMM) layer is too abstract. That is, at the VMM layer there is nothing but just raw bits and bytes. Thus, we often have to bridge the semantic gap while developing VMM services. This problem was first raised by Professor Peter Chen and Professor Brian Noble from University of Michigan in 2001.

It is often tedious, error-prone and time-consuming to bridge the semantic gap. Given the importance of this problem, several research groups from organizations such as VMware, Stanford, University of Michigan, Georgia Tech, University of Maryland, Purdue, and North Carolina State University have attempted to solve it. Their solutions range from purely manual processes to semi-automatic processes. Assistant Professor Zhiqiang Lin at The University of Texas at Dallas, together with his student Mr. Fu, has developed novel techniques that make the semantic-bridging entirely automatic.

"What we've done represents a significant advancement in virtualization technology, and it might change the daily practice for many virtualization services," says Dr. Lin. "Anyone interested in system virtualization will be very interested in our work."

Of particular application for this technique will be the virtual machine introspection (VMI) and memory forensics. VMI is a widely-used technique by cloud providers to inspect the state of the guest OS, such as detecting any intrusions. "By using our software, we automatically generate a number of VMI tools for free, whereas previously, software developers had to manually write such tools," Dr. Lin said. "Also, they can now natively develop the new VMI software and can fully use the rich semantic APIs offered by the Operating System". In other words, there will be no semantic gap when using their techniques to perform the VMI. Similarly in the critical areas of memory forensics, the investigators can now fully re-use the legacy memory state inspection tool to examine the memory without worrying about the semantic gap.

The foundations of their technique are based on the observation that a program is usually composed of code and data. A program running on one machine typically consumes the data within that machine. "Our techniques make the program consume the data in a second machine at the VMM layer without the awareness that the data is actually from other machines. This is how we automatically bridge the semantic gap," added Dr. Lin.

"Fu and Lin have developed an interesting way to take existing code from a trusted guest operating system and automatically use it for virtual-machine introspection. The ability to leverage existing code goes a long way in solving the semantic gap problem inherent to many types of virtual machine services," said Dr. Peter Chen, Arthur F. Thurnau Professor of Electrical Engineering and Computer Science at the University of Michigan, who first posed the "semantic gap" problem in 2001.

"The UT Dallas technique for virtual machine introspection -- a dynamic monitoring techniques for guest operating systems that might be malware infected -- offers new and unprecedented opportunities for secure system operation," said Dr. Virgil Gligor, Professor of Computer Science and Director of Carnegie Mellon University's CYLAB. "Briefly, Dr. Lin's technique enables the design of new and novel tools in intrusion detection, malware analysis, dynamic monitoring of process execution, and memory forensics. Using tools based on his technique, anti-virus software companies and forensic investigators may no longer need to write new software for every new type of malware in a cloud environment, and they can reuse legacy binary code to do the introspection" he added.

"The work by Dr. Lin and his student is significant in cloud security. They have demonstrated a solution for a hard problem in virtualization – this solution automatically bridges the semantic-gap by reusing the legacy binary code. Once the semantic-gap is bridged, there will be many new opportunities for virtualization research," said Dr. Elisa Bertino, Professor of Computer Science at Purdue University and Director of Purdue's Center for Education and Research in Information Assurance and Security (CERIAS).

Dr. Lin and his student Mr. Fu recently presented their paper based on this breakthrough research titled *Space Traveling across VM: Automatically Bridging the Semantic Gap in Virtual Machine Introspection via Online Kernel Data Redirection* at the 33rd IEEE Symposium On Security And Privacy in San Francisco, California. "This is the top conference in cyber security," said Dr. Bhavani Thuraisingham, Executive Director of UT Dallas' Cyber Security Research and Education Center and the Louis A. Beecherl, Jr. I Distinguished Professor of Computer Science. "It is a major breakthrough that VMI developers no longer need to write any code to bridge the semantic gap by using the technology invented by Dr. Lin and his student Mr. Fu. This research has given us tremendous visibility among the cyber security research community around the world," added Dr. Thuraisingham.

For more information, contact Rhonda Walls at rhonda.walls@utdallas.edu or 972.883.2731.